WILEY BIOTECHNOLOGY ENCYCLOPEDIAS

ENCYCLOPEDIA OF ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

EDITED BY
THOMAS H. MURRAY
MAXWELL J. MEHLMAN

ENCYCLOPEDIA OF

ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

VOLUME 1

ENCYCLOPEDIA OF

ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

VOLUME 1

WILEY BIOTECHNOLOGY ENCYCLOPEDIAS

Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation
Edited by Michael C. Flickinger and Stephen W. Drew

Encyclopedia of Molecular Biology
Edited by Thomas E. Creighton

Encyclopedia of Cell Technology
Edited by Raymond E. Spier

Encyclopedia of Ethical, Legal, and Policy Issues in Biotechnology
Edited by Thomas H. Murray and Maxwell J. Mehlman

ENCYCLOPEDIA OF ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

EDITORIAL BOARD

ENCYCLOPEDIA OF

ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

VOLUME 1

Thomas H. Murray
The Hastings Center
Garrison, New York

Maxwell J. Mehlman
Case Western Reserve University
Cleveland, Ohio

A Wiley-Interscience Publication
John Wiley & Sons, Inc.
New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

This book is printed on acid-free paper. ♾

Published simultaneously in Canada.

For ordering and customer service, call 1-800-CALL-WILEY.

Library of Congress Cataloging in Publication Data:

Murray, Thomas H. (Thomas Harold), 1926–

Encyclopedia of ethical, legal, and policy issues in biotechnology / Thomas Murray, Maxwell J. Mehlman.

p. cm.

Includes index.

ISBN 0-471-17612-5 (cloth : alk. paper)

1. Biotechnology — Moral and ethical aspects — Encyclopedias. 2. Biotechnology industries — Law and legislation — Encyclopedias. 3. Biotechnology — government policy — United States — Encyclopedias. I. Mehlman, Maxwell J. II. Title.

TP248.16.M87 2000

174′.96606 — dc21 00-021383

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

PREFACE

The Wiley Biotechnology Encyclopedias, composed of the *Encyclopedia of Molecular Biology*; the *Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation*; the *Encyclopedia of Cell Technology*; and the *Encyclopedia of Ethical, Legal, and Policy Issues in Biotechnology* cover very broadly four major contemporary themes in biotechnology. The series comes at a fascinating time in that, as we move into the twenty-first century, the discipline of biotechnology is undergoing striking paradigm changes.

Biotechnology is now beginning to be viewed as an informational science. In a simplistic sense there are three types of biological information. First, there is the digital or linear information of our chromosomes and genes with the four-letter alphabet composed of G, C, A, and T (the bases guanine, cytosine, adenine, and thymine). Variation in the order of these letters in the digital strings of our chromosomes or our expressed genes (or mRNAs) generates information of several distinct types: genes, regulatory machinery, and information that enables chromosomes to carry out their tasks as informational organelles (e.g., centromeric and telomeric sequences).

Second, there is the three-dimensional information of proteins, the molecular machines of life. Proteins are strings of amino acids employing a 20-letter alphabet. Proteins pose four technical challenges: (*1*) Proteins are synthesized as linear strings and fold into precise three-dimensional structures as dictated by the order of amino acid residues in the string. Can we formulate the rules for protein folding to predict three-dimensional structure from primary amino acid sequence? The identification and comparative analysis of all human and model organism (bacteria, yeast, nematode, fly, mouse, etc.) genes and proteins will eventually lead to a lexicon of motifs that are the building block components of genes and proteins. These motifs will greatly constrain the shape space that computational algorithms must search to successfully correlate primary amino acid sequence with the correct three-dimensional shapes. The protein-folding problem will probably be solved within the next 10–15 years. (*2*) Can we predict protein function from knowledge of the three-dimensional structure? Once again the lexicon of motifs with their functional as well as structural correlations will play a critical role in solving this problem. (*3*) How do the myriad of chemical modifications of proteins (e.g., phosphorylation, acetylation, etc.) alter their structures and modify their functions? The mass spectrometer will play a key role in identifying secondary modifications. (*4*) How do proteins interact with one another and/or with other macromolecules to form complex molecular machines (e.g., the ribosomal subunits)? If these functional complexes can be isolated, the mass spectrometer, coupled with a knowledge of all protein sequences that can be derived from the complete genomic sequence of the organism, will serve as a powerful tool for identifying all the components of complex molecular machines.

The third type of biological information arises from complex biological systems and networks. Systems information is four dimensional because it varies with time. For example, the human brain has 1,012 neurons making approximately 1,015 connections. From this network arise systems properties such as memory, consciousness, and the ability to learn. The important point is that systems properties cannot be understood from studying the network elements (e.g., neurons) one at a time; rather the collective behavior of the elements needs to be studied. To study most biological systems, three issues need to be stressed. First, most biological systems are too complex to study directly, therefore they must be divided into tractable subsystems whose properties in part reflect those of the system. These subsystems must be sufficiently small to analyze all their elements and connections. Second, high-throughput analytic or global tools are required for studying many systems elements at one time (see later). Finally, the systems information needs to be modeled mathematically before systems properties can be predicted and ultimately understood. This will require recruiting computer scientists and applied mathematicians into biology—just as the attempts to decipher the information of complete genomes and the protein folding and structure/function problems have required the recruitment of computational scientists.

I would be remiss not to point out that there are many other molecules that generate biological information: amino acids, carbohydrates, lipids, and so forth. These too must be studied in the context of their specific structures and specific functions.

The deciphering and manipulation of these various types of biological information represent an enormous technical challenge for biotechnology. Yet major new and powerful tools for doing so are emerging.

One class of tools for deciphering biological information is termed high-throughput analytic or global tools. These tools can be used to study many genes or chromosome features (genomics), many proteins (proteomics), or many cells rapidly: large-scale DNA sequencing, genomewide genetic mapping, cDNA or oligonucleotide arrays, two-dimensional gel electrophoresis and other global protein separation technologies, mass spectrometric analysis of proteins and protein fragments, multiparameter, high-throughput cell and chromosome sorting, and high-throughput phenotypic assays.

A second approach to the deciphering and manipulating of biological information centers around combinatorial strategies. The basic idea is to synthesize an informational string (DNA fragments, RNA fragments, protein fragments, antibody combining sites, etc.) using all combinations of the basic letters of the corresponding alphabet, thus creating many different shapes that can be used to activate, inhibit, or complement the biological functions of designated three-dimensional shapes (e.g., a molecule in a signal transduction pathway). The power of combinational chemistry is just beginning to be appreciated.

A critical approach to deciphering biological information will ultimately be the ability to visualize the functioning of genes, proteins, cells, and other informational elements within living organisms (in vivo informational imaging).

Finally, there are the computational tools required to collect, store, analyze, model, and ultimately distribute the various types of biological information. The creation presents a challenge comparable to that of developing new instrumentation and new chemistries. Once again this means recruiting computer scientists and applied mathematicians to biology. The biggest challenge in this regard is the language barriers that separate different scientific disciplines. Teaching biology as an informational science has been a very effective means for breeching these barriers.

The challenge is, of course, to decipher various types of biological information and then be able to use this information to manipulate genes, proteins, cells, and informational pathways in living organisms to eliminate or prevent disease, produce higher-yield crops, or increase the productivity of animals for meat and other foods.

Biotechnology and its applications raise a host of social, ethical, and legal questions, for example, genetic privacy, germline genetic engineering, cloning of animals, genes that influence behavior, cost of therapeutic drugs generated by biotechnology, animal rights, and the nature and control of intellectual property.

Clearly, the challenge is to educate society so that each citizen can thoughtfully and rationally deal with these issues, for ultimately society dictates the resources and regulations that circumscribe the development and practice of biotechnology. Ultimately, I feel enormous responsibility rests with scientists to inform and educate society about the challenges as well as the opportunities arising from biotechnology. These are critical issues for biotechnology that are developed in detail in the *Encyclopedia of Ethical, Legal, and Policy Issues in Biotechnology*.

The view that biotechnology is an informational science pervades virtually every aspect of this science, including discovery, reduction to practice, and societal concerns. These Encyclopedias of Biotechnology reinforce the emerging informational paradigm change that is powerfully positioning science as we move into the twenty-first century to more effectively decipher and manipulate for humankind's benefit the biological information of relevant living organisms.

Leroy Hood
University of Washington

LIST OF ARTICLES

CONTRIBUTORS

Terrence F. Ackerman, *University of Tennessee, Memphis, Tennessee,* Human Subjects Research, Ethics, Compensation of Subjects for Injury

Sheri Alpert, *University of Notre Dame, Notre Dame, Indiana,* Genetic Information, Ethics, Ethical Issues in Tissue Banking and Human Subject Research in Stored Tissues

Mary R. Anderlik, *University of Houston, Houston, Texas,* Genetic Information, Legal, Genetic Privacy Laws; Genetic Information, Legal, Genetics and the Americans with Disabilities Act

Mary Jane Angelo, *St. John's River Water Management District, Palatka, Florida,* Agricultural Biotechnology, Law, and EPA Regulation

Adrienne Asch, *Wellesley College, Wellesley, Massachusetts,* Reproduction, Ethics, Prenatal Testing, and the Disability Rights Critique

Richard E. Ashcroft, *Imperial College, London, England,* International Aspects: National Profiles, United Kingdom

Patricia Backlar, *Portland State University, Oregon Health Sciences University, Portland, Oregon,* Human Subjects Research, Ethics, Research on Vulnerable Populations

Dianne M. Bartels, *University of Minnesota, Center for Bioethics, Minneapolis, Minnesota,* Reproduction, Ethics, Sex Selection

Françoise Bieri, *Swiss Priority Programme Biotechnology of the Swiss National Science Foundation, Basel, Switzerland,* International Aspects: National Profiles, Switzerland

Barbara Bowles Biesecker, *National Human Genome Research Institute, Bethesda, Maryland,* Reproduction, Ethics, the Ethics of Reproductive Genetic Counseling: Nondirectiveness

Michael R. Bleavins, *Warner-Lambert Company, University of Michigan Medical School, Ann Arbor, Michigan,* Pharmacogenetics

David Blumenthal, *Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts,* Academic Industry Relationships in Biotechnology, Overview

Mary Anne Bobinski, *University of Houston Law Center, Houston, Texas,* Genetic Information, Legal, ERISA Preemption, and HIPAA Protection

Robert A. Bohrer, *California Western School of Law, San Diego, California,* Federal Regulation of Biotechnology Products for Human Use, FDA, Orphan Drug Act

Andrea L. Bonnicksen, *Northern Illinois University, DeKalb, Illinois,* Gene Therapy, Ethics, and International Perspectives

Jeffrey R. Botkin, *University of Utah School of Medicine, Salt Lake City, Utah,* Reproduction, Law, Wrongful Birth, and Wrongful Life Actions

Dan W. Brock, *Brown University, Providence, Rhode Island,* Cloning, Ethics

Paul Brodwin, *University of Wisconsin, Milwaukee, Wisconsin,* Professional Power and the Cultural Meanings of Biotechnology

Baruch Brody, *Baylor College of Medicine, Houston, Texas,* Human Subjects Research, Law, FDA Rules

William P. Browne, *Central Michigan University, Mount Pleasant, Michigan,* Agricultural Biotechnology, Law, and Social Impacts of Agricultural Biotechnology

Jeffrey Burkhardt, *University of Florida, Gainesville, Florida,* Agricultural Biotechnology, Ethics, Family Farms, and Industrialization

David L. Cameron, *Willamette University College of Law, Salem, Oregon,* Scientific Research, Policy, Tax Treatment of Research and Development

Courtney S. Campbell, *Oregon State University, Corvallis, Oregon,* Patents and Licensing, Ethics, Moral Status of Human Tissues: Sale, Abandonment or Gift; Religious Views on Biotechnology, Protestant

Eric G. Campbell, *Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts,* Academic Industry Relationships in Biotechnology, Overview

Benjamin Capps, *University of Bristol, Bristol, England,* International Aspects: National Profiles, United Kingdom

Catherine Chabert, *Alain Bensoussan—Avocats, Lyon, France,* International Aspects: National Profiles, France

R. Cole-Turner, *Pittsburgh Theological Seminary, Pittsburgh, Pennsylvania,* Patents and Licensing, Ethics, Organizations with Prominent Positions on Gene Patenting

Robert Mullan Cook-Deegan, *Georgetown University, Annapolis, Maryland,* Medical Biotechnology, United States Policies Influencing its Development

George C. Cunningham, *State of California-Department of Health Services, Berkeley, California,* Genetic Information, Legal, Regulating Genetic Services

Elliot N. Dorff, *University of Judaism, Bel Air, California,* Religious Views on Biotechnology, Jewish

Daniel W. Drell, *U.S. Department of Energy, Washington, District of Columbia,* Informing Federal Policy on Biotechnology: Executive Branch, Department of Energy

Elisa Eiseman, *RAND Science and Technology Policy Institute, Washington, District of Columbia,* Cloning, Policy Issues

Susan S. Ellenberg, *U.S. Food and Drug Administration, Washington, District of Columbia,* Human Subjects Research, Ethics, Stopping Rules for Randomized Clinical Trials

Lisa M. Gehrke, *Nilles & Nilles, S.C., Milwaukee, Wisconsin,* Ownership of Human Biological Material

Oreste Ghisalba, *Swiss Priority Programme Biotechnology of the Swiss National Science Foundation, Basel, Switzerland,* International Aspects: National Profiles, Switzerland

Paul Giannelli, *Case Western Reserve University, School of Law, Cleveland, Ohio,* Genetic Information, Law, Legal Issues in Law Enforcement DNA Databanks

Henry T. Greely, *Stanford University Law School, Stanford, California,* Human Genome Diversity Project

Ronald Green, *Darmouth College, Hanover, New Hampshire,* Human Subjects Research, Ethics, Research on Human Embryos

Louis M. Guenin, *Harvard Medical School, Boston, Massachusetts,* Patents, Ethics, Human Life Forms

Kathi E. Hanna, *Science and Health Policy Consultant, Prince Frederick, Maryland,* Federal Policy Making for Biotechnology, Executive Branch, National Human Genome Research Institute; Cloning, Overview of Human Cloning

Mats G. Hansson, *Uppsala University, Uppsala, Sweden,* International Aspects: National Profiles, Scandinavia

Sarah C. Hensel, *Mayo Clinic, Rochester, Minnesota,* Animal Medical Biotechnology, Policy, Would Transgenic Animals Solve the Organ Shortage Problem?

Göran Hermerén, *Lund University, Lund, Sweden,* Patents and Licensing, Ethics, International Controversies

Sharona Hoffman, *Case Western Reserve University, School of Law, Cleveland, Ohio,* Genetic Information, Law, Legal Issues in Law Enforcement DNA Databanks

Rachelle D. Hollander, *National Science Foundation, Arlington, Virginia,* Scientific Research, Ethics, Values in Science

Cynthia M. Ho, *Loyola University of Chicago School of Law, Chicago, Illinois,* International Intellectual Property Issues for Biotechnology

Anny Huang, *Wachtell, Lipton, Rosen, and Katz, New York, New York,* Genetic Information, Legal, FDA Regulation of Genetic Testing

Richard Huxtable, *University of Bristol, Bristol, England,* International Aspects: National Profiles, United Kingdom

Gail H. Javitt, *Johns Hopkins School of Hygiene and, Public Health, Baltimore, Maryland,* Gene Therapy, Law and FDA Role in Regulation

Eric T. Juengst, *Case Western Reserve University, Cleveland, Ohio,* Genetic Information, Ethics, Family Issues; Human Enhancement Uses of Biotechnology, Ethics, the Ethics of Enhancement

Marshall B. Kapp, *Wright State University School of Medicine, Dayton, Ohio,* Human Subjects Research, Law, Law of Informed Consent

Othmar Käppeli, *Swiss Priority Programme Biotechnology of the Swiss National Science Foundation, Basel, Switzerland,* International Aspects: National Profiles, Switzerland

Nancy E. Kass, *John Hopkins University, Baltimore, Maryland,* Human Subjects Research, Ethics, Informed Consent in Research

Dana Katz, *University of Pennsylvania, Philadelphia, Pennsylvania,* Patents and Licensing, Policy, Patenting of Inventions Developed with Public Funds

Matthias Kettner, *Cultural Studies Center of Northrhein-Westfalia (KWI), Essen, Germany,* International Aspects: National Profiles, Germany

John Kleinig, *City University of New York, New York, New York,* Research on Animals, Ethics, Principles Governing Research on Animals

Bartha Maria Knoppers, *University of Montreal, Montreal, Canada,* Human Subjects Research, Ethics, and International Codes on Genetic Research

Loretta M. Kopelman, *Brody School of Medicine at East Carolina University, Greenville, North Carolina,* Human Subjects Research, Ethics, and Research on Children

William Lacy, *University of California, Davis, California,* Agricultural Biotechnology, Socioeconomic Issues, and the Fourth Criterion

Robert P. Lawry, *Case Western Reserve University, Cleveland, Ohio,* University-Industry Research Relationships, Ethics, Conflict of Interest

Susan K. Lehnhardt, *Morrison & Foerster LLP, New York, New York,* Transferring Innovations from Academic Research Institutions to Industry: Overview

Josh Lerner, *Harvard University and National Bureau of Economic Research, Boston, Massachusetts,* Strategic Alliances and Technology Licensing in Biotechnology

Bonnie S. Leroy, *University of Minnesota, Institute of Human Genetics, Minneapolis, Minnesota,* Reproduction, Ethics, Sex Selection

Robert J. Levine, *Yale University School of Medical, New Haven, Connecticut,* Human Subjects Research, Ethics, Principles Governing Research with Human Subjects; Human Subjects Research, Ethics, Stopping Rules for Randomized Clinical Trials

Susan Lewis, *Case Western Reserve University, University Hospitals of Cleveland, Cleveland, Ohio,* Human Subjects Research, Ethics, Family, and Pedigree Studies

Darryl R.J. Macer, *Institute of Biological Sciences, University of Tsukuba, Tsukuba Science City, Japan,* International Aspects: National Profiles, Japan

David R. Mackenzie, *University of Maryland, College Park, Maryland,* Agricultural Biotechnology, Law, APHIS Regulation

David Magnus, *University of Pennsylvania, Philadelphia, Pennsylvania,* Eugenics, Ethics

Michael J. Malinowski, *Widener University School of Law, Wilmington, Delaware,* FDA Regulation of Biotechnology Products for Human Use

Wendy K. Mariner, *Boston University, Boston, Massachusetts,* Human Subjects Research, Law, Common Law of Human Experimentation

Cynthia R. Marling, *Ohio University, Athens, Ohio,* Human Enhancement Uses of Biotechnology, Ethics, Cognitive Enhancement

Anne L. Matthews, *Case Western Reserve University, Cleveland, Ohio,* Genetic Counseling

Charles R. McCarthy, *Kennedy Institute of Ethics, Georgetown University, Washington, District of Columbia,* Human Subjects Research, Law, HHS Rules

Jean E. McEwen, *National Human Genome Research Institute, Bethesda, Maryland,* Genetic Information, Ethics, and Information Relating to Biological Parenthood

Glenn McGee, *University of Pennsylvania, Philadelphia, Pennsylvania,* Eugenics, Ethics

Joseph D. McInerney, *Foundation for Genetic Education and Counseling, Baltimore, Maryland,* Education and Training, Public Education about Genetic Technology

Gerald P. McKenny, *Rice University, Houston, Texas,* Gene Therapy, Ethics, Religious Perspectives; Human Enhancement Uses of Biotechnology, Ethics, Therapy vs. Enhancement

Maxwell J. Mehlman, *Case Western Reserve University, Cleveland, Ohio,* Human Enhancement Uses of Biotechnology, Law, Genetic Enhancement, and the Regulation of Acquired Genetic Advantages

Richard A. Merrill, *University of Virginia, Charlottesville, Virginia,* Gene Therapy, Law and FDA Role in Regulation

Jon F. Merz, *University of Pennsylvania, Philadelphia, Pennsylvania,* Patents and Licensing, Policy, Patenting of Inventions Developed with Public Funds

Henry I. Miller, *Stanford University, Stanford, California,* Agricultural Biotechnology, Law, and Food Biotechnology Regulation

Barbara Mishkin, *Hogan Hartson L.L.P., Washington, District of Columbia,* Scientific Research, Law, and Penalties for Scientific Misconduct

Ronald Y. Nakasone, *Graduate Theological Union, Berkeley, California,* Religious Views on Biotechnology, Buddhism

Curtis Naser, *Fairfield University, Fairfield, Connecticut,* Genetic Information, Ethics, Ethical Issues in Tissue Banking and Human Subject Research in Stored Tissues

Dorothy Nelkin, *New York University, New York, New York,* Media Coverage of Biotechnology

Robert M. Nelson, *University of Pennsylvania, Philadelphia, Pennsylvania,* Professional Power and the Cultural Meanings of Biotechnology; Gene Therapy, Ethics, Germ Cell Gene Transfer

Marion Nestle, *New York University, New York, New York,* Agricultural Biotechnology, Policy, and Nutrition

Gilbert S. Omenn, *University of Michigan, Ann Arbor, Michigan,* Federal Policy Making for Biotechnology, Executive Branch, Office of Management and Budget

F. Barbara Orlans, *Kennedy Institute of Ethics, Georgetown University, Washington, District of Columbia,* Research on Animals, Law, Legislative, and Welfare Issues in the Use of Animals for Genetic Engineering and Xenotransplantation

Stuart M. Pape, *Patton Boggs, LLP, Washington, District of Columbia,* Animal Biotechnology, Law, FDA Regulation of Genetically Modified Animals for Human Food Use

Erik Parens, *The Hastings Center, Garrison, New York,* Reproduction, Ethics, Prenatal Testing, and the Disability Rights Critique

Kayhan P. Parsi, *Institute for Ethics, American Medical Association, Chicago, Illinois,* Reproduction, Law, Regulation of Reproductive Technologies

Rebecca D. Pentz, *M.D. Anderson Cancer Center, Houston, Texas,* Genetic Information, Legal, Genetic Privacy Laws

Jeffrey L. Platt, *Mayo Clinic, Rochester, Minnesota,* Animal Medical Biotechnology, Policy, Would Transgenic Animals Solve the Organ Shortage Problem?

Cynthia Powell, *University of North Carolina, Chapel Hill, North Carolina,* Reproduction, Ethics, Prenatal Testing, and the Disability Rights Critique

Madison Powers, *Georgetown University, Washington, District of Columbia,* Genetic Information, Ethics, Privacy and Confidentiality: Overview

Kimberly A. Quaid, *Indiana University School of Medicine, Indianapolis, Indiana,* Genetic Information, Ethics, Informed Consent to Testing and Screening

Philip R. Reilly, *Shriver Center for Mental Retardation, Inc., Waltham, Massachusetts,* Eugenics, Ethics, Sterilization Laws

Herbert Reutimann, *Swiss Priority Programme Biotechnology of the Swiss National Science Foundation, Basel, Switzerland,* International Aspects: National Profiles, Switzerland

Laura F. Rothstein, *Louis D. Brandeis School of Law, University of Louisville, Louisville, Kentucky,* Reproduction, Law, Is Infertility a Disability?

Maria Ruano, *Alain Bensoussan—Avocats, Lyon, France,* International Aspects: National Profiles, France

Paul D. Rubin, *Patton Boggs, LLP, Washington, District of Columbia,* Animal Biotechnology, Law, FDA Regulation of Genetically Modified Animals for Human Food Use

Lilly-Marlene Russow, *Purdue University, West Lafayette, Indiana,* Research on Animals, Ethics, and the Moral Status of Animals

Kenneth J. Ryan, *Harvard Medical School, Boston, Massachusetts,* Scientific Research, Ethics, Scientific Misconduct

Ani B. Satz, *Princeton University, Princeton, New Jersey, University of Michigan Law School, Ann Arbor, Michigan,* Disability and Biotechnology

Jennifer Scott, *Case Western Reserve University, University Hospitals of Cleveland, Cleveland, Ohio,* Human Subjects Research, Ethics, Family, and Pedigree Studies

Harold T. Shapiro, *Princeton University, Princeton, New Jersey,* Federal Policy Making for Biotechnology, Executive Branch, National Bioethics Advisory Commission

Michael H. Shapiro, *University of Southern California Law School, Los Angeles, California,* Human Enhancement Uses of Biotechnology, Policy, Technological Enhancement, of Human Equality

Michael M. Shi, *Warner-Lambert Company, University of Michigan Medical School, Ann Arbor, Michigan,* Pharmacogenetics

Anita Silvers, *San Francisco State University, San Francisco, California,* Disability and Biotechnology

Rivers Singleton, Jr., *University of Delaware, Newark, Delaware,* Transgenic Animals: An Overview

Dominique Sprumont, *University of Neuchâtel, Neuchâtel, Switzerland,* Human Subjects Research, Ethics, and International Codes on Genetic Research

Bonnie Steinbock, *SUNY Albany, Albany, New York,* Reproduction, Ethics, Moral Status of the Fetus

Michele Svatos, *Iowa State University, Ames, Iowa,* Patents and Licensing, Ethics, Ownership of Animal, and Plant Genes

Carol A. Tauer, *College of St. Catherine, St. Paul, Minnesota,* Gene Therapy, Ethics, Gene Therapy for Fetuses and Embryos; Human Enhancement Uses of Biotechnology, Ethics, Human Growth Hormone

Paul B. Thompson, *Purdue University, West Lafayette, Indiana,* Agricultural Biotechnology, Ethics, Food Safety, Risk, and Individual Consent

Françoise Touraine-Moulin, *Immunology Laboratory—Hôpital Neurologique Pierre Wertheimer, Lyon, France,* International Aspects: National Profiles, France

Andrew Trew, *John Carroll University, Cleveland, Ohio,* Animal Medical Biotechnology, Legal, Laws and Regulations Governing Animals as Sources of Human Organs

Gary Varner, *Texas A&M University, College Station, Texas,* Cloning, Overview of Animal Cloning

Marissa Vicari, *La Tour de Salvagny, Lyon, France,* International Aspects: National Profiles, France

Robert Wachbroit, *University of Maryland, College Park, Maryland,* Genetic Determinism, Genetic Reductionism, and Genetic Essentialism; Human Enhancement Uses of Biotechnology: Overview

Wendy E. Wagner, *Case Western Reserve University, Cleveland, Ohio,* Federal Policy Making for Biotechnology, Congress, EPA; Genetic Information, Law, Legal Issues in Law Enforcement DNA Databanks

LeRoy Walters, *Kennedy Institute of Ethics, Georgetown University, Washington, District of Columbia,* Federal Policy Making for Biotechnology, Executive Branch, RAC; Gene Therapy: Overview

David T. Wasserman, *University of Maryland, College Park, Maryland,* Behavioral Genetics, Human

Charles Weiner, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Recombinant DNA, Policy, Asilomar Conference

Dorothy C. Wertz, *The Shriver Center, Waltham, Massachusetts,* Public Perceptions: Surveys of Attitudes Toward Biotechnology

Peter J. Whitehouse, *Case Western Reserve University, Cleveland, Ohio,* Human Enhancement Uses of Biotechnology, Ethics, Cognitive Enhancement

Georgia L. Wiesner, *Case Western Reserve University, University Hospitals of Cleveland, Cleveland, Ohio,* Human Subjects Research, Ethics, Family, and Pedigree Studies

Michael D. Witt, *Technology Enterprise Development Company, Troy, Michigan,* Transferring Innovations from Academic Research Institutions to Industry: Overview

Nelson A. Wivel, *University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania,* Gene Therapy, Ethics, Somatic Cell Gene Therapy

Audra Wolfe, *University of Pennsylvania, Philadelphia, Pennsylvania,* Federal Policy Making for Biotechnology, Executive Branch, ELSI

INTRODUCTION

If future generations mark the third millennium as the age of biotechnology, they may well regard the dawn of the twenty-first century as its birth. The mapping and sequencing of the human genome is almost complete. Products made with biotechnology, ranging from drugs to pest-resistant crops, are becoming commonplace. Human gene therapy is poised to make clinically significant strides.

In the midst of these astounding developments, a growing cadre of scientists and scholars are struggling to understand the profound ethical, legal, and social implications. Never before has humanity been able so directly to manipulate the code of life. The result is both opportunity and danger on an unprecedented scale. How safe is biotechnology? In what ways will genetic information affect our conceptions of who we are and our relationships with each other? Will the promise of eradicating genetic disease lead us to take untoward experimental risks, or to impose these risks on vulnerable subjects? Can democracy exist in a society in which only the wealthy obtain access to new genetic advances? At what point does a genetically modified human being cease to be a member of the human species?

Our objective in producing the *Encyclopedia of Ethical, Legal, and Policy Issues in Biotechnology* is to bring together the best minds to describe these issues, analyze their implications, and present public policy options. The Encyclopedia contains 112 entries arranged in broad alphabetical order. Related entries are cross-referenced. Each entry includes a list of sources. Virtually all of the entries have undergone peer review by two independent reviewers. For the most part, we have asked authors to be objective. Some entries, however, reflect partisan viewpoints due to their subject matter and the background of their authors. We hope we have made sure that, whenever this occurs, it is obvious to the reader.

The potential audience for this Encyclopedia is extremely broad, ranging from individuals with substantial knowledge and experience in these fields to those just embarking on their journey of understanding. We have endeavored to make all of the entries useful to the former while still accessible to the latter. Yet many entries address complex, technically demanding subjects, and we apologize to readers who find specific entries either too elementary or too abstruse. In addition, completing a project of this scope takes time. We recognize that, in a field as dynamic as biotechnology, descriptions of developments in science, ethics, law, and public policy rapidly become out-of-date. We have undertaken to make entries as current as possible.

Our ultimate goal has been to make this a comprehensive reference work. We began by forming an advisory board of renowned experts, headed by David Blumenthal. With their assistance, we created an exhaustive list of topics and identified potential contributors. We aimed for learned, highly respected authors, individuals who are actively involved in their fields and consequently in great demand. It was not always possible to engage their participation. As a result, there are gaps in coverage, which we mention here to dispel the notion that we simply failed to identify important topics for inclusion. For example, while individual entries describe a number of key government agencies and offices that affect biotechnology policy, we were unable to secure entries for some other government agencies, government offices, industry groups, and interest groups. We include profiles of a number of countries with extensive involvement in biotechnology, but were unable to obtain profiles for some of the countries on our list. The Encyclopedia contains discussions of the views of a number of religions toward biotechnology, but we were unable to obtain discussions for Islam or Roman Catholicism. Moreover, we originally intended to include a separate entry on each topic from an ethical, legal, and public policy perspective. This was not always possible. Nevertheless, we feel that, for the most part, entries that we were able to obtain from one or two of the perspectives provide adequate coverage of the major issues from the other viewpoints.

We wish to thank the authors, the members of the advisory board, our editors at Wiley, our staffs, and our families for their dedication and support.

Thomas H. Murray
Maxwell J. Mehlman

A

ACADEMIC INDUSTRY RELATIONSHIPS IN BIOTECHNOLOGY, OVERVIEW

David Blumenthal
Eric G. Campbell
Massachusetts General Hospital
Harvard Medical School
Boston, Massachusetts

OUTLINE

INTRODUCTION

Relationships between academic institutions and industries have become central to the biotechnology enterprise, and indeed, to all the life sciences in the United States. Academic–industry relationships (AIRs) in biotechnology are thought to serve a variety of purposes for participating academic institutions (universities, their medical schools, and their associated clinical teaching and research facilities and their faculty), for industries, and for the larger society. It is widely believed that AIRs facilitate the transfer of new knowledge from the academic to the industrial sector, and thereby the application of that knowledge to the practical needs of human beings in this country and around the globe. Funds from industry and from the commercial sales of intellectual properties owned by academic institutions support research and training in those institutions, and thus may spur the development of new knowledge and young investigators needed by both universities and industries. Industry, in turn, benefits not only from the profits realized from transferred academic intellectual property but also from increased productivity in its internal research resulting from enhanced opportunities to recruit talented scientists from academia and from exposure of its investigators to ongoing academic research.

The apparent growth of AIRs has also generated ongoing concerns about their risks. The greatest potential risks affect the academic enterprise. There are fears that AIRs will retard scientific progress through a number of effects: by involving the nation's most talented academic investigators in commercially relevant work, thus distracting them from the pursuit of fundamental questions whose answers will set the stage for the next biological revolution; by promoting secrecy in academic science, which will undermine scientific exchange that is essential to optimal progress in the life sciences; by involving young investigators in commercial projects of lesser scientific interest and import, and thereby compromising the quality of their training. Industrial partners of AIRs, of course, also face potential downsides, though these are primarily business risks of a type that is routine in any for-profit enterprise. Academic institutions may turn out to unproductive partners either because cultural differences between universities and industries cannot be successfully bridged, or because academic work produces little commercializable intellectual property, or because garrulous academics prove unable to protect industrial secrets until they can be commercialized.

Given the stakes involved, it is not surprising that AIRs in biotechnology and the life sciences generally continue to attract considerable attention in the popular press and to be the subject of both praise and deprecation (1–4). This article reviews the status of AIRs in biotechnology at the close of the twentieth century. We cover the following relevant topics:

1. The history of AIRs in biotechnology.
2. The definition of AIRs.
3. Their current prevalence.
4. Evidence of their benefits and risks.
5. Evidence concerning their evolution over time.
6. The policy implications of these findings.

RECENT HISTORY OF AIRs IN BIOTECHNOLOGY

Relationships between academic life scientists and industrial organizations have existed through much of the twentieth century (5). Prior to the 1970s these interactions consisted predominantly of consulting by academicians retained by pharmaceutical companies to help with the solution of particular research problems. Academically based clinicians also participated in clinical trials to test newly developed industrial products. Furthermore universities had been commercializing their intellectual property on a modest scale since the 1920s, when the University of Wisconsin created the Wisconsin Alumni Research Foundation to hold the patent on the technology for irradiating dairy products to instill them with vitamin D (6).

In the late 1970s and early 1980s, however, events in science, law and public policy combined to expand the potential value of AIRs in biotechnology. First and most important was the biotechnology revolution itself. This revolution represented the flowering of research investments by the federal government over 30 years following the end of the Second World War. The signature breakthrough heralding the new era in biotechnology was, of course, the development of recombinant DNA technology

by Cohen and Boyer in 1973, but parallel breakthroughs were occurring in monoclonal antibody technologies, large-scale fermentation, genetic sequencing, and genetic synthesis. The arrival of these new techniques suggested that academic research might have much more commercial relevance, both short term and long term, than had been supposed in the past, and also suggested that the next generation of dramatic pharmacological breakthroughs might be based in the biological rather than, as previously, the chemical sciences. During much of the early history of the pharmaceutical industry, the primary source of new agents was the chemical isolation and synthesis of naturally occurring compounds that had been found to have biological activity. Thus major pharmaceutical companies had developed deep expertise in chemistry, but were ill-prepared to take advantage of the biological revolution occurring in university laboratories. It became a pressing business priority therefore for the pharmaceutical industry to develop relationships with universities. AIRs would allow pharmaceutical companies to capture new intellectual property arising in universities, and would facilitate retraining of current industry investigators and/or recruitment of new talent who could then work within industry to exploit the new biotechnologies.

The dawn of the biotechnology revolution coincided with policy developments that facilitated AIRs in biotechnology. The 1970s were a time of deep economic anxiety for the United States (7,8). Rising oil prices and decades of complacency on the part of major U.S. industries had combined to produce rapid inflation and stagnant productivity at a time when Japanese industries were thriving. A crisis of confidence in the U.S. economy ensued, and the U.S. government began examining strategies for restoring the vitality of the country's economy. Policy makers concluded that as one such strategy, the United States should take better advantage of its large investment in university research (9). They also concluded that lack of incentives for universities and their scientists to exploit the commercial potential of their work was a major impediment to the success of this strategy. Congress in 1980 enacted the Bayh-Dole Act, which enabled universities to claim ownership to intellectual property resulting from federally sponsored research. The law also required that inventors within universities receive a share of the gains from commercialization of these properties. At least in theory, Bayh-Dole gave universities and their scientists a financial motive to cooperate with industrial partners.

Still another development, this one in patent law, gave an additional boost to the development of AIRs. In 1980 the Supreme Court ruled in the *Diamond v. Chakrabarty* case that it was legal to patent new life forms created as a result of biotechnological manipulation. Since many of the most promising products of the biotechnology revolution resulted from the creation of novel cells and organisms that yielded valuable biological agents, the Chakrabarty case reassured industries and universities that the intellectual property likely to result from AIRs (and other biotechnology endeavors) could be protected under existing patent law.

Following these changes in science, policy, and law, a number of highly publicized AIRs ensued. One of the first, the subject of critical congressional hearings, was an arrangement between the Massachusetts General Hospital and a German chemical and drug company, Hoechst A.G. Hoechst funded not only research at the MGH but also the creation of a new department of genetics and the construction of a research building. Other large relationships developed in the early 1980s as well: between Harvard Medical School and the Dupont Co., Monsanto and the Washington University, Bristol-Myers and Yale University, and others. Equally interesting was the emergence of small, biotech start-up companies founded by and with university faculty members. These included a number of enterprises that have survived to this day: Amgen, Biogen, Genentech, Immunex, and Chiron, to name a few. It is estimated that university faculty participated in the founding of 500 such companies over the 1980s and 1990s (8).

The participation of university faculty in founding new companies was hardly unprecedented. During the 1960s a number of engineers had left universities such as Massachusetts Institute of Technology, Stanford University, and California Institute of Technology to found the computer companies that gave rise to Silicon Valley and Route 128. However, for the most part, these professors cut their formal ties to the university, becoming full-time business persons. In contrast, many university-based biotechnology entrepreneurs wanted to maintain their faculty positions even while they held major roles in start-up companies. Some faculty became major equity holders in for-profit enterprises, and then conducted university-based research funded by those companies. This reflected not only the business aspirations of faculty but also the wishes of the venture capital companies that funded their new start-ups. In many cases venture capitalists felt that the success of the new businesses was dependent on providing faculty with strong incentives to stay involved and support the work of the start-up. Whatever the cause, the participation of faculty in founding new companies created a new and more intimate form of AIR.

There was also ample precedent prior to the biotechnology revolution for the funding of research in universities by industry. Research relationships between universities and companies were common in chemistry and engineering disciplines prior to the biotechnology revolution, and even after the arrival of the new biotechnologies, AIRs in chemistry and engineering were more prevalent than in the biological sciences. In 1985, 43 percent of faculty principal investigators in chemistry and engineering had research relationships with industries, compared with 23 percent involved in biotechnology (10).

Nevertheless, AIRs provoked greater controversy and more soul-searching in the life sciences than had AIRs in other fields. There were several reasons for this. First and most important were the potential implications of AIRs in biotechnology for health care services. The products of AIRs were likely to yield pharmaceuticals and devices that would be used in the treatment of patients. Observers worried that conflicts of interest on the part of university researchers might cause research bias that would ultimately hurt patients. More immediately, some research

sponsored by biotechnology companies required the use of human subjects. Concerns arose about whether university-based researchers would let financial interests compromise their management of research subjects in ways that could adversely affect their health in the short term.

Second, university research in biomedicine benefits from billions of dollars annually in federal support. Universities and researchers were very anxious that AIRs, which support only about 14 percent of research in U.S. academic health centers compared to 67 percent from the federal government (11), not in any way compromise the federal government's or voter's trust in academic biomedical research (12). The federal government, through the National Institutes of Health (NIH), was further concerned that the participation of faculty in AIRs not reduce the effectiveness of federal investments in research or training (13).

Third, the fact that university researchers were founding companies at a great rate, and staying in the university, seemed to set AIRs in biotechnology apart from AIRs in other fields. AIRs in biotechnology were creating a cadre of faculty entrepreneurs who continued to have teaching and administrative responsibilities in universities, and continued to participate in federal peer review and consulting roles, as though nothing had changed. The conflicts of interest and conflicts of commitment among such entrepreneurs were likely to be particularly intense, but there was no way to ensure, under then existing university policies, that colleagues or outside clients would be aware of the existence of such conflicts.

Fourth, the above concerns focused attention on the fact that there was virtually no data in field of biotechnology or any other academic discipline on the extent and consequences of AIRs in universities. The balance of this article explores these issues.

DEFINITIONS OF AIRs IN BIOTECHNOLOGY

Before trying to describe the extent and consequences of AIRs in biotechnology, it is useful to define more precisely what AIRs consist of from our standpoint: What exactly we are describing. For the purposes of this discussion, AIRs consist of arrangements between for-profit corporations and academic institutions (or their faculty, staff, and trainees) in which something of value is exchanged. Commonly universities provide a service (e.g., research or training) or intellectual property (in the form of a patent, license, or advice) in return for financial considerations of various types (research support, honoraria, consulting fees, royalties, or equity) (14).

AIRs in the health sciences can assume a variety of forms. The following types of AIRs are among the most common but by no means exhaust the alternatives:

1. Academic-industry research relationships (AIRRs): the support by industry (through grant or contract) of university-based research.
2. Consulting relationships: the compensated provision of advice or information, usually by individual faculty members, to commercial organizations.
3. The sale or licensing of patents by university to industries.
4. The participation by academic institutions or their faculty in the founding and/or ownership of new companies commercializing university based research: AIRs of this type often occur when cash-poor start-up companies use small amounts of equity to compensate faculty for consulting or other services. However, academic institutions or their faculty may also participate in the founding of new commercial entities, sometimes taking much larger amounts of equity in return for contributions of intellectual property (14–16).
5. Academic-industry training relationships (AITRs): industries provide support for the research or educational expenses of graduate students or postdoctoral fellows, or contract with academic institutions to provide various educational experiences (e.g., seminars or fellowships) to industrial employees.

These and other forms of AIRs may occur singly or in combination. The mixed forms of AIRs (e.g., those involving AIRRs and consulting or equity holding) often raise the most troubling concerns about conflict of interest because multiple relationships often involve more money (both real and potential) than single forms of AIRs.

Most current information on the dimensions and consequences of the AIR phenomenon in biotechnology concerns AIRRs, and much of that information pertains to AIRRs in the life sciences generally rather than specifically in the subfield of biotechnology. For that reason the ensuing discussion focuses particularly on AIRRs, and often references the field of life sciences generally. Nevertheless, where data specific to the biotechnology area and data concerning other types of relationships are available, we convey these as well.

PREVALENCE AND MAGNITUDE OF AIRs

The most recent nationally representative data on the prevalence and magnitude of AIRs stems from surveys of industries and faculty members in 1994 and 1995 (15,17). A 1994 survey of senior executives in a representative sample of life-sciences companies revealed that over 90 percent participated in some form of AIR. The most prevalent form was retention of university faculty as consultants (88 percent). Fifty-nine percent participated in AIRRs and 38 percent in AITRs. Seven percent of companies reported that faculty members were significant equity holders in their companies (17).

A contemporaneous survey of 2052 faculty members at the 50 most research intensive U.S. universities revealed that 28 percent of respondents reported receiving some research support from industrial sources (15). The prevalence of support was greater for clinical (36 percent) than nonclinical (21 percent) departments. Among a subgroup of faculty whose research involved what the authors defined as "biotechnologies" (recombinant DNA technology, monoclonal antibodies, gene synthesis, gene sequencing, tissue culture, enzymology, and large-scale fermentation), 21 percent of principal investigators on

research grants reported receiving research support from industry. This subsample was specifically chosen to be comparable to a 1986 survey of faculty using the same biotechnology techniques (see below) (10,15). Thus, as of the middle of the 1990s, between 20 and 30 percent of life-sciences and biotechnology faculty in research-intensive U.S. universities participated in AIRRs. There is no reason to suppose that this number has declined since that time, so current levels of faculty participation in AIRRs are likely to be at least that high.

Characteristics of AIRRs in the life sciences suggest that relationships tend on average to be small in size and short in duration. Industry respondents indicated that 71 percent of AIRRs in 1994 and 1995 were funded at less than $100,000 a year. Only 6 percent of responding firms provided annual funding of $500,000 or more. For 84 percent of respondents whose firms had relationships with academe, the typical relationship lasted two years or less. The generally short duration of AIRRs and the small funding levels suggest that at that time, the research they supported tended to be targeted — that is, applied rather than fundamental (17). AIRRs also constituted a relatively small proportion of the total research funding to universities in the mid-1990s: about 12 percent of the total. (The 14 percent figure cited above refers to academic health centers, which include teaching hospitals and which are more heavily weighted toward clinical departments, where the prevalence of AIRRs is higher).

Recent data about the prevalence of other types of AIRs are scarce. The 1985 survey of biotechnology faculty cited above indicated that 7 percent held equity in a biotechnology company related to their own work, while 47 percent consulted to industry. In a separate survey of nearly 700 graduate students and fellows in life-sciences departments at six leading universities, 19 percent reported receiving some research support from industry (18). Krimsky et al. showed in 1988 that as many as 31 percent of scientists in certain life-sciences departments had some form of link to outside firms (19).

BENEFITS OF AIRRs

The best documentation of the benefits that result from the relationships between academic institutions and industry derive from studies of AIRRs. For this reason we primarily focus AIRRs; however, we will comment on the benefits that derive from other forms of AIRs when such data are available.

The most obvious benefit of AIRRs is that these relationships provide funds to support the research conducted in academic institutions. In a 1994 survey of senior research executives at 306 life-sciences companies in the United States, respondents reported that their companies supported more than 1500 academe based research projects at a cost of over $340 million (17). Based on these reports it was estimated that the life-sciences industry as a whole supported more than 6000 life-sciences, projects and expended $1.5 billion for academic research in the life sciences.

Receipt of industry funds is not associated with detectable adverse effects on academic productivity. Indeed, if anything, AIRRs are associated with enhanced productivity on the part of involved university investigators. Faculty involved AIRRs exhibit higher levels of research productivity than faculty without such relationships. In a 1994-95 survey of over 2000 life-sciences faculty in the 50 most research intensive universities, faculty with funding from industry published significantly more articles in peer-reviewed journals in the previous three years than faculty without AIRRs (15). Faculty benefit from increased publications, since articles in peer-reviewed journals represent one of the main criteria by which faculty are awarded the trappings of academic success including promotions, tenure, prizes, future research grants, positions in professional organizations, and ultimately a place in the history of the scientific endeavor (21). At an institutional level more publications by faculty translate into greater prestige and, perhaps, an increased ability to attract top students, faculty, and future research funding for universities.

In addition to publications, AIRRs are associated with an increased likelihood of commercial activities on the part of faculty and their institutions. Blumenthal and colleagues (1996) found that compared to faculty without AIRRs, those with industry funding were significantly more likely to report that they had applied for a patent (24 vs. 42 percent), had a patent granted (12.6 vs. 25 percent), had a patent licensed (8.7 vs. 18.5 percent), a product under review (5.5 vs. 26.7 percent) a product on the market (10.8 vs. 26.1 percent), or a start-up company (6.0 vs. 14.3 percent) (15). A number of benefits may accrue to faculty as a result of the commercial opportunities that are associated with AIRRs, including financial returns, the opportunity to see the results of their research developed into useful products and services, and perhaps enhanced career opportunities in the industrial sector. Universities benefit from faculty commercialization, since their polices often provide the institution with the option to participate in commercial ventures such as supporting the costs of filing a patent in exchange for a portion of the licensing revenues or by providing venture capital funding for a start-up in exchange for a share of the future profits of that firm.

Several other benefits accrue to universities, faculty, and students as a result of AIRRs. For example, 66 percent of faculty reported that research grants and contracts from industry involved less “red tape” than those from federal sources, 60 percent felt that AIRRs enhanced career opportunities for students, 49 percent felt AIRRs increased the prestige of their department, 37 percent felt AIRRs contributed to their promotion and tenure, and 34 percent reported that these relationships led to salary increases (15).

Like universities and their faculty, companies also benefit from AIRRs. In our 1994 survey of life-sciences companies 60 percent of firms with AIRRs have realized patents, products and sales as direct result of those relationships (17). In addition to direct benefits such as patents, products, and sales, companies receive access to ideas, knowledge, and a pool of talented potential researchers. For example, 56 percent of companies with AIRRs reported that they depend very much on these

relationships to "keep staff current" with important research, 53 percent depend on them to provide ideas for new products and services, and 37 percent to aid in recruiting able researchers. Only 29 percent reported that they rely somewhat or very much on AIRRs to invent the products that the company will license (17).

Perhaps one of the most important benefits of AIRRs to industry may be that these relationships provide sponsors with access to the most recent research results of faculty — often months or years ahead of competitors. It is common for most AIRRs to allow a sponsor 30 to 90 days to review the results of the research they sponsored prior to submission for publication. An executive of a company said that in his field the published literature is "miles behind the front line of what is happening in universities" (16). For companies, especially those rapidly developing fields such as human genetics, using AIRRs as a means of access to new knowledge that is not yet public may constitute a considerable competitive advantage over companies without AIRRs.

Society ultimately benefits from AIRRs in terms of the increased flow of research results from universities into the industrial setting — a process often called *technology transfer*. A study by Cohen, Florida, and Goe (1994) suggested that collaborative research and development and other forms of intimate interactions between university researchers and industry personnel were more effective in transferring information into the industrial sector than communication through traditional academic channels such as publications and presentations (22). Cohen and his colleagues studied the extent, characteristics, and consequences of university–industry research centers (UIRCs). Based on a national survey of UIRC directors (response rate 48 percent), they found that the overwhelming majority of UIRCs were created with government support (70 percent) and continued to derive about 86 percent of their research funds from governmental sources. On average UIRCs derived 46 percent of their funds from the federal government compared to only 31 percent from industrial sources. These data suggested the important role of federal funding in creating and maintaining UIRCs as a potential mechanism for technology transfer and indicated that federal and industrial research coexist in close proximity to university settings. Cohen and his colleagues also found that UIRCs reported generating 211 patents in 1990. UIRC patent productivity per dollar invested was about one-third of that observed in industrial research and development (R&D) laboratories. Patent productivity tended to be higher in small UIRCs, those predominately funded by industry and UIRCs in the fields of biotechnology and advanced materials. The biotechnology field was also most productive among all scientific fields of new products from UIRCs. Twenty-two percent of UIRCs reported spin-off companies resulting from their work (22).

Broader evidence of the localized benefits of AIRRs have resulted from studies of what economists refer to as "spill over." In a seminal study Jaffe (1989) found a positive association between university-based and industrial innovation among companies in the same state, as measured by the number of patents issued to the firms between 1972 and 1986. In the drug industry, a 1 percent increase in university-based, biomedical research was associated with a 0.28 percent increase in the number of patents issued to drug firms (23). Additional work by Jaffe, Trajtenberg, and Henderson (1993) provided further confirmation of spillovers in a study of whether patents were more likely to cite earlier patents that originated in the same geographic location compared to earlier patents originating in other geographic settings (24). They found that later patents were significantly more likely to cite earlier patents that originated nearby than they were to cite a control group of patents from the same field and technological area that resulted from work done in a different state, standard metropolitan statistical area, or county.

In addition there are real, visible effects of AIRRs on local economies. For example, academic researchers have played a seminal role establishment of high tech industries. Etzkowitz (1988), Etzkowitz and Peters (1991), and Dorfman (1993) have documented the role of investigators at the Massachusetts Institute of Technology, Stanford University, and the University of California in founding and staffing local biotechnology and electronics companies such as Raytheon, Data General, Digital Equipment Corporation, Genetics Institute, and Biogen in the Route 128 area and Genentech in the Silicon Valley (25–27). Creation of these firms/industries no doubt included the creation of high-paying technical and professional jobs, an increased source of tax revenues for local economies, and an increased inflow of venture capital and other research-related services into the local economies.

RISKS OF AIRRs

As with all forms of individual and institutional behavior AIRRs have risks that must be addressed. According to Derek Bok, former president of Harvard University, AIRs may "...divert the faculty. Graduate students may be drawn into projects in ways that sacrifice their education for commercial gain. Research performed with an eye towards profit may lure investigators into conflicts of interest or cause them to practice forms of secrecy that hamper scientific progress. Ultimately, corporate ties may undermine the university's reputation for objectivity" (16). This quote encapsulates many of the worst fears about the potential negative influences of AIRs on the academic enterprise.

A frequently cited risk of AIRs is the potential for increased secrecy in academic science. Secrecy in scientific research can take a number of forms including delaying publications for an extended period of time, faculty refusing to share research results and materials when asked by other academic scientists, and the keeping of trade secrets. Blumenthal and colleagues (1997) found in their 1994–95 survey of academic life scientists that 27.2 percent of researchers with AIRRs delayed publication of their research longer than 6 months compared to 16.5 percent of faculty without industry funding ($p < 0.001$) (15). Faculty with AIRRs were significantly more likely than those without AIRRs to report having denied other university faculty access

to their research results or biomaterials such as cell lines, tissues, reagents, and so on (11 vs. 8 percent, $p < 0.01$). Participation in trade secrecy, defined as information kept secret to protect its proprietary value was significantly more prevalent among researchers with AIRRs (14.5 percent) compared faculty without AIRRs (4.7 percent, $p < 0.001$).

While secrecy may stem from individual scientists' motivations and career aspirations, it is likely that the policies of universities and corporate research sponsors encourage secrecy as well. According to an NIH study in the early 1990s, 20 percent of AIRRs permitted companies to delay publications for longer than six months, so that companies can review findings and secure rights to commercializable products (28). More than 80 percent of life science companies supporting research in universities in 1994 reported that their agreements sometimes require academic researchers to keep research results secret prior to filing a patent (17). Cohen and colleagues found that 41 percent of UIRCs had restrictions on their ability to communicate their research results to the general public, 29 percent on their communication with faculty at other universities, 21 percent on sharing information with faculty in their own institution ,and 13 percent on sharing information with scientists within their own research center (22).

Another risk mentioned by Bok was that AIRRs may have a negative impact on scientists in training. A 1985 survey of 693 advanced trainees in the life sciences at 6 universities found that 34 percent of respondents whose projects were supported by industry felt constrained in discussing their research results with other scientists (18). Further this study found that graduate students and post-doctoral fellows whose projects were supported by industry reported significantly fewer publications on average (2.62) than those with no industry support (3.67).

A third risk of AIRRs is that these relationships may lure faculty away from basic research, which has long been the mainstay of academics, towards research that has commercial applications. Unpublished research by Blumenthal and colleagues conducted in 1994 found that more than half (54 percent) of all life scientists felt AIRRs created pressures on faculty to "spend too much time on commercial activities" (29). Blumenthal et al. (1996) found that faculty members with industrial support were significantly more likely than those without AIRRs to report that their choice of research topics had been influenced somewhat or greatly by the likelihood that the results would have commercial application (35 vs. 14 percent, $p < 0.001$) (15).

A fourth risk associated with AIRRs is that too much funding from industry may be associated with lower research productivity on the part of involved faculty. Faculty who received more than 66 percent of their funding from industry published significantly fewer articles over a three-year period, published in less influential journals, and were less likely to report commercial outcomes from their research than faculty with less support from industry (15). This finding may reflect the faculty that faculty with more than two-thirds of their funding from industry are less able than others to attract peer-reviewed support from governmental and nonindustrial funding sources.

As with all business relationships there is the risk that AIRRs may not produce the outcome(s) industrial sponsors had predicted or hoped for. There are some data to suggest that the behavior of faculty may cause some AIRRs to be less useful to companies than when they were conceived. In 1995, 33 percent of life science firms with AIRRs reported that academic scientists had changed the direction of research conducted under an AIRR to the extent that the results were less useful to the corporate sponsor than had been originally expected (17).

Bok articulated the greatest potential risk of AIRRs when he wrote that these relationships may "undermine the university's reputation for objectivity." The public's generous support for research is founded on the belief that the results of research represent faculties' best effort to detect the truth, untainted by commercial interests. Recent research regarding the effects of AIRs on the outcomes of studies that examined the efficacy and safety of calcium-channel antagonists in the treatment of cardiovascular disorders suggests there may be cause for some concern (30). Between March 1995 and September 1996 more than 70 studies were published that were either supportive, neutral, or critical with respect to the safety and efficacy of using calcium-channel antagonists in the clinical setting (30). Stellfox and colleagues surveyed the authors of these 70 papers about their relationships with companies producing the antagonists or with companies producing competing products. He found that 96 percent of authors whose research was supportive of the use of calcium-channel antogonists had financial relationships with companies that produced antagonists, compared with only 60 percent of those whose research was neutral, or 37 percent of those whose findings were critical. Further he found that despite the fact that 44 out of the 70 authors had AIRs only 2 of the studies disclosed the authors' relationships with industry. As Stellfox and colleagues wrote, "We wonder how the public would interpret the debate over calcium channel antagonists if it knew that most of the authors participating in the debate had undisclosed financial ties with pharmaceutical manufacturers. . . .Full disclosure of relationships between physicians and pharmaceutical manufacturers is necessary to affirm the integrity of the medical profession and maintain public confidence" (30).

Despite concerns, several of the risks associated with AIRRs have not been substantiated. First, there is no evidence that AIRRs have resulted in a diversion of faculty effort from academic and administrative commitments — so-called conflicts of commitment. Data from Blumenthal and colleagues (1996) show that faculty with AIRRs spent as much time per week teaching undergraduates, graduate students, and postdoctoral fellows as those without AIRRs (15). Also AIRRs were associated with increased rather than decreased service activities on the part of faculty to their institution and their discipline. Faculty with AIRRs were significantly more likely than those without AIRRs to have been chairs of a departments, universitywide administrators, members of review panels for federal agencies, editors or editorial

board members of journals, the heads or associate heads of research institutes, chairs of a universitywide committees, or officers of professional associations (15).

AIRRs IN GENETICS (A SPECIAL CASE)

AIRRs in genetics differ significantly from the other life-sciences fields in terms of their prevalence, magnitude, benefits and risks. First, AIRRs are significantly more prevalent in the field of genetics than the other life-sciences fields. Based on a survey of 210 life-sciences companies in the United States, Blumenthal and colleagues (1997) found that after controlling for firm size, companies conducting genetics research were significantly more likely than nongenetics firms to support research in universities (69 vs. 45 percent, $p < 0.005$) (31). Also genetics firms were significantly more likely to support research training than other life science firms (46 vs. 33 percent, $p < 0.005$).

Second, AIRRS in genetics are longer in duration and involve more money that AIRRs in other life-sciences fields (31). Among genetics firms, 19 percent of AIRRs lasted three years or more compared to 14 percent among nongenetics companies. Agreements of one year or less were significantly less common in among genetics firms than nongenetics firms (15 vs. 34 percent, $p < 0.05$). Among large companies the median amount of research support provided to universities by genetics companies was $102,000 compared to $70,000 for large, nongenetics companies.

Third, there is some evidence suggesting that AIRRs in genetics have greater benefits in some respects than AIRRs in other fields (31). Among faculty with AIRRs, genetics researchers reported publishing more articles in peer reviewed in the preceding three years (18 vs. 14.5), participating in more service related activities within their institution or discipline and publishing in more influential journals than nongeneticists. Also genetics researchers with industry support were significantly more likely than other researchers with AIRRs to have applied for a patent, received a patent, licensed a patent, or started a new company.

Fourth, AIRRs in genetics are more prone to the risks of data-withholding than those in the other life sciences (31). Genetics firms with AIRRs were more likely than nongenetics firms with AIRRs to report that their agreements with universities required researchers to keep results secret beyond the time to file a patent. Also genetics researchers with AIRRs were significantly more likely to report that trade secrets resulted from their university research, to have delayed publication of their results in order to file for a patent, and to have denied direct requests from other scientists for access to their research results and materials than nongenetics researchers with AIRRs.

AIRRs IN BIOTECHNOLOGY, 1985–1995

Since the mid-1980s the rate of faculty participation in AIRRs has remained about the same (10,15). Based on 1985 and an 1995 survey of academic biotechnology researchers (faculty using recombinant DNA, monoclonal antibodies, gene synthesis, gene sequencing, cell tissue and culture, enzymology, and large-scale fermentation), 23 percent of biotechnology faculty in 1985 reported that they were principal investigators on research grants or projects funded by industry compared to 21 percent in 1995. For these faculty, industry supplied 7.4 percent of their total research budgets in 1985, as compared with 5.8 percent in 1995.

The experiences of faculty members in 1985 and 1995 were similar in other ways as well. From 1985 to 1995 the percentage of biotechnology researchers with AIRRs who reported that trade secrets had resulted from their research increased slightly from 12 percent to 17.2 percent. However, among those without AIRRs the percentage who had engaged in trade secrecy doubled from 3 percent in 1985 to 6.6 percent in 1995. Similar results were found regarding biotechnology researchers' choice of investigational topics. In 1985 and 1995, 30 percent of those with AIRRs reported that their choice of research topics had been influenced to some extent or to a great extent by the likelihood that the results would have commercial application. However, among those without AIRRs the percentage who reported that their choice of research topics had been influenced to some extent or to a great extent by the likelihood that the results would have commercial application doubled from 7 percent in 1985 to 14 percent in 1995.

IMPLICATIONS FOR POLICY AND MANAGEMENT

Persistent uncertainties about the scope and consequences of AIRs in biotechnology somewhat complicate the tasks of regulating and managing these relationships. Nevertheless, it is possible to draw some reasonable conclusions concerning these issues from existing data on the prevalence, magnitude, benefits, risks, and historical development of AIRs.

First, AIRs in biotechnology and the other life sciences have documented benefits that constitute a persuasive argument for continuing and even promoting AIRs of certain types and in certain situations. These benefits are best demonstrated for AIRRs, patent and licensing arrangements, and academe–industry training relationships, and include increased funding for academic research, possible increases in rates of patenting and publishing on the part of academic investigators, income from patents and licenses, and the apparent enhancement of the educational experiences of trainees.

Second, many types of AIRs also pose real risks for the academic institutions that participate in them. These risks include reductions in the openness of communication among investigators, channeling of research in more applied directions, and threats to the public credibility of the life sciences. With the exception of certain limited situations, however, these risks seem not to present a clear and immediate danger to the conduct of science in universities or to their educational missions and do not justify at this time limiting the freedom of academic institutions to establish AIRs. This conclusion could change as further information emerges concerning the

long-term benefits and risks of AIRs both for universities and for the scientific enterprise. For the time being, however, policy and management should emphasize disclosure on the part of life scientists, vigilance on the part of academic and public administrators, and further research into the positive and negative effects of these relationships. In addition it would seem prudent for academic institutions to avoid excessive dependence on industrial relationships for research support. This will minimize the chances that the effects of AIRs on the norms and behaviors of universities will have significant, lasting impact on the character of life science research.

Third, certain forms or combinations of AIRs pose qualitatively greater concerns than others. Intense conflicts of interest arise in equity-holding AIRs (especially those where potential gains to investigators are large) and in major, sustained consulting arrangements in which faculty members derive appreciable amounts of their annual income from one particular company or a small number of companies. Public and private managers are justified in subjecting these relationships to a higher level of scrutiny, in limiting the size and number of such relationships among their faculty, and, in certain cases, in forbidding them altogether.

Another situation requiring different treatment is AIRs in which patients are directly involved, such as research involving living human subjects. When academic clinical investigators have financial relationships with companies (usually in the form of substantial equity positions or major consulting income) that may benefit from their clinical research, the resulting conflicts of interest create the appearance or reality that the interests of human subjects may in some way be compromised. This could occur, for example, if financially involved clinical investigators failed to fully inform prospective research subjects of the benefits and risks of the clinical protocols, inappropriately pressed subjects who wished to withdraw to remain in a research protocol, or engaged in biased patient recruitment to increase the chances of a successful outcome. The chances of such occurrences are less for large-scale, multicenter clinical trials than for more exploratory types of clinical investigation. Nevertheless, even large-scale clinical trials may not be exempt from such concerns when they occur in a discipline with a small number of leading investigators who may all have relationships with sponsoring companies.

One final situation that raises special issues is when academic administrators develop personal financial relationships with outside firms with which their faculty are also involved. Such relationships have no documented benefits but jeopardize the real or apparent ability of universities to manage objectively their faculty members' AIRs.

A fourth conclusion is that it is neither practical nor desirable for the federal government to dictate detailed rules for management of AIRs in biotechnology or the other life sciences across the United States. Past experience with oversight of research involving human subjects suggests the feasibility of permitting individual institutions to take responsibility for overseeing AIRs. However, experience with institutional review boards also suggests the need for continuing federal supervision and review of universities as they attempt self-regulation of sensitive ethical and policy issues related to research. Such supervision and review will undoubtedly be required for federal sponsors of research in the area of AIRs.

With these conclusions in mind, the following specific recommendations for universities and for federal research sponsors seem appropriate. For academic institutions:

1. All universities conducting biotechnology and health science research should require regular disclosure by faculty (including those not receiving federal funds) and senior administrators of financial relationships with companies that have life-sciences or health care interests. These disclosures should be reviewed carefully and confidentially by academic managers. Disclosure constitutes the minimal acceptable response of academic institutions to the demonstrated risks posed by AIRs in biotechnology and other fields. It is also impossible for academic institutions to learn from experience with AIRs if they do not know they exist.
2. Academic institutions should develop explicit policies for deciding which AIRs in biotechnology are desirable and undesirable. In this regard it would be prudent for universities to prohibit research involving living human subjects on the part of investigators with major financial interests in companies that may benefit from the results of that research.
3. Academic institutions should avoid excessive dependence on industrially sponsored research, given the proven risks of such relationships. The definition of *excessive* will undoubtedly vary from institution to institution, but a reasonable rule would be to keep industrial research support below one-third of total funds for biotechnology and other life-sciences research. Given dramatic recent increases in NIH funding, meeting this target will be appreciably easier than it was in the 1990s.

For federal sponsors of research:

1. The federal government should not fund clinical research when the principal investigator has a personal financial relationship with a company that may be affected by the outcome of that research. Exceptions may be made in cases where the relationships are minimal according to standards defined by the federal government.
2. The federal government should not fund research at institutions that do not have formal policies and procedures governing academe–industry relationships, or where such policies and procedures cannot be fairly and effectively enforced. The latter circumstance would arguably exist when there is no effective enforcement process, or where academic administrators themselves share financial interests in companies in which faculty members are also involved.

Academic industry relationships in biotechnology and the life sciences generally are part of the modern life-sciences

economy. They cannot and should not be prevented. But their benefits should not be exaggerated, nor their risks minimized. Academic institutions are priceless resources whose integrity and independence are critical to the long-term health of the American people and the American economy.

BIBLIOGRAPHY

1. E. Lau, *Washington Times*, D2 (1999).
2. R.L. Hotz, *L.A. Times*, A1 (1999).
3. R. Saltus, *Boston Globe*, A6 (1998).
4. C.G. Stolberg, *N.Y. Times*, A17 (1998).
5. J.P. Swan, *Academic Scientists and the Pharmaceutical Industry*, The Johns Hopkins University Press, Baltimore, MD, 1988.
6. D. Blumenthal, S. Epstein, and J. Maxwell, *N. Engl. J. Med.* **314**, 1621–1626 (1986).
7. W.E. Deming, *Out of the Crisis*, MIT Press, Cambridge, MA, 1982.
8. M. Kenney, *Biotechnology: The University-Industry Complex*, Yale University Press, New Haven, CT, 1988.
9. D.J. Prager and G.S. Omenn, *Science* **207**, 379–84 (1980).
10. D. Blumenthal et al., *Science* **232**, 1361–1366 (1986).
11. The Commonwealth Fund Task Force on Academic Health Centers, *From Bench to Bedside: Preserving the Mission of Academic Health Centers*, The Commonwealth Fund, New York, 1999.
12. Hearings Before the Human Resources and Intergovernmental Relations Subcommittee, 101st Congr., 1st Sess., Is Science for Sale? Conflicts of Interest v. the Public Interest, 1989.
13. General Accounting Office, *Controlling Inappropriate Access to Federally Funded Research Results*, U.S. General Accounting Office, Washington, DC, 1992.
14. D. Blumenthal, *J. Am. Med. Assoc.* **268**, 3334–3339 (1992).
15. D. Blumenthal et al., *N. Engl. J. Med.* **335**, 1734–1739 (1996).
16. N. Bowie, *University-Business Partnerships: An Assessment*, Lanham: Rowan & Littlefield Publishers, London, 1994.
17. D. Blumenthal et al., *N. Engl. J. Med.* **334**, 368–373 (1996).
18. M. Gluck, D. Blumenthal, and M.A. Stoto, *Res. Policy* **16**, 327–366 (1987).
19. S. Krimsky, J.G. Ennis, and R. Weissman, *Sci. Technol. Hum. Values* **16**, 275–287 (1991).
20. D. Blumenthal et al., *Science* **231**, 242–246 (1986).
21. M.F. Fox, in J.C. Smart, ed., *Higher Education: Handbook of Theory and Research*, Agathon Press, New York, 1985, pp. 255–282.
22. W.R. Cohen, R. Florida, and W.R. Goe, *University-industry Research Centers in the United States*, Carnegie Mellon University, Pittsburgh, PA, 1994.
23. A.B. Jaffe, *Am. Econ. Rev.* **79**, 957–990 (1989).
24. A.B. Jaffe, R. Trajtenberg, and R. Henderson, *Q. J. Econ.* **108**, 557–598 (1993).
25. N.S. Dorfman, *Res. Policy* **12**, 299–316 (1983).
26. H. Etzkowitz, in E. Mendelsohn, M.R. Smith, and P. Weingart, eds., *Science, Technology and the Military*, Kluwer Academic Press, Boston, MA, 1988, pp. 515–540.
27. H. Etzkowitz and L. Peters, *Minerva* (Summer) 133–66 (1991).
28. B. Healy, *N. Engl. J. Med.* **329**, 725–727 (1993).
29. D. Blumenthal, Unpublished results from a 1995 survey of life science faculty, 1999.
30. H.T. Stelfox et al., *N. Engl. J. Med.* **339**, 101–106 (1998).
31. D. Blumenthal, N. Causino, and E.G. Campbell, *Nat. Genet.* **16**, 104–108 (1997).

See other entries TRANSFERRING INNOVATIONS FROM ACADEMIC RESEARCH INSTITUTIONS TO INDUSTRY: OVERVIEW; UNIVERSITY-INDUSTRY RESEARCH RELATIONSHIPS, ETHICS, CONFLICT OF INTEREST.

AGRICULTURAL BIOTECHNOLOGY, ETHICS, FAMILY FARMS, AND INDUSTRIALIZATION

JEFFREY BURKHARDT
University of Florida
Gainesville, Florida

OUTLINE

INTRODUCTION

After nearly a century of neglect, agriculture is receiving increased attention from intellectuals concerned with social ethics. Along with the ethical dimensions of agriculture's impact on the environment and nonhuman animals, one fundamental ethical concern is the plight of the family farm, the traditional farming unit around the globe. The family farm is considered by many ethicists to have special moral, cultural, or political-economic significance, and various social and economic forces — especially new agricultural technologies — appear to threaten to drive the family farm to extinction. The increased reliance on high technology approaches to farming, and the increasing dependence of agriculture on other sectors of the economy, especially manufacturing and petrochemical refining, is referred to as the industrialization of agriculture. While

the industrialization of agriculture has been a century-long trend, the emergence of agricultural biotechnology in the 1970s heightened concerns that the family farm was becoming even more threatened. This is because agricultural biotechnology was thought to benefit primarily (or exclusively) large, already industrialized farms. If, as some people argue, society has some ethical obligation to protect or save family farms, then industrialization overall, and the new agricultural biotechnologies in particular, are cause for serious ethical concern. Independent of religious or environmental objections to biotechnology, there are three "family farm critiques" of agricultural biotechnology. These are based on the potential damage biotechnology might inflict on (*1*) an important political-economic entity, (*2*) a cherished symbol if not the embodiment of basic moral values, and (*3*) the solution to long-term natural resource problems.

INDUSTRIALIZATION OF AGRICULTURE

Industrialization of Agriculture: A Brief Overview

The "industrialization of agriculture" is a catchword for a broad set of changes that have occurred in agriculture in the United States and other developed countries over the last one hundred years. Its main features are as follows:

- The transition from animal powered farming techniques to machine power, with its attendant need for electrical or petrochemical fuel.
- The transition from the use of inputs (seed, fertilizers, and pest control measures) produced on the farm and reused yearly to the purchase of inputs from nonfarm sources, such as seed companies.
- The transition from small- to medium-sized farms, worked by a farm family and a few hired hands, to large-scale farms where all workers are hired and the farm manager may not even be the farm owner.
- The transition from localized and seasonal farm markets to regional, national and international markets, and "seasonal" produce available year-round.
- The transition from numerous independent farm producers supplying markets and processing firms to a relatively small number of large farm producers supplying commodities under contract to processing firms.
- The overall integration of agriculture into the larger industrial society with farming conceived of as "just another industry" akin to manufacturing or retail.

One could argue that the "industrialization" of agriculture began when humans first began using fabricating tools for farming, such as stone hoes, metal plows, or leather harnesses for draft animals. Most observers of agriculture, however, identify the beginnings of industrialization with the emergence of agricultural chemicals, particularly fertilizers, in the late 1800s. A more significant change in agriculture occurred with the development and widespread adoption of mechanical technologies, especially the gasoline-powered tractor, in the 1920s in the United States. Refinements and major breakthroughs in agricultural machinery (combines, harvesters, and postharvest storage technologies) and agricultural chemicals (pesticides, herbicides, and animal pharmaceuticals) continued throughout the middle twentieth century. After World War II the pace of these developments accelerated. In recent decades additional improvements in machines and chemicals have been accompanied by the growing use of computers in agriculture. Now we are witnessing what some see as the culmination of technological industrialization, the introduction of various biotechnologically produced products for agriculture.

The industrialization of agriculture is not simply the transition from a primarily manual-labor-based enterprise to a more technologically based production system. The biological, economic, and sociological dimensions of agricultural have changed in important ways as well. For example, the basic biological unit of crop production — the seed — has undergone significant changes. Soybeans and corn are no longer produced from the best seed saved from the farmer's previous year's harvest. Seeds are now biologically altered by scientists, patented (or given other patent-like protections) and sold to farmers annually. Hybridization, while augmenting desired traits in plants (e.g., salt tolerance and drought resistance), has made farmers dependent on external agents — seed stores, seed companies, and researchers in agricultural colleges — for the acquisition of their basic production input. Similarly chemicals and gas-powered machinery have to be purchased, often at considerable cost. Computer usage requires expensive operating programs and training. Products of biotechnology such as bioengineered seed or genetically altered animals also have to be purchased. Each of these changes in farm technology carries the attending consequence that agriculture increasingly must rely on nonfarm sources for the biological and mechanical inputs necessary for farming in its current form.

These changes have brought socioeconomic consequences. Three in particular stand out: (*1*) The size of farms has steadily increased since the turn of the twentieth century, (*2*) the number of farms in the United States has rapidly declined since around 1930, and (*3*) there has been a sharp decline in the number of people involved in agriculture — both owners and hired labor. The trend in farm size is directly attributable to the nature of the technological changes that have occurred in farming. It is a generally recognized fact that technologies are not what economists refer to as "scale-neutral." Certain technologies tend to favor production systems or enterprises of a particular size or scale. In agriculture, most of the technologies introduced over the last 40 years tend to favor large farm operations. Cost is a factor. Purchasing chemicals or seed in large quantities reduces the per unit or marginal cost of those expenses. Large operations are typically in a better financial position to buy mass qualitites and thereby realize a savings. Cost is not the only factor, however. While a large 400 horsepower grain harvester is expensive, it is also better suited to large fields. Most of the important new technologies introduced into farming since World War II have tended to be (and it is frequently claimed were intended to be) more useful to

large-scale farming operations. As large-scale farms have grown larger, the total number of farms has declined along with the number of people involved in farming. There are several reasons for these related trends. The departure of people from agriculture and the decline in the number of farms from 1930 to 1945 has been attributed to the economic hardships farmers faced during the Great Depression. Many lost their farms because of an inability to purchase critical inputs and/or receive credit from failed banks. Around 1940 many young farmers and male children of older farmers went to serve in the war. Of those who returned home, many preferred to forsake the hard life of farming for alternative employment in newly suburbanizing America. Others pursued advanced education with the help of the GI Bill. The result was a rapid decline in the total number of farms and farm-based employment.

Even 50 years after World War II farm numbers and farm employment continue to decline. Urbanization, loss of farmland, and the less-than-glorious nature of farm work undoubtedly continue to contribute to this decline. It has been argued that there has been a systematic attempt by the federal government and nonfarm agribusiness industries to consolidate production into fewer and larger farms. The development of non-scale-neutral technologies have contributed to that. Many small farms are now what amounts to "hobby" farms that do not contribute in a significant way to the total output of agricultural as a whole. Middle-sized farms are forced either to get big or be reduced to hobby-type farms.

Current Structure of Agriculture

The term "structure of agriculture" refers to the overall nature of the agricultural industry in terms of its economic, sociological, and demographic features. The industrialization of agriculture has resulted in a restructuring of the nature of farming in the United States and other developed nations. The composition of farming, for example, the number and size of farms, who farms, and what the relationship is between farms and nonfarm enterprises, has changed. Currently the United States has what is called a "bipolar" agricultural structure. Of the nearly two million farms in the United States, fewer than 25 percent produce more than 80 percent of foodstuffs Americans consume. About 15 percent of farms account for three quarters of total agricultural sales (in dollars). The average income for roughly 75 percent of U.S. farms is around $17,000. Nearly all U.S. farms supplement their income with nonfarm income, and for most farms the nonfarm income exceeds farm income. However, for most small farms, the nonfarm income actually covers net farm losses. In terms of size, nearly a million and a half of U.S. farms are less than 80 acres, whereas some farms in California, Texas, and Florida are as large as 200,000 acres. Ownership patterns have changed as well. While most small- and mid-size farms are single-family proprietorships or family-owned corporations, many of the largest and greatest revenue-producing farms are owned by corporations whose primary enterprise is not agricultural. For example, many farms are owned by petrochemical companies, restaurant chains, and a consortium of urban-based investors. A legitimate question that can be raised is why agricultural industrialization or the increasingly bipolar structure of agriculture should be of any concern, given that (*1*) less than 3 percent of the U.S. population is engaged in farming and (*2*) Americans spend the smallest portion of their income (15 to 18 percent) on food compared with the rest of the world. One could argue that the structural changes occurring are simply the logical result of the technological transformation of agriculture. As long as consumers continue to spend a relatively small portion of their income on agricultural products and remain happy with the outcome of agriculture's industrialization, there is little cause for concern, economic, ethical, or otherwise. Some people believe, however, that there is cause for concern.

Ethics and Industrialization

Consumers are generally satisfied with the relatively low price of food and fiber products in the United States. This establishes, for some, an ethical justification for the industrialization of agriculture and its attendant structural effects. Utilitarian ethical theory holds that actions or policies are ethically sound if they result in "the greatest good of the greatest number," which many utilitarian ethicists interpret to mean "most satisfied preferences." The longstanding goals of agricultural policy and practices — including the development of agricultural technologies — have been (*1*) enough food, (*2*) a safe food supply, and (*3*) inexpensive food (the cost of the food should be at a rate that will return a reasonable profit to farmers). These three goals can be referred to as the "QQP" outcome: sufficient quantity, adequate quality, affordable (and profitable) price. It is generally held that the U.S. food system has met these goals and any utilitarian evaluation of the system would conclude that the current nature of agriculture is not only acceptable but also "ethically best." However, this line of thinking precludes customer awareness of agricultural processes. Consumers may never see the process or even think about the process that their food goes through before it reaches their table. Some would say as long as people get their food it does not matter to them what is going on in the agricultural process. Therefore agriculture and the technologies employed in food and fiber production can remain "invisible" tools for the satisfaction of consumer preferences and still be justified on utilitarian grounds.

According to utilitarian reasoning, it is only the increasing visibility of some of the results of industrialization (agricultural or otherwise) which accounts for questions about the ethics of contemporary practices. The noticeable environmental effects of agriculture processes, for instance, water pollution, soil erosion, or even the smell of diary or swine production, have begun to erode, though not completely undermine, the "satisfied preferences" basis for industrialized agriculture's legitimacy. In the 1960s heavy air pollution and fouled rivers led to concerns about the legitimacy of the smokestack industry and forced changes in practices and policies in that industry. If the impacts of industrialization on small farms were to become as apparent to the public as they are to residents of rural communities, one might expect a change in attitudes and policies toward the industrialization of agriculture. For the

present, however, it appears that industrialization will be endorsed despite the attempts of contemporary Agrarian philosophers and small-farm activists to alert the public to the negative impacts of industrialization. After all, in the utilitarian calculus, the livelihoods of less than 2 million people weigh lightly against the satisfaction of 260 million others. Even so, there are strong ethical arguments about why we should be concerned about industrialization and the restructuring of agriculture.

ETHICS AND THE FAMILY FARM

Why Focus on the Family Farm?

In the 1960s and early 1970s, people concerned with social ethics were alerted to the significance of agricultural practices and policies through the publication of Rachel Carson's *Silent Spring* (1) and Ruth Harrison's *Animal Machines, the New Factory Farming Industry* (2). During this time the focus of agricultural ethics was on the environmental consequences of agriculture and the treatment of nonhuman animals. The scope of agricultural ethics began to broaden in the late 1970s, and its focus began to shift. Films and television documentaries in the United States highlighted the conditions of farm labor. The public's attention was caught by the successful unionization of California farm labor. Food safety became a matter of larger public concern as questions were raised about the carcinogenicity of artificial food additives. U.S. foreign aid, trade, and development policies came under increased scrutiny after the sale of U.S. grain surpluses to the Soviet Union and after the famine in Bangladesh. By the early 1980s ethical analysis and critique of agriculture was well underway. It was Wendell Berry's *The Unsettling of America* (3), however, that ignited the most focused philosophical and ethical works in agricultural ethics. Berry articulated and critiqued the whole philosophy of modern agriculture. His work prompted social scientists, ethicists, and philosophers to study the trends, meanings, and normative implications of the modern agricultural production with a focus on its impact on traditional family farms.

Today agricultural or food system ethics encompasses a broad range of ethical concerns — chemical use, farm labor and management practices, impacts on animals, environmental pollution and resource depletion, the health of the land, food safety, and issues relating to international aid, trade, and development. Yet, for many agricultural ethicists, the key to understanding the ethical depth and complexity of agriculture — and how many of the practices of contemporary farming are ethically objectionable — is the family farm. Berry argued that the traditional family farm represents a way of life that precludes its contributing to environmental or cultural problems associated with industrialized agriculture. Most agricultural ethicists see the family farm as Berry characterized it: a standard against which most ethical issues or problems in agriculture can be "tested." There are three ways in which the family farm serves as this standard: (*1*) from a political-economic perspective, the question of whether family farms can continue to exist represents a test of the fairness or justness of democratic, market-based societies, (*2*) from a cultural or moral-value perspective, the question of whether values associated with traditional family farms continue to be viable serves as a test of the moral or spiritual health of modern society, and (*3*) from the perspective of our responsibility to future generations, the question of whether environmentally sustainable practices of family farms can be employed in modern farming tests the legitimacy of current production practices. Acceptance of the family farm as an ethical standard in any of these three domains gives agricultural ethicists reasons to take issue with biotechnology: Agricultural biotechnology may threaten or undermine the family farm. As such, these three perspectives together form what can be referred to as "the family farm critique" of agricultural biotechnology. Before proceeding to that matter, a brief examination of the arguments supporting the ethical standing of the family farm is in order. It should be noted that the following sections, the terms *agrarian populism* and *agrarian traditionalism*, when referring to contemporary philosophies, are derived from Thompson et al. (4).

Agrarian Populism and Farmers' Rights

Agrarianism is a philosophy that holds that farming is an important social good. Farming is a profession or occupation that ought to be respected, and in the legal-political realm, protected. The United States has had a long history of political and social concern for the plight of farmers. In his *Notes on the State of Virginia*, Thomas Jefferson argued that small-scale agricultural freeholders are good citizens and essential for the political-cultural life of the new nation being formed. They embody the "spirit of independence," but more important, a class of land-based, geographically dispersed, independent laborers serve as a political check against powerful urban interests and other threats to democracy. According to Jefferson, farmers and farm communities are to be encouraged and protected. Jefferson's ideas continued to be voiced in the cultural practices and government policies of the United States for more than a hundred years. It was not until the late 1860s that small-scale farmers realized that the Jeffersonian agrarian vision was being undermined. It was during this time, that the combined effect of U.S. Department of Agriculture programs, federal monetary policies, and the growing political strength of banking and manufacturing interests began to place small farmers in political and economic jeopardy.

At the turn of the century, Agrarian *Populism* (after the People's Party, circa 1870–1920) resurrected political-cultural arguments concerning the importance of small-scale farms. According to Populists, farmers' fundamental rights (property rights and the right to self-determination) were being deliberately violated or threatened by large business enterprises and government programs. Populists argued that small-scale farmers were entitled to "equal protection under the law," at the very least. Indeed, big business and big government at the time were acting to undermine basic principles of democracy and free-market capitalism. Populists demanded reforms and found some concessions. Nevertheless, the Agrarian

Populist movement failed to protect small farms over the longer term.

If the Populists had succeeded in all their demands regarding the protection of small farms, industrialization and structural change would not be an issue. Indeed, it is partly the result of the *failure* of Agrarian Populism as a political reform movement that Agrarian Populist arguments remains philosophically significant position to this day. As things stand, the philosophical tenets of Agrarian Populism are the basis for an ethical analysis and critique of contemporary agricultural practices and policies, including biotechnology and biotechnology policy. According to one contemporary spokesperson for Agrarian Populism, Jim Hightower (5), the family farm remains an important ethical, political, and economic entity: The family farm is one of the last, if not the last, holdouts in the attempt to secure fundamental values of democratic societies — freedom, self-determination, and equality of opportunity. While it may be no more important than, for example, hardware stores or plumbers, family farms are at the very least entitled to not be discriminated against in markets or in public policy (including research and development policies). The claim (like that of the turn-of-the-century Agrarians) is that the government and big business have conspired to drive family farms out of business and out of existence. Government policies encourage large farms, and large-scale agribusiness firms are always waiting in the wings to scavenge the remains of small farms unable to stay in production. It is unethical that individual farmers' rights and opportunities are systematically being violated, whether it is deliberate or simply the result of the socioeconomic system. It is also a harbinger of the death of democracy and free-market capitalism. According to Hightower and fellow Populist Marty Strange (6), family farms are potentially economically and politically viable despite what defenders of current agricultural practices claim. This is true, however, only if the government takes steps to protect family farms. Family farms may have no intrinsic or special ethical value to some, but they are valuable to the greater society. As a matter of justice, governments should guarantee that markets and policies are fair.

Agrarian Traditionalism and the Moral Value of Farms

Somewhat in contrast to Populism, Agrarian *Traditionalism*, exemplified in Wendell Berry's work, holds that traditional family farms have intrinsic or special ethical significance. Traditionalists agree with Populists that family farms should be preserved and protected but for different reasons. According to Agrarian Traditionalism, the family farm is at once the embodiment as well as the symbol of a set of values and virtues that have inherent worth. Among these values or virtues are self-reliance, community, and communion with nature. Traditionalists argue that the traditional family farm, by the very nature of the activities that occur thereon, promotes those virtues. Family farms engaged in the difficult labor of harnessing nature's power in order to survive, foster strength of will, courage, and self-determination among family members. The nature of farm work and the need for members of the larger farm community to help each other during times of adversity fosters a sense of community values — sharing benefits and burdens, joys and sufferings. Those values and virtues are self-reinforcing and above all healthy for body and soul. Traditionalists argue that the modernization of farming — agriculture as a business — degrades and threatens farming as a "way of life." There are differences within Traditionalism concerning the precise meaning of the family farm as a moral ideal in this regard. To some, family farms, even in the present age, embody and promote those values or virtues. If family farms are intrinsically good in so doing, an implicit moral judgment might follow that everyone should engage in family farming. Since this is not possible in modern society, some traditionalists hold that while family farms do embody these virtues, the more important point is that they serve as symbols or paradigms for how fundamental ethical virtues such as community and respect for nature should be regarded. The family farm, in other words, is a metaphor for the good life, ethically conceived, rather than a profession or occupation to which all people ought to devote themselves.

Whether Agrarian Traditionalism is understood as advocating for the intrinsic value of actual family farms or suggesting that the family farm serves as a metaphor for the ethical life, a critique of modern agriculture or modern society necessarily follows from its tenets. Traditionalism holds that modern agricultural practices are essentially inimical to self-reliance, moral character, family values, and communities. Modern society is business-oriented and materialistic, and decidedly out of sync with basic human needs, especially spirituality. Moreover modern agriculture and society are insensitive to the rhythms of nature and the organic or holistic features of the traditional family farm. If there are fundamental ecological, psychological, and philosophical truths, modern agriculture and society are alienated from these. Perhaps it is not possible for everyone to experience these by actually farming. They are nevertheless fundamental ethical goods that would, if followed, lead to significant changes in the way modern people live their lives.

Family Farms and the Future

Agrarian Populists and Agrarian Traditionalists share the belief that since family farms have political-economic or intrinsic ethical value, public policies should at least preserve and protect family farms, if not actively promote or encourage them. Agrarians' arguments are increasingly being joined by proponents of "sustainable agriculture," who find the structure of traditional family farms and their farming techniques to be more consistent with long-term environmental stewardship which is essential for a sustainable planet. According to proponents of sustainability, human beings have fundamental ethical obligations to the future. Among these obligations is the duty to not exploit natural resources to the point that future generations will be unable to sustain themselves though food and fiber production. Given that many if not most modern industrial agricultural practices are resource-depleting there is an ethical obligation to change those practices. Taking this one step further, it has been argued that present people have an obligation to

leave for posterity sound democratic institutions and a heritage of deep cultural and moral values. In each case, then, sustainability advocates strike notes similar to the Agrarians: Save the family farm, for political-economic, cultural-moral, and environmental reasons. This also is a matter of intergenerational justice.

In each of the positions described above, the ethical argument implies a critique of many if not most contemporary agricultural practices and policies. Agriculture has become in most developed nations a fairly large-scale, high technology, inputs-dependent industry. Family farm advocates have challenged large-scale corporate farmers, government policy makers, and agribusiness inputs manufacturers (chemical, mechanical, biotechnological) on ethical grounds. They argue that not only should family farms be protected or promoted, but many of the practices and practices associated with modern agriculture must be rejected. Family farm proponents have begun to target corporate and government/university actors involved in research and development for foisting ever-increasing industrialization on the farm sector. In essence, Agrarians and advocates of sustainable agriculture have come to find high technology — machines, chemicals, computers, and now biotechnology — ethically indictable in the apparent continuing demise of the traditional family farm. There is a shared critique and a shared vision of what is wrong with modern, industrial agriculture.

BIOTECHNOLOGY, INDUSTRIAL AGRICULTURE, AND THE FAMILY FARM

Emergence of Agricultural Biotechnology

The first applications of biotechnology were primarily in the medical/pharmaceutical field, such as in producing bioengineered insulin. The earliest agriculturally related products to emerge from genetic engineering were enzymes for fermentation (mainly for cheese) and agents for the biological control of pests, for example, *bacillus thuringiensis* (Bt). These products of biotechnology were not exactly the breakthroughs that the biotechnology community and agriculturalists have envisioned in the early 1970s (7). Rather, they were simply organisms whose commercial use value was enhanced through biotechnology, since genetic engineering made their large-scale production more efficient and less costly.

In 1970 the commercial prospects for agricultural biotechnology were made more attractive to the scientific community and to industry with the passage of the Plant Variety Protection Act (PVPA). PVPA provided for patentlike protections for novel plant species, whether the new or improved species was produced through conventional plant breeding or through bioengineering. PVPA's legal protections were extended as a result of two U.S. Supreme Court decisions (1980 and 1985), which allowed complete patent protection of novel organisms and plant species. The U.S. Patent and Trademark Office's decision (1988) to allow researchers at Harvard University to patent a mouse that had been generically altered, and this brought to completion the legal protections necessary for researchers involved in biotechnology to forge ahead with the development and commercial release of biotechnology products and processes, including those related to agriculture.

Agricultural researchers interviewed in a 1983 U.S. National Science Foundation-sponsored study believed that genetically engineered varieties of tomatoes and wheat, for example, would be in farmers' fields within five to ten years (7). In fact those bioengineered species are only now becoming commercially widespread. The overall process of bringing commercial agricultural biotechnologies to the marketplace or to the farm has been very slow. While pharmaceuticals and "biologics" (e.g., growth hormones) for animal agriculture are becoming increasingly available, the pace of the introduction of these has been far less rapid than early proponents of agricultural biotechnology (researchers and corporate marketing agents) had predicted and promoted. At present, the actual number of products from biotechnology currently being used in agriculture is not large. Nevertheless, given patents and patentlike protections, and the prospects for large profits from bioengineered agricultural products, that number is likely to increase at a rapid rate. Estimates are that market for U.S. agricultural biotechnology will grow from approximately $400 million in 1998 to over $2 billion by 2008 (8,9).

Despite the risks generally associated with biotechnology, there has been little public controversy over agricultural biotechnology. Although the most sophisticated application of genetic engineering to date (1998), cloning, was performed on an agricultural animal, controversies about cloning focused the possibilities of cloning human beings and not on agricultural applications. Perhaps the most contentious development in agricultural biotechnology per se has been the introduction of bovine somatotropin (bST), the nonsteroidal hormone capable of increasing milk production in dairy cattle. Yet even the main objections to bST, mostly at the public policy level and initiated by consumer groups, have had to do with its safety relative to human consumption of milk from bST-treated cows. Its role in agricultural industrialization and restructuring, while noted by some activist groups, has received little public or governmental attention. This is understandable given the relatively little attention the larger public devotes to *anything* related to agriculture — unless it directly affects human health or the environment.

Reach of Agricultural Biotechnology

The industrialization of agriculture has had the effect of blurring lines between research activities, industries, and governmental activities that formerly were separable and relatively isolated from each other. Agricultural animal science research is now closely tied to human medicinal and pharmaceutical research. Petrochemical and pharmaceutical firms own seed companies and animal-breeding facilities. Food safety oversight and regulation now includes involvement from the Department of Agriculture, the Food and Drug Administration, and the Environmental Protection Agency. Each of these connections suggest that agricultural production has become more and more integrated into the larger industrial society.

One major goal of industrialization in any sector of an economy is *control*. Industrialization requires and enhances control over a production-distribution system, whether it is the factory, the transportation of products or marketing. This goal has permeated industrial agriculture. Farmers and the firms supplying agricultural technologies have always looked for ways of increasing control over the production of food and fiber. Part of the appeal of agricultural *bio*technology is the promise of even greater control over farming. Farmers and agribusiness concerns differ, however, in what kinds of control they wish to exercise over farming.

Agricultural biotechnologies can be categorized in terms of what sorts of control they allow the farmer over the production process. The most basic units in agriculture are, for crops, the soil, water, and solar resources and the plants or seeds to be grown. For animal agriculture, they are the animals (swine, cattle, chickens, etc.) and the foodstuffs necessary for producing animal products. The longstanding goal of traditional plant and animal breeding has been the introduction of traits into plants and animals that would allow them to be more productive relative to the conditions under which they are grown, which includes both natural circumstances (soil, water, etc.) as well as inputs such as feed. Biotechnology, at least in theory, makes those goals more attainable because genetic engineering is quicker and more precise in the transference of the genes controlling those traits. Plant varieties that are pest resistant, drought resistant, better able to absorb nutrients in the soil, and the like, are a desirable outcomes of crop biotechnology. Similarly animals bioengineered to withstand hostile climatic conditions or resist diseases while continuing to produce milk, eggs, or lean muscle tissue are important goals for animal biotechnology. Plant varieties or animal species so engineered are in less need of constant management of inputs and external conditions. Therefore biotechnologically improved plants and animals should give agricultural producers more control over their production processes. The most desirable products of biotechnology for industrialized farms thus are bioengineered plant and animal species.

In the absence of a plant variety that is high-yielding and resistant to pests or climatic stresses, a second level of agricultural biotechnologies is desirable. These are genetically engineered organisms or biotechnologically produced substances that assist the plant in resisting pests or diseases or taking up nutrients from the soil. *Bacillus thuringiensis* (Bt, mentioned above) is one such product. Bt is a bacterium, engineered so as to be deadly to various species of caterpillar that are harmful to vegetable (tomato, bean) plants but not harmful to the plant nor other species (including humans). Organisms that prevent frost damage to fragile young plants (e.g., potatoes, strawberries) have also been bioengineered. For crop farming, these "external" control agents (sprayed on crops) are obviously less desirable than a crop with pest resistance or frost tolerance "built in," but they are nevertheless useful in helping control the environment in which production occurs.

For animal agriculture, the second tier in biotechnology is similar: organisms or chemical substances that might be injected or added to feed in order to help prevent disease, augment nutrition, and increase control of milk or egg or meat production. *Bovine somatotrophin* (bST, also called bovine growth hormone (BGH)) is one of these. When bST is administered to dairy cattle, it increases milk output without a corresponding need for increased animal feed intake.

A third tier of agricultural biotechnology is concerned with postharvest control over the products of agriculture, whether plant or animal. Though some of these directly benefit farmers, most are designed to reduce spoilage in vegetables and grains and keep animal products fresh and safe during transportation and marketing. Some are designed to help control or speed up processing, for example, bioengineered enzymes that are more efficient than traditionally used biochemicals in the fermentation of cheeses or beer. While some of these third-tier products have become available for commercial use in transportation, processing, and marketing, most postharvest biotechnology products are still in the developmental stage.

In fact most of the desired agricultural biotechnology products and processes are still in the developmental stage, and many will never reach commercial applicability. Nevertheless, the energies and monies invested in agricultural biotechnology suggest that the long-touted "biotechnological revolution" in agriculture is quite possible. Given the longstanding goal of increased productivity and its corollary, increased control in the production–processing–distribution system, increased agricultural biotechnology is desirable at least for its corporate producers. This suggests an additional issue of control associated with agricultural biotechnology. That is the potential control of agriculture, including both large and small farms, associated with the corporate entities who are by and large the major proponents and suppliers of agricultural biotechnology.

Biotechnology, Corporations, and the Family Farm

Technological change and economics of scale have placed small- and medium-sized farms (both of which most likely to be family farms) in a precarious position. According to the theory of the "technology treadmill," farms must be able to adopt the latest, efficiency-enhancing technologies as these technologies emerge from the research and development process. Large industrialized farms are most likely to be able to keep up with the acquisition of "new and improved" technologies. Middle-sized and small farms must either figure out ways to move toward adoption, or fall off the treadmill. Small farms are most likely to fall.

The issue with agricultural biotechnology is the same. Some of biotechnology's real and potential products and processes for agriculture may be scale neutral, equally capable of benefiting small and large farms. However, it has been claimed that most, including, for example, bST, are not scale neutral, and appear to be *designed* for large farming operations. Research and development of agricultural biotechnology in universities (especially in colleges of agriculture) has been criticized for being financially and politically influenced by large-scale farm operators. What large farms desire is what is produced.

While the intention may not be to deliberately harm small or medium family farms, the nature of the products and processes developed does so anyway by exacerbating the technology treadmill.

More significant is the fact that final product development, production, and marketing of biotechnology products and processes is in the hands of large, in many cases, multinational corporations. In fact much of the university-based research is being funded by those corporations, which include pharmaceutical firms, seed companies, petrochemical giants, and global food-and-fiber products distributors. The primary concern of these agribusiness firms is to increase profits and market shares for their products, and one way of doing so is to bring "new and improved" agricultural biotechnology to the market as quickly, and as often, as possible. In this regard corporate producers of agricultural biotechnology can control the speed and direction of the technology treadmill. In so doing, these corporations effectively control the further industrialization of agriculture, with attendant implications for the restructuring of agriculture and the fate of the family farm.

There are two likely outcomes of the way in which agricultural biotechnology is currently being developed and commercialized: Only large-scale farming operations will be able to afford them, and only large-scale enterprises will be able to efficiently and effectively employ the products. The structural effects of these outcomes will be reinforced by other current trends. For example, in the broiler chicken industry, producers must purchase inputs (chicks, feed) from a given firm and sell their products back to that firm under a contractual tie between corporate actors and on-farm producers. The small grower who is unable either to afford the inputs or able to guarantee a particular quantity at a particular weight at a fixed price is effectively excluded from the market. With no viable market for its products, the smaller farm will fall out of agricultural production (other things being equal). In the case of some of the new agricultural biotechnology products, a similar contractual arrangement is taking place. Growers who want to plant a particularly high-yielding soybean hybrid bioengineered for tolerance to a particular herbicide must agree in writing to purchase both seed and the herbicide from the same corporation. This may be contrary to growers' best interests, given the fact that the herbicide's patent protection has run out and other companies are producing identical herbicides at a lower price. Moreover the seed-herbicide "package" is expensive, effectively limiting its market to large producers.

The nature of agricultural biotechnology's products then, combined with the fact that these products and processes will be the domain of large corporate actors in the food system, does not portend well for family farms. Indeed, it may raise concerns for all of agriculture as even large-scale farms find themselves increasingly tied to nonfarm agribusiness corporations. Biotechnology per se may not be the ultimate cause of further industrialization or increasing bipolarization in the structure of agriculture. Agricultural biotechnology may not be the cause of the ultimate demise of the family farm in the United States. Nevertheless, as a further and soon to be more pervasive tool in the toolbox of agricultural researchers, on-farm producers, and corporate producers and suppliers, it is and will continue to be a contributing factor to the decrease in the competitiveness and viability of small family farms. Given the fact that only 20 percent or so of the large farm producers account for over 80 percent of farm output and sales in the United States, there is little incentive for the producers and marketers of agricultural biotechnology to focus their research and development efforts on small farms' needs or interests. Instead, the trend is likely to continue that the big will get bigger and the small will be placed in jeopardy if not driven to extinction. This trend will be aided by the researchers and corporate actors who have a vested interest in the success of the "biotechnological revolution" in agriculture, that is, if the trend is not halted by public policy or changes in consumer tastes and preferences.

CONCLUSION

Agricultural ethics is the systematic application of disciplinary tools from philosophy and ethics to the problems and issues of agriculture. It focuses on specific problems as well as on the ethical aspects of agriculture or the agricultural/food system as a whole. Agricultural ethics, as in other areas of applied ethics (business ethics, medical ethics, etc.), is not simply about identifying issues and concerns, however. That would be only half its philosophical task, that is, the *description* of problems, conflicts, values, and orientations in agriculture — *descriptive ethics*. The other more important task is prescriptive (proscriptive), the work of *normative ethics*. Although agricultural ethicists may be in no authoritative position to tell farmers, agribusinesses, consumers, or public policy makers what to do concerning such things as using pesticides, managing livestock, or adopting new technologies, it is nevertheless part of the responsibility of ethicists to articulate the normative implications of actual or potential decisions. If, for example, there is general agreement that people have an obligation to future generations to leave a habitable environment (and there seems to be some such consensus), then when an ethicist shows how a particular agricultural or natural resource-related practice potentially endangers the environment, the ethicist "proves" that the practice is unethical and therefore must be stopped. The difficulties are (*1*) finding the consensus we individually and collectively might have regarding ethical obligations and (*2*) accurately describing and analyzing the facts and ethical implications of specific actions or general practices.

In the discussion of the utilitarian justification for industrial agriculture, for example, a conclusion was drawn that "quantity, qaulity, and price" represent widely held and ethically acceptable goals for agriculture. This is a legitimate inference, because with only a few exceptions related to the environment or food safety, public actions and public policy have not challenged agriculture so long as QQP has been achieved. The task for the agricultural ethicist in this case is to carefully analyze actions or

developments in agriculture to see if there are any that may be inconsistent with QQP in the present (e.g., the use of some pesticides threaten food quality even if unbeknownst to consumers) or in the future (e.g., the control over agricultural production by a few corporations raises the prospect of monopoly pricing). The normative implications of finding such inconsistencies should be clear: From a utilitarian perspective, inconsistencies between practices or trends and QQP (or whatever goals the public has for the food system) entail that those practices or trends are ethically wrong and should be corrected or stopped.

Rights-based analysis involves similar reasoning. Suppose that we can identify a widely held rule or principle concerning peoples' rights as human beings or as participants in a democratic, market-based society, This rights-defining rule or principle may be implicit in current policies or public thinking, or may be only very general and vague, for example, "People should have equal opportunities" or "People should be treated fairly." However implicit or vague, if we can also show that something is happening in agriculture that may undermine or infringe on those rights or entitlements, we have arrived at a basis for judging that this particular action, practice, or trend is ethically unacceptable.

In sum, it is not the practice of ethics or agricultural ethics to preach or dictate. Rather, it is to show that *if* there are dimensions of agriculture — including how agriculturalists are themselves treated by nonagricultural actors including governments — that are at odds with basic ethical principles such as utility maximization, fairness, or respect for the future, *then* these dimensions should be seen for what they are — unethical. It is up to the relevant actors to decide on the basis of this moral knowledge whether or not they will do the right thing.

In any event, there is one thing normative ethical analysis, in agriculture or elsewhere, a priori rejects: Retreat to the claim "that's the way it is." Reviewing past occurrences can give us perspective as to how present circumstances have developed. However, we must be careful to see "developments" or "trends" for what they are: the collective results of individual decisions. Sometimes those collective results are unintended. A farmer purchasing a tractor or a sack of hybrid corn seed in the 1930s could not have predicted that 50 years hence the development of the tractor or hybrid seed would be early occurrences in the process of transforming farming into a high technology, international, corporate-controlled agribusiness system. Nevertheless, in purchasing that tractor and seed, the farmer made a decision that affected the development of agriculture into the system before as today. While normative ethicists cannot ask decision makers to acquire predictive powers in all their actions, ethicists can ask, indeed demand, that each of us be more circumspect in our decisions and actions about the potential grand-scale outcomes of small decisions. Trends begin and end when individuals begin or end them.

From a normative ethical perspective, we are left with a set of questions about the family farm, the industrialization and restructuring of agriculture, and the role of biotechnology and the purveyors of biotechnology.

- Has the industrialization of agriculture over the last century been a good thing? For whom? According to what ethical criteria (rights, utility, etc.)?
- Is the continuing industrialization a good thing? Again, for whom and on what ethical basis?
- Is it ethically justifiable that small family farms have been the major losers in the process of the industrialization of agriculture? On what basis?
- Indeed, does society have any obligations to small family farms?
- Are the new agricultural biotechnologies (individually or as a whole) good for agriculture? Are they good for society? On what grounds?
- Are there ethical problems associated with the research, development, or marketing goals or strategies of the biotechnology entrepreneurs?
- What should society's ethical judgment be regarding the new agricultural biotechnologies and their governmental and corporate "sponsors"? What should the public's response be?

These are the big questions to consider as the industrialization of agriculture nears completion in the United States and most other developed nations. They merit ethical reflection.

BIBLIOGRAPHY

1. R. Carson, *Silent Spring*, Houghton Mifflin, Boston, MA, 1962.
2. R. Harrison, *Animal Machines: The New Factory Farming Industry*, V. Smart, London, 1964.
3. W. Berry, *The Unsettling of America*, Sierra Club Books, San Francisco, CA, 1977.
4. P.B. Thompson, R.J. Matthews, and E.O. Van Ravenswaay, *Ethics, Public Policy and Agriculture*, Macmillan, New York, 1994.
5. J. Hightower, *Eat Your Heart Out: Food Profiteering in America*, Crown Publishers, New York, 1975.
6. M. Strange, *Family Farming: A New Economic Vision*, University of Nebraska Press, Lincoln, 1988.
7. L. Busch, W.B. Lacy, J. Burkhardt, and L.R. Lacy, *Plants, Power, and Profit: Social, Economic and Ethical Consequences of the New Biotechnologies*, Basil Blackwell, London, 1991.
8. Biotechnology Industrial Organization (BIO), 1998. Industry website: *http://www.bio.org*
9. BioTech Navigator, 1998. Industry online newsletter: *http://www.biotechnav.com*

See other entries AGRICULTURAL BIOTECHNOLOGY; ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE.

AGRICULTURAL BIOTECHNOLOGY, ETHICS, FOOD SAFETY, RISK, AND INDIVIDUAL CONSENT

PAUL B. THOMPSON
Purdue University
West Lafayette, Indiana

OUTLINE

INTRODUCTION

Agricultural biotechnology has spawned heated controversy. While human and biomedical applications raise clear ethical issues, the ethics of food biotechnology are less obvious, and concern may appear entirely misguided. Changes to food stimulate reactions of intense emotion, resentment, and resistance among some consumers. These reactions are themselves complex, combining feelings based on religious beliefs about genetic technology in general, environmental impact, sympathy for animals, or solidarity with small farm, organic and sustainable agriculture movements with genuine concern about the hazards and uncertainties associated with the consumption of genetically modified foods and food products. For some individuals, fear and resentment about human gene technologies may be manifest in attitudes toward food biotechnology (1,2). Whatever the ultimate basis, these concerns about agricultural biotechnology often surface as concerns about consuming food (3).

Scientifically and analytically trained observers of the debate apply a narrower set of criteria to the evaluation of food safety. Even when concern about environmental impact is seen as valid, for example, experts do not translate this into a legitimate basis for concern about the consumption of food. As such, experts often dismiss the broader public's reaction to food biotechnology as muddled, uninformed, and even irrational (4,5). Yet it should not be surprising that public resistance to genetic technology might center on food. Most people will rarely or never face a personal choice about the more exotic and ethically troubling applications of genetic technology, and few may be willing to undertake what is necessary to influence policy in a political forum. Everyone, however, makes food choices everyday. It is natural that many of those who experience the greatest anxiety and moral opposition to genetic technology would express their concern most vehemently with respect to food.

The inconsistency between expert and lay approaches to the risk of genetically modified foods is the overarching philosophical issue for food safety. Experts understand food safety almost exclusively as a problem framed by the conditional probability of injury or disease as a result of consuming a food product, either once or over a lifetime. For the expert, issues are relevant to food safety only to the extent that they affect these probability estimates. To be sure, the expert approach to risk entails ethical questions of its own (discussed below). Nevertheless, it is reasonable to conclude that the likelihood of injury or disease from eating genetically modified foods is low, and can be reduced to minimal levels with a modicum of cautionary practices. To most experts, this conclusion establishes a sound basis for repudiating many of the lay public's concerns about consuming genetically modified food. However, a different philosophical interpretation of risk would countenance a much broader array of factors, and would make each individual the sovereign judge of whether the interests threatened by genetically engineered food are vital. Such an approach would emphasize an individual's right to choose whether or not to consume genetically engineered food. Safety, in the sense of probable harm, would clearly be relevant in this alternative view, but the appropriate ethical response to risk might be to secure conditions of individual consent, rather than to minimize probability of injury or disease.

There has been comparatively little discussion of these issues by professional bioethicists. Some of the key philosophical questions can be highlighted, and a sketch of each of the two contrasting approaches to risk can be made. Perhaps the crucial policy questions revolve around labeling of genetically modified foods, but legal and economic issues contend with ethical concerns in fixing labeling policies. Even the definition of a "genetically engineered food" is open to debate.

BIOTECHNOLOGY AND FOOD

Recombinant DNA techniques have many applications in the food system, and many of the key terms are ambiguous. The phrase "whole foods" indicates foods that have not been combined or adulterated; other foods are "processed foods." Apples, beef, whole-wheat flour, and whole milk are whole foods. Sausages, breads, and cheeses are processed. However, some foods, such as fruit juices, vegetable oils, or skimmed milk, fit uncomfortably in this simple dichotomy. The term "constituent" will be used to identify the various parts (fat, fiber, vitamins, etc.) of whole foods, while "ingredient" will be used to indicate a food or additive used in processed food. "Contaminants" such as pesticide residue, insect parts, or fecal matter are unintentionally incorporated in both whole and processed foods. "Adulterates" are impurities that have been intentionally introduced, generally to enhance the bulk or appearance of a commodity. Regulatory agencies approve certain substances (e.g., dyes) as additives. Nevertheless, philosophical controversies arise because one person's additive is another's adulteration.

The term "genetically modified organism" (GMO) is used commonly to include whole foods from plants or animals whose germ plasm has been modified using recombinant DNA techniques. The expression "genetically modified food" would normally be given a broader interpretation to include processed foods with GMO ingredients. Such processed foods would involve consumption of a GMO as an ingredient, though the chemical properties of the GMO may be substantially

affected by processing. The use of genetic engineering in the food chain, however, is often indirect, as in the case of recombinant rennet, where a bacteria is modified to produce an enzyme used in making cheese. The use of terminology in such cases can be controversial, but here we will refer to all cases with the phrase "GMO." However, all foods are "genetically modified" in a broad sense. Virtually all plants and animals consumed for human food are the product of crossbreeding and genetic selection. Such practices of genetic modification have been used since antiquity, and for our purposes they do not count as creating a GMO.

Key ethical and philosophical judgments depend on the interpretation of terminology. One use of genetic engineering is to develop DNA probes designed to increase the accuracy and efficiency of meat and produce inspection. The availability of such probes will significantly change the procedures for monitoring food pathogens and enforcing food safety regulations. These changes in procedure may, in turn, have an impact on food safety. DNA probes are, indeed, intended to enhance food safety, though whether this result will occur can be disputed. However, DNA probes do not in themselves alter the composition of foods on which they are used. At the opposite end of the spectrum are genetically modified plants and animals. Maize and soybean, for example, have been modified through genetic engineering to make the crops resistant to plant viruses or the use of herbicide, and to synthesize the naturally occurring substance bacillus thuringensus (Bt), which is toxic to leipedoptera. The grain and oil derived from these genetically engineered crops are consumed directly by humans, and indirectly when the crops are fed to animals who in turn produce milk and meat products.

A number of other food technologies stand in between diagnostics, which do not alter the composition of food, and genetically engineered crops or food animals, which clearly have been altered. For example, recombinant rennet, the enzyme essential to cheese making, was one of the first commercial products from genetic engineering outside of medicine. In nature, rennet is produced within the gut of calves. Traditional rennet for cheese making has been harvested from the entrails of slaughtered calves. Genetically engineered bacteria modified to synthesize the enzyme produce recombinant rennet in a process similar to fermentation. Other bacteria have been modified to synthesize bovine somatotropin (BST), which can be administered to lactating dairy cows as a stimulant to milk production. Such modifications can reasonably be interpreted as changing the constituents of food, though they do not involve genetic manipulation of the organism that is thought of as the food source itself.

DNA itself is present in virtually all foods and food ingredients, excepting only minerals (e.g., salt) and water. The high temperatures used in many forms of food preparation destroy DNA. Fresh fruits or vegetables and raw milk or meat contain DNA in the forms in which they are normally consumed. New food biotechnologies thus do not introduce DNA into the human food chain. DNA is nontoxic and is thoroughly metabolized in normal human food consumption. Some applications of food biotechnology result in direct human consumption of altered genetic material. Some do not. Alterations in genetic material are normally made to affect an organism's production of proteins, or to affect the regulation of the organism's cellular processes. Such changes affect the organism at a phenotypic level. Phenotypic alterations could affect the suitability of an organism for use as food.

Whether some, none, or all of these applications trigger ethical issues depends on pragmatic circumstances. In regulatory contexts, authority may be constrained by legislation or bureaucracy. Regulatory decision makers must decide whether a DNA probe or an intermediate product (e.g., recombinant rennet or BST) is defined as a food, an additive, an ingredient, or a contaminant in order to know which regulatory criteria to apply. Often such a judgment determines which agency or division has jurisdiction. If reducing public exposure to injury or disease is the overriding objective, applications that affect food safety inspection procedures are more significant (for good or ill) than genetically modified crops such as herbicide tolerant soybeans or Bt maize. An entirely different set of criteria may be appropriate when the circumstances that lead to individual consumption of a genetically modified substance are at issue, rather than general public health. For example, a person who believes that religious dietary rules prohibit the consumption of genetic materials derived from specific animals will be far more interested in the source of genes that ultimately find their way into the food chain than in the material impact of biotechnology on the probability of disease or injury.

FOOD: SAFE, PURE, AND WHOLESOME

Dietary rules have occupied a minor place in religious and philosophical discourse from time immemorial. Semitic dietary rules are well known. Christian rules have been associated with specific rituals and seasons such as Lent. In ancient Greece the Pythagorean cult of which Plato was a member had a rule against eating beans. In addition to explicit rules, all human cultures adopt implicit beliefs about what is and is not considered edible, and in what combinations or at what seasons edibles may be eaten. Though the basis and meaning for food regimes is a subject of debate among anthropologists, there is no doubt that such regimes fulfill a minimal social function. Every human society must have some means for avoiding poisons. Such knowledge directs ordinary food choices to plants and animals that are not acutely toxic, and encourages practices of harvest, storage, and preparation that minimize risk. Furthermore, since the type and availability of edibles will vary according to season, location, and climatic conditions of drought or pestilence, this knowledge must be reproduced from day to day, year to year, and generation to generation. Any successful human society will have developed a food regime that satisfies these conditions (6,7).

Food regimes thus represent an implicit knowledge system of enormous complexity, and one that is highly sensitive to technological transformation. Claude Lévi-Strauss attributed deep significance to the emergence of cooking, suggesting that a culture's entire system of

signs is rooted in this fundamental food technology. Lévi-Strauss's structuralism represents a view that might be interpreted to entail hidden meanings wrapped up in food beliefs and the potential for serious ethical considerations when that system of belief is challenged. Yet even without structuralism's backing, it is easy to see why beliefs about what is and is not food would become deeply interwoven with a culture's ideas about purity, hierarchy, and the sacred. It is also easy to see why such beliefs would become a minor battleground when distinct cultures come in contact with one another and why minorities would nurture food regimes as components of cultural identity. Given this background, the emergence of the modern system for state regulation of foods is one of industrialization's more remarkable achievements.

Pythagorean rules on beans aside, philosophers have largely neglected food regimes, and the history of food safety during industrialization has yet to be written in a definitive fashion. Any concise overview is thus necessarily speculative to some degree. Over a few centuries, industrialized societies evolved a conception of food safety in which impurities were thought to be the primary cause of food-borne illness, a view that evolved into the germ theory of disease in the nineteenth century. Avoiding germ contamination was consistent with a traditional emphasis on pure and wholesome foods. In both traditional and early industrial food regimes, pure and wholesome foods are "good for you" in the broadest sense, meaning that they promoted physical health, a positive mental outlook, and were socially acceptable. In traditional societies, this was assured by following implicit or culturally based rules about what could and could not be eaten. Under industrialization, "pure and wholesome" foods were those not contaminated or adulterated by germs, residues or foreign substances (7).

In the emerging industrial food regime, science was deployed first as a means of identifying impurities, and then of measuring the risk associated with them. Regulatory policy was to protect the food supply from adulteration by contaminants, and to weigh the benefit and risk from additives and residues, removing offending substances that posed significant risk. Many members of the lay public probably retain this mental picture of food safety today. It shares two important features with traditional food regimes. First, the determination that a substance is or is not food is the primary basis for a judgment of purity. Second, risk is associated with a compromise in purity. Until comparatively recently, most people would probably have also presumed a strong link between purity and wholesomeness, or between nonadulteration of a food and its nutritional or social acceptability.

Toxicologists, nutritionists, and other experts on food have so modified this picture during the last quarter of the twentieth century as to constitute its gradual abandonment. Scientists have put complex interaction among the chemicals that make up whole foods in place of the idea that purity is equivalent to safety (8). Dietary induced cancer is now thought to be caused by interaction between two groups of food constituents, mutagens and antimutagens, along with the genetic disposition of the individual consumer. Furthermore mutagens and antimutagens are thought to be found in virtually all foods. The more general link between nutrition and health has also shifted from eating pure foods derived from groups, to the proper balance of fats, protein, carbohydrate, fiber, and other food constituents. Most whole foods contain a combination of all these constituents. This reductionist view of food safety has made some experts skeptical of regulatory approaches that require identification and elimination of alleged carcinogens. It plays an important role in the reasoning of scientists who evaluate the safety of genetically modified foods.

At the same time physicians are diagnosing more and more individuals as suffering from allergic reaction to specific foods or to chemical sensitivities. True allergies cause toxic and sometimes fatal reactions, usually within a few hours of consuming the allergen. Food sensitivities such as lactose intolerance cause less critical reactions such as gastrointestinal distress. Individuals with specific allergies or sensitivities have reasons to avoid foods of uncertain origin, and they have raised a number of questions about the effects of genetic modification. However, some experts question the rising tide of allergy diagnoses, claiming that most people with sensitivities can consume moderate amounts of the given food, especially in combination with other foods. There is thus controversy as to whether the increase in allergy diagnosis is biologically, psychologically, or culturally based, and an individual's view on this question often colors their assessment of genetically modified food.

In summary, the starting point for any discussion of food safety and biotechnology is an environment where science has outrun cultural attitudes on purity and wholesomeness, attitudes that have evolved as the primary social basis for food safety over centuries, if not millennia. The science itself is dynamic, and debates are frequent. With this background in mind, it is remarkable that there is relatively little debate among food safety experts about the low probability of injury or disease from consuming GMOs. Yet it is not surprising that the broader public debate over GMOs should be affected by cultural and technical disputes that attend food safety generally.

RISK, SAFETY, AND DELIBERATIVE RATIONAL CHOICE

Experts approach food safety as a problem of risk assessment framed by the parameters of deliberative rational choice. Some experts also presume a public health philosophy that reflects a utilitarian approach to public policy. Deliberative rational choice presumes that decision makers see each course of action open to them as a means for bringing about consequences. These consequences can then evaluated in such a way that each course of action can be understood as having an expected value. A given course of action commonly has two or more possible outcomes. In such cases the expected value of a given course of action is derived by considering both the expected value of each possible outcome, and the likelihood or probability that it will occur. In common parlance, the *risk* of a given course of action is a function of the probability that unwanted consequences will occur, and the harm or loss associated with those consequences (9).

Deliberative rational choice can be applied to many different kinds of choice situation. At the policy level, the choice might concern whether to allow *any* genetically modified foods on the market or whether to allow a specific application of the technology (e.g., recombinant rennet) on the market. In either case, the *risk* of the policy would be found by assessing both the likelihood of unwanted consequences and the harm or loss expected to be associated with those consequences. Individuals might also apply deliberative rational choice in making individual food decisions. Here each food purchase or consumption decision might be evaluated as having an expected value (10–12). The risk of purchasing or consuming a genetically modified food would, for the individual consumer, be a function of the probability that unwanted consequences would occur, and of the harm or loss associated with those consequences, should they occur.

Regulators, scientists, and others who have contributed to the literature on food safety appear to be applying a framework of deliberative rational choice, though few say so explicitly. They assume that consumers either apply or intend to apply such a framework in their food choices. Experts contributing to this debate rarely question the assumption that consumers see food choice in terms of instrumental rationality. Policy choices that end in general dietary health and no increase in the rate of dietary diseases or disorders are assumed to meet all relevant ethical criteria. Such policies are seen as consistent with public health and consumer wishes, and there is rarely any acknowledgment that these two criteria could diverge.

Approaches to Food Safety

Given this general approach to risk, one might develop any of several approaches to food safety, but three paradigms dominate debate over food safety. *Risk thresholds* are routinely applied to determine criteria for food safety in most industrial countries. In the United States the Food and Drug Administration applies threshold criteria to genetically modified food. *Risk-benefit* averaging adapts broadly utilitarian criteria to administrative decision making. It differs dramatically from a threshold approach in conceptual terms, but practical differences consist in the fact that benefits are taken into consideration along with risk. A third approach would allow *market forces* to determine acceptable levels of risk.

Threshold Approaches. A food might be deemed safe if the risk of any policy decision to allow it on markets falls below a given threshold. Note that the level of risk might be driven below a threshold by many different characteristics.

- An event that occurs relatively frequently but with trivial harm or loss might be considered "low risk."
- An event that occurs very infrequently may be considered "low risk," even if harm or loss in those infrequent occurrences is comparatively significant. In particular, an event likely to affect only a small percentage of the population may be assessed as "low risk," even when it is very likely to affect that small minority.
- Even the potential for catastrophic effects may not bring a risk above the threshold, if the probability of catastrophe is sufficiently small.

In practice, however, food safety policy makers have adopted very conservative approaches to setting thresholds. The U.S. debate over thresholds has centered on the Delaney Clause, which stipulates that no carcinogenic substances may be used as additives. Regulators have interpreted the Delaney Clause as requiring "zero risk," or a threshold of zero for acceptable risks.

The zero-risk threshold has created enormous problems for regulators. Given the general parameters of deliberative rational choice, regulators must consider any scenario that leads to an unwanted outcome in making a risk assessment. The only way that such a scenario can be found acceptable under a zero-risk threshold is to prove that there is a zero probability of its occurrence. But statistical methods do not support such a proof. As such, regulators have treated zero risk as "no measurable risk." Even this approach has become problematic as the general philosophy of food safety has become more reductionistic. It is increasingly difficult to design experiments that could establish meaningful probabilities, while eliminating the confounding effects of other mutagens and antimutagens that are natural constituents of food. Regulators and experts share a general consensus that the zero-risk threshold is impossible to support with modern scientific methods (see (13) for a related discussion on workplace hazards).

Risk–Benefit. Frustration with thresholds has led to a surge of interest in risk–benefit criteria. In this approach a decision maker must weigh both risks and benefits in assessing policy. Economists and nutritionists would be consulted to assess the benefits of a genetically modified food. Foods that return economic benefits to consumers, farmers and the food industry, or that improve nutritional quality might be found acceptably safe, even when risks are nonzero, or even above *de minimus* levels. Such a standard is currently used to assess risks associated with chemicals regulated under the U.S. Federal Insecticide, Fungicide and Rodenticide Act (FIFRA). A risk–benefit approach to genetically engineered foods could be given one of several interpretations:

- Safety could be deemed acceptable whenever benefits outweigh risks.
- Policies could be required to produce the optimal ratio of benefit over risk.
- Policy could reflect a "mixed mode" of thresholds and consideration of benefits, so options entailing significant risk would not be considered, no matter what level of benefit might be associated with them.

Additional technical parameters would have to be specified before a formal risk–benefit can begin, as well.

Risk–benefit is attractive in part because it seems more consistent with the utilitarian philosophy that some associate with deliberative rational choice. Though utilitarians differ over how to assess the expected value

(utility) of a choice, they would advocate an approach of predicting consequences for all affected parties, assigning utility or value to these consequences, then selecting the course of action that maximizes utility. The emphasis on including consequences for all affected parties provides a rationale for considering benefit as well as risk. The struggle over which way to interpret risk–benefit requirements applied to GMOs would be seen as part of a larger problem in interpreting the maximization requirement: Is it maximal total utility, is it average utility, or is there some hybrid notion?

Market Solutions. Both approaches described above assume that administrative decision makers will assess risk and apply a policy decision rule to determine when and whether consumers will be exposed to risks associated with GMOs. An alternative approach would stress providing individual consumers with information about risks, and then letting market forces determine the level of acceptable risk. In practice, such an approach has many pitfalls. It would require difficult and contestable judgments about how to provide information on the probability of disease or injury that might be associated with any consumer's food decision. Furthermore, since consumers are in the habit of assuming that foods on grocery shelves meet standards of food safety, it is likely that they would be slow to apply a more critical approach to risk in their individual food choices. There is also doubt that the average consumer has the ability to make appropriate risk judgments.

For all these reasons, market solutions are often treated more as a foil in technical debates over food safety than as a serious alternative. However, experts tend to presume that those who advocate labels for GMOs are advocating a market solution (14–17). As such, they evaluate the question of labels in light of standards derived from deliberative rational choice. Labeling is seen as a policy option that should be evaluated in light of whether individuals use information on labels to satisfy their personal values with respect to risk, safety, price, taste, and the other characteristics affecting the expected value of a food decision. The mere occurrence of a label might lead some consumers to assume that a food is risky, even when the substantive information would support a different comparison. Labels for GMOs thus have a clear potential for to suboptimal policy performance from a deliberative rational choice perspective. This theme will be revisited again in the discussion below.

Assessing Risk

Deliberative rational choice demands an assessment of the probability and value of the consequences that might be expected to ensue from general public consumption of GMOs. Such assessments require epistemological and methodological judgments that have ethical implications. Although only a few authors have addressed the questions of food safety for GMOs, there is an extensive philosophical literature on risk and probability that is relevant to this general problem. First, risk assessors must settle questions about the *interpretation of probability*. Second, they must have a general philosophy for distinguishing *risk and uncertainty*, and must have norms for responding to each. Third, they must make assumptions about minimizing *type I and type II errors*. Finally, they must decide between *formal and informal* risk assessment.

Interpretation of Probability. Three general approaches to probability can be found in the literature, and considerable innovation in the philosophy of probability has taken place in the last decade. Classical probabilities are derived from formal properties of the systems under study. An ordinary die has six sides, and a "fair die" can be defined as one for which the likelihood for each face turning up on a given role is equal, or 1/6. From this one can derive probabilities for combinations arising from several dice thrown at once. Relative frequency is a probability stated as the frequency with which a given outcome occurs in a given population of trials. Subjective probability refers to the confidence or expectation that a given person has that a predicted event will occur, or that a given proposition obtains.

Classical probability provides a formal specification of probability that permits substantial development of statistical theory. One may then treat both relative frequency measurements and subjective probabilities as situations where the analyst simply lacks an adequate specification of the formal system. Bayes's theorem provides a way to combine relative frequencies with subjective judgments, as well as to update results from several trials. Uncertainty then reflects a measure of how likely it is that any given statistical measurement is wrong, given available evidence. Although classical probability is generally thought to be inapplicable in most problems of empirical risk assessment, statistical theory and methods generally render philosophical questions about which theory of probability is the true one moot in practical situations.

Nevertheless, the potential for subjective interpretations of probability opens the door for dispute about the legitimacy of any given risk assessment. Whose subjective judgments are to count? Why should expert judgments supplant lay judgments? These are easy questions to answer in a purely instrumental context. Experts generally make more reliable judgments in virtue of their expertise. Yet the questions may be asked as a challenge to the authority of experts in making regulatory decisions that affect the lay public. Here the above-mentioned assumption that consumers are themselves deploying a deliberative rational choice model becomes crucial, for if that is so, relying on expert judgment may result in better choices than if individuals make their own judgments about risk. Even a resort to relative frequency does not settle this issue, for there are value judgments embedded in how to construct and evaluate the populations on which relative frequency trials are based. Such concerns might be more productively pursued within the framework of a general rejection of some assumptions common in deliberative rational choice (18).

Risk and Uncertainty. Although statistical theory provides one approach to uncertainty, it demands evidence or assumptions about the relationship between observations

and the total population of instances for which observations might be made. In practical terms, this means that one must have at least speculative knowledge of the mechanism or correlation thought to lie at the root of a risk. This approach to uncertainty does nothing to address the possibility that there has been some scenario or possibility that no one has thought of. Yet entirely unknown risks do materialize. When scientists developed the feeding strategies that are now thought to have led to the variant of encephalopathy associated with mad cow disease, prions were not known as potential risk agents. It is clear that many who raise questions about uncertainty associated with human consumption of GMOs through the food chain are referring to this kind of uncertainty, rather than the sort addressed through standard statistical approaches.

Though it would be difficult to say how such uncertainty should be measured, it is common to apply relative quantitative judgments to the uncertainty that surrounds GMOs as compared to traditional foods. Common practice would also lead to the judgment that practices associated with great uncertainty are thereby "risky." Given the vagueness and unmeasurability of this uncertainty, it must be approached through subjective probability, if it is to be incorporated into deliberative rational choice at all. Here, again, the problem of whose judgments to use (discussed immediately above), and whether Bayesian techniques may be applied, reasserts itself. In general, philosophers have concluded that experts are prone to dismiss uncertainties too quickly, and to conclude that highly novel practices are inherently risky. For their part, risk experts have tended to demand at least a plausible scenario for how an unwanted event might transpire before they will seriously consider reviewing it in a risk assessment (18).

Type I and Type II Errors. In standard scientific research, conservative practice demands that a result be rejected unless uncertainty (in the technical sense) can be reduced to the point that the research is 95 percent certain that the result is true. Accepting a false result is a type I statistical error. A type II error occurs when one fails to accept a result that is, in fact, true. In risk assessment, early results often indicate risk but are not adequate to corroborate that result with 95 percent confidence. Should scientists minimize the chance of a type I error, and withhold judgment until further studies are done? Or should they minimize the chance of a type II error, and announce that a risk is present, even though future studies may well show that it is not?

Within the regulatory system for genetically engineered foods, decision makers would certainly apply a principle of caution (minimizing type II error) prior to the approval or release of a given product. The situation can be far more difficult if evidence for risk appears after a substance is already on the market. Here the potential for needless panic and economic loss often leads regulators to regard relatively unconfirmed results pointing toward risk as premature. Though no clear cases of type I/type II dilemmas have yet occurred in the food safety regulation of GMOs, uncertainty is itself treated as evidence for type II error by some critics. It is not clear how regulatory agencies would or should handle such cases (19).

Formal or Informal Risk Assessment. Risks for pesticides, food additives and drugs are subjected to laboratory and clinical research trials that establish measures of risk. Many have argued that such formal procedures introduce unnecessary costs in the development of a GMO. Instead, they argue that adequate risk assessment can be done simply by reviewing the nature of the planned alteration subjectively. Reviewers would note that interventions involving known allergens or that had a known potential to create new proteins in food should undergo formal risk assessment, but other GMOs would pass directly into the food chain. Reviews would be done by individual researchers or by Institutional Review Boards (IRBs) at universities, laboratories, and within the food industry. This is, in effect, the procedure currently used for assessing the food safety risk of consuming organisms developed through conventional plant and animal breeding. Critics of agricultural biotechnology have vociferously opposed the informal approach.

These two views split two key value assumptions of deliberative rational choice. On the one hand, deliberative rational choice is supposed to be deliberative, which would imply careful review and objective assessment of all risks for all affected parties. On the other hand, rational choice is supposed to produce the best outcome, and if otherwise acceptable products are never developed because process of policy approval is too costly, that is a result hardly in keeping with its spirit, (5,20–22).

RISK, PURITY, AND CONSENT

Traditional approaches to food choice combine the safety of food with culturally based judgments about purity. Anthropologists have studied these approaches, but no philosophical literature articulates the ethical principles on which they rest. One possibility is that people are attempting to emulate deliberative rational choice through their food traditions. An alternative possibility is that individuals and groups have been thought to have the right to apply whatever standards of purity they deem appropriate in making food choices. This alternative view finds philosophical support in two complementary positions. First, purity norms may function to promote rational ends, but through a nondeliberative mechanism. If so, it may be rational to rely on purity rules, even when deliberative calculations indicate otherwise. Second, many bioethicists have long argued that risks may only be imposed on subjects with their consent. Together, purity rules and consent criteria establish a procedural burden of proof for the safety of genetically engineered foods and food products.

Risk and Purity

Anthropologist Mary Douglas has approached risk from the standpoint of cultural norms that establish the most basic categories of acceptable behavior. In any society certain patterns of conduct are established as accepted and unexceptional. Cultural norms and expectations determine the boundaries for accepted conduct. Since behavior that falls within these bounds is expected, it does not occasion special consideration or deliberation. Douglas

notes that food regimes and purity rules constitute an important part of the implicit norms that provide background rules for acceptable conduct in any society. Conduct that challenges these boundaries is defined as risky. Such conduct will either be repressed, or it will require justification according to burdens of proof that are also culturally determined (23,24).

From the standpoint of rational choice, cultural norms function as pre-deliberative filters that limit the circumstances in which deliberation will be applied. Although the range of deliberative choices in industrial societies is quite broad, it is impossible for individuals or organizations to apply the calculation and ranking entailed by deliberative rational choice to every potential choice situation. People simply do not have enough time and mental energy to weigh the consequences of every possible action. As such, rational behavior presupposes the existence of cognitive filters that sort life's options into categories. Some actions require no deliberative attention, others do. Douglas's purity rules function as pre-rational filters that sort life's options into the unexceptional and the risky. Risky actions require further consideration; they fall under a bias that demands proof of their acceptability. Actions that are consistent with purity norms do not trigger these additional burdens of proof, which is to say that they are not risky.

Although this approach is tantalizingly close to the rational choice paradigm, it is important to see that it utilizes an altogether different conception of risk. Crucially, there is no contradiction in saying that a given action has a nonzero likelihood of causing harm, but is not risky in virtue of the fact that it would not call for deliberative consideration. To say that an action entails risk in Douglas's sense is simply to say that it is out of the ordinary. Many daily activities — walking down stairs, making a pot of coffee — are undertaken without deliberative, conscious calculation. Bad things can happen as a result of walking down stairs or making coffee, but we do not apply a calculation of the probability of bad consequences in ordinary daily pursuit of these activities.

A culturally based set of food purity rules would have to be functionally rational in the sense that they would have to limit the number of cases of accidental poisoning as well as short-term disorders and dietary deficiencies. A society with too many such incidents would experience any number of weaknesses that would threaten its survival. Indeed, if a food regime appears to functioning adequately, it might be irrational for the cultural elite to expend time and energy on a deliberative review of it. However, anyone who violates these rules or who challenges them in any way would be engaging in conduct that *does* call for deliberative review, and quite possibly sanction. Such conduct would be classified as risk (18).

The time-honored response to risk is to repress it, to ban risky actions altogether. Responses to risk in complex societies are more varied. One response is to initiate the conscious, deliberative review of choice that leads one eventually to an assessment of the probability that a given course of action will result in harm or loss. However, other responses may also be reasonable, including policy norms that permit risky actions when affected parties have been given the opportunity to give or withhold their consent. New foods would thus not constitute a risk to an existing food regime so long as practitioners of traditional purity rules could continue their usual practices, experimenting with novelty only under conditions of consent.

Purity rules have two important philosophical ramifications. First, they blur the distinction that experts draw between safety understood as probability of harm or loss and broader, culturally based views about what is or is not acceptable food. When an observant Jew or Muslim violates dietary laws, their conduct challenges tradition in a manner that might be called "risky," even though it may have little objective probability of causing illness or injury. The social, cultural, and individual objectives being served through dietary rules may be much broader than the expert's conception of safety, and the attempt to supplant them with a reductive approach to risk may itself be perceived as a challenge to the cultural integrity of a food regime (25). The policy debate that occurs in response to this situation leads to political and ethical issues about how expected value assessments of food safety should be deployed, and whether they are consistent with consent criteria, discussed below.

Second, while functionally rational, purity rules suggest an approach to risk that is conceptually incompatible with deliberative rational choice. It is meaningful to claim that a given food or diet poses "no risk" on this view. Advocates of deliberative rational choice translate this as "zero risk." They go on to assert that the zero risk goal is irrational, implying that anyone who continues to address dietary choice through a framework of purity rules is irrational, hysterical, or at least profoundly misinformed. In itself, this may not represent a deep philosophical problem. We have two ways to use our general concept of risk, one in a general classificatory sense, the other to specifically call attention to the probabilistic dimension of unintended consequences. Pragmatic and contextual circumstances should determine which sense is in play at any given moment. However, this pragmatic or contextualist approach to risk is itself challenged by many who defend a more essentialist analysis of risk. Here genetically engineered foods become a case for a larger philosophical dispute (26).

Risk and Consent

Literature on risk and consent emphasizes situations in which individuals will be exposed to hazards as a condition of employment, medical treatment, or involvement in scientific research. In the almost unanimous opinion of scholars who have written on the subject, such risks may not be imposed on conscious and competent individuals without their consent. Individuals may claim a right to the information needed to evaluate the likely consequences of such risks, and to withhold consent, or *exit* from the risky situation. When such risks are imposed without consent, the party imposing risk may be held responsible for damages, and it is morally culpable for imposing risk, even when damages do not actually occur.

Classic environmental risks from air or water pollution present an altogether different situation. The entire population in a region is typically exposed to such risks, and it is in no position to claim a right of exit. There is debate as to

the significance of these cases for consent criteria. On one view, such risks differ in their nature, and must be evaluated according to criteria of general public good, rather than consent. On another view, pollution risks differ only in virtue of the fact that it is difficult to identify the source of pollution, and hence difficult to discern who should be culpable for imposing risk without consent (27–29).

Food safety risks introduced through the food chain bear some similarity to environmental risks. Once a genetically engineered variety of corn or soybean is combined with other bulk crops, it is very difficult to trace the source and content of any given commodity lot. Processed foods with variable ingredients have long been sold, and it would be impossible to argue that consumers have been able to claim a right to reject food constituents on an ingredient by ingredient basis. However, it is also true that religious minorities have successfully maintained an ability to exercise dietary rules. Individuals with very idiosyncratic beliefs about diet and health have been able make food choices based on those beliefs, mainly by eating a diet consisting in whole foods. GMOs challenge their opportunity to do this, for there may be no way to determine whether any given whole food commodity may contain grains, meats, or milk products derived from GMOs (30).

Food biotechnology's challenge to consent is procedural. One need not show that GMOs pose measurable risk to individual consumers in order to show that they challenge an individual's right of consent. As noted above, this right is typically exercised against a broad range of challenges to an existing dietary regime, not simply against the probability that the individual will suffer from injury or disease. As such it is useful to review some of the reasons why individuals might prefer not to eat genetically engineered foods.

1. *Religious objections*. Genetic engineering raises religious issues for many individuals. At least three types of religious concern may be relevant to food.
   - *Genetic engineering is wrong*. Clearly, individuals who believe that all forms of genetic engineering are wrong may have a legitimate reason to avoid genetically engineered food.
   - *Dietary rules*. The question of whether a particular food biotechnology is consistent with a given sect or congregation's interpretation of dietary rules (e.g., *kashrut*) must be left to religious authorities.
   - *Sanctity of life*. Some critics of biotechnology have extended a religiously based concern about commercial exploitation of genetic technology to animals and plants.
2. *Mistrust of science*. Many do not trust scientists or scientific pronouncements in the wake of well-publicized mistakes and deceptions. This mistrust takes at least three forms.
   - *Safety concerns*. Though no evidence suggests that GMOs increase the probability of disease or injury, many are unwilling to rely on existing studies or the word of scientists.
   - *Reflexive risk inferences*. Some people infer that if scientists and regulators are unwilling to provide information through food labeling (discussed below), the technology must be dangerous.
   - *Increasing power of scientific elites*. Some may resist genetic technology because the see it as one instance of a general loss of individual autonomy in complex society.
3. *Broader consequences*. Consumers may feel that food choices provide them the best opportunity to voice concerns about environmental, social, or animal impacts of food biotechnology.
4. *The yuk factor*. Many find genetic engineering aesthetically repulsive. Since individual aesthetics are an intrinsic dimension of food choice, it is reasonable for individuals to cite their aversion as a basis for withholding consent.

Any of these reasons might provide an individual with a reason to reject food biotechnology as a personal choice, and to regard it as "impure" or "unwholesome" (31).

FOOD LABELS

The larger philosophical issue is that experts and a segment of the lay public may be applying different philosophical frameworks to food safety. Experts define food safety quite narrowly, with respect to the conditional probability of disease or injury associated with the consumption of genetically modified foods. Technical problems of risk assessment aside, they see little ethical basis for concern about GMOs. Some in the lay public are applying a form of purity rules or have one or more of the concerns noted above. They claim a right to exit from the emerging food regime that includes GMOs. This tension is manifesting itself as a dispute over the need for labeling of genetically modified foods.

The labeling dispute replicates issues in the larger dispute. On the one hand, food labels can be seen as instruments that enable rational choice. In this view, information placed on the label of a food has ethical significance to the extent that it helps consumers realize the optimizing objectives of deliberative rational choice. Information must be true, but it must also be usable. It must help consumers reach the goals they seek to implement. To the extent that these goals are limited to health concerns, the only basis for distinguishing one product from another is when there is some reason to associate a measurable probability of disease or harm. The paradigm cases in the United States have been tobacco and saccharin, both subjected to mandatory labels in the wake of scientific studies. Lacking studies that identify risk from GMOs, there is no basis for requiring labels.

On the other hand, those who claim a right to know whether any given food is the product of genetic engineering or other forms of food biotechnology are demanding a right of exit. They see food labels as mechanisms for securing consent, without regard to whether or how they will use the information on labels in making food choices. Even those who plan to eat GMOs

may believe that the right of consent should be protected. Their mistrust of food biotechnology may, ironically, be based on suspicions that arise in the wake of resistance to labeling that arises on the part of the expert community.

The philosophical perspectives of rational choice and of consent thus present radically different burdens of proof for evaluating a policy on the labeling of GMOs, just as they do for evaluating the broad questions of risk and safety themselves. The likely consequences of labeling may be largely irrelevant to an advocate of consent, and the rational choice view that labels must be constructed so as to enable better choices will be seen as paternalistic. On the other side of the controversy, experts fear that poorly informed individuals will misinterpret labels, that they will make unwise food choices, and that labels will stigmatize GMOs without basis (32).

BIBLIOGRAPHY

1. P. Sparks, R. Shepherd, and L.J. Frewer, *Agric. Hum. Values* **11**(1), 19–28 (1994).
2. L.J. Frewer, R. Shepherd, and P. Sparks, *Br. Food J.* **9**, 26–33 (1994).
3. T.J. Hoban, in D. Maurer and J. Sobal, eds., *Eating Agendas: Food and Nutrition as Social Problems*, de Gruyter, New York, 1995, pp. 189–209.
4. I.C. Munro and R.L. Hall, in J.F. MacDonald, ed., *Agricultural Biotechnology, Food Safety, and Nutritional Quality for the Consumer*, National Agricultural Biotechnology Council, Ithaca, NY, 1991, pp. 64–73.
5. S. Huttner, in R.D. Weaver, ed., *U.S. Agricultural Research: Strategic Opportunities and Options*, Agricultural Research Institute, Bethesda, MD, 1993, pp. 155–168.
6. N. Elias, *The Civilizing Process*, Basil Blackwell, Oxford, UK, 1976.
7. W.C. Whit, *Food and Society: A Sociological Approach*, General Hall, Dix Hills, NY, 1995.
8. B. Ames, *Science* 1256–1263 (1983).
9. P.B. Thompson, in T.B. Mepham, G.A. Tucker, and J. Wiseman, eds., *Issues in Agricultural Bioethics*, Nottingham University Press, Nottingham, England, 1995, pp. 31–45.
10. S. Stich, *Philos. Public Affairs* **7**, 187–205 (1978); reprinted in M. Ruse, ed., *Philosophy of Biology*, Macmillan, New York, 1989, pp. 229–243.
11. D. Johnson, *Agric. Hum. Values* **3**(1&2), 171–179 (1986).
12. R. Wachbroit, in M.A. Levin and H.S. Strauss, eds., *Risk Assessment in Genetic Engineering*, McGraw-Hill, New York, 1991, pp. 368–377.
13. J.D. Graham, L.C. Green, and M.J. Roberts, *In Search of Safety: Chemicals and Cancer Risk*, Harvard University Press, Cambridge, MA, 1988.
14. J. Halloran, *Agric. Hum. Values* **3**(1&2), 5–9 (1986).
15. M.P. Doyle and E.H. Marth, in B. Baumgardt and M. Martin, eds., *Agricultural Biotechnology: Issues and Choices*, Purdue Research Foundation, Lafayette, IN, 1991, pp. 55–67.
16. H.R. Cross, in J.F. MacDonald, ed., *Animal Biotechnology: Opportunities and Challenges*, National Agricultural Biotechnology Council, Ithaca, NY, 1992.
17. M.J. Dobbins et al., *J. Nutr. Educ.* **26**(2), 69–72 (1994).
18. P.B. Thompson, *Food Biotechnology in Ethical Perspective*, Chapman & Hall, London, 1997.
19. C.F. Cranor, *Regulating Toxic Substances: A Philosophy of Science and the Law*, Oxford University Press, New York, 1993.
20. A.L. Caplan, *Agric. Hum. Values* **3**(1&2), 180–190 (1986).
21. G.L. Johnson and P.B. Thompson, in B. Baumgardt and M. Martin, eds., *Agricultural Biotechnology: Issues and Choices*, Purdue Research Foundation, Lafayette, IN, 1991, pp. 121–137.
22. H.I. Miller, *Trends Biotechnol.* **13**, 123–125 (1995).
23. M. Douglas, *Implicit Meanings*, Routledge, London, 1975.
24. M. Douglas and A. Wildavsky, *Risk and Culture*, University of California Press, Berkeley, 1982.
25. K. Taylor, in J.F. MacDonald, ed., *Agricultural Biotechnology, Food Safety, and Nutritional Quality for the Consumer*, National Agricultural Biotechnology Council, Ithaca, NY, 1991, pp. 96–102.
26. R. von Schomberg, in R. von Schomberg, ed., *Contested Technology: Ethics, Risk and Public Debate*, International Centre for Human and Public Affairs, Tilburg, 1995, pp. 13–28.
27. D. MacLean, in *Values at Risk*, Rowman & Allanheld, Totawa, NJ, 1986, pp. 17–30.
28. H. Shue, *Agric. Hum. Values* **3**(1&2), 191–200 (1986).
29. K. Shrader-Frechette, *Risk and Rationality*, University of California Press, Berkeley, 1991.
30. J. Burkhardt, in H.D. Hafs and R.G. Zimbelman, eds., *Low-Fat Meats: Design Strategies and Human Implications*, Academic Press, San Diego, CA, 1994, pp. 87–111.
31. P.B. Thompson, *Hastings Cent. Rep.* **27**(4), 34–38 (1997).
32. P.B. Thompson, *Choices: Mag. Food, Farm, Resour. Issues*, First Quarter, 11–13 (1996).

See other entries AGRICULTURAL BIOTECHNOLOGY; ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE.

AGRICULTURAL BIOTECHNOLOGY, LAW, AND EPA REGULATION

MARY JANE ANGELO
St. John's River Water Management District
Palatka, Florida

OUTLINE

INTRODUCTION

As we cross the threshold into the new millennium, agriculture is on the brink of a new technological era. Biotechnology is dramatically changing the landscape of modern agriculture. In the mid-1990s there were more than 1200 biotechnology companies in the United States (1). This number is expected to increase dramatically in the new millennium. In the year 2000, farm-level sales of biotechnology products are expected to be in the tens of billions of dollars (2). One of the most promising areas of agricultural biotechnology is the area of pest resistance. Biotechnology has made it now possible to produce naturally occurring proteins that act as pesticides in quantity in microbial organisms or plants. The ability of a microbe or crop plant to produce pest resistance may obviate the need for harsher and less selective synthetic pesticides (3). In addition to pest resistance, a variety of crops have been engineered to produce increased levels of desired nutrients or to impart them with other desirable characteristics such as cold tolerance (3). Regulatory agencies, such as the U.S. Environmental Protection Agency (EPA) have been struggling to keep up with these rapidly developing technologies and to establish regulatory programs that promote the beneficial uses of the new products, while attempting to protect the public health and the environment from the associated potential risks. This article focuses on EPA's regulation of agricultural biotechnology under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) (4) and the Federal Food, Drug, and Cosmetic Act (FFDCA) (5).

AGRICULTURAL BIOTECHNOLOGY

Since 1962, when Rachel Carson's seminal work, *Silent Spring* (6), first awakened the country to the risks of chemical pesticides, the public has been skeptical of the government's ability to protect them and their environment from the hazards of pesticide use. While in recent years public attention has increasingly focused on the risks of pesticides in food, particularly the risks to small children, many pesticides may also pose significant risks to farm workers, consumers, and the natural environment. Within three decades following the publication of *Silent Spring*, environmentalists and consumer groups repeatedly called upon the government, in particular EPA, to push for the reduction of pesticide use. Despite these efforts, in 1991 alone, approximately 4.1 billion dollars worth of pesticides, roughly 320 million kilograms of pesticides, were used in the United States (7). EPA has registered over 19,000 pesticides, containing 913 different active ingredients (8).

In the 1990s dramatic changes began to take place, especially with regard to an ever-increasing number of pesticide products derived from biotechnology processes. The promise of these pesticides stems from the potentially lower risk to humans and the environment. This is due to their greater specificity to the target pest than chemical pesticides, their tendency to have lower toxicity than chemical pesticides, and the tendency of many biotechnology pesticides to have limited persistence in the environment. The universe of biotechnology pesticides is large and diverse, and it includes microorganisms such as bacteria, fungi, algae, protozoa, and viruses, which act as pesticides by producing toxins, acting as parasites, or acting through competition and macroorganisms such as parasitic wasps or plants that produce substances that exert a pesticidal effect.

What is Biotechnology? Biotechnology, in its broadest sense, is the use of living organisms, be they plants, animals, or microorganisms, to make or modify products. While there does not appear to be one standard definition of "biotechnology," most definitions are broad enough to cover a wide array of processes including genetic engineering and more traditional processes such as plant breeding and fermentation. The U.S. government has defined "biotechnology" as the "use of various biological processes, both traditional and newly devised, to make products and perform services from living organisms or their components" (9). For centuries, biotechnology has been used to manufacture products such as bread, beer, wine, yogurt, and cheese (10). For thousands of years traditional plant breeding has enabled the production of crop plants with desired traits such as high seed yields and increased resistance of pests and environmental stresses (11). For example, early farmers are believed to have created wheat over 5000 years ago by combining traits from three different species (12). By repeatedly selecting plants that exhibit the desired traits and crossbreeding them with closely related plants over several generations, traditional plant breeders were able to create plants with a desired combination of traits (11). However, traditional plant breeding is limited by two major constraints: (*1*) Removing undesirable traits from the original cross can take generations and often takes years; (*2*) only closely related plant species can be directly bred together, severally limiting the gene pool available. Genetic engineering enables plants to be developed that cannot be produced through traditional plant breeding (12).

In recent years, through the use of recombinant DNA (rDNA), researchers have been able to "genetically engineer" organisms by moving genes from one organism to another. Recombinant DNA technology allows the isolation and characterization of specific pieces of DNA from one organism and transfer of the DNA sequences into another organism. The term "genetic engineering" generally refers to the use of recombinant DNA (rDNA), cell fusion, or other novel bioprocessing techniques (13). Recombinant DNA technology has dramatically increased the speed of inserting a desired trait into a plant (11). Moreover rDNA techniques eliminate the problem of undesirable traits being introduced into the plant along with the desired genes (14). Additionally rDNA techniques can be used to move desired genes from virtually any types of living organism, be it plant, animal, or microorganism, into the plant (14).

How is Biotechnology Used in Agriculture? For the past 10 years, EPA has exercised regulatory oversight over genetically engineered microbial organisms that act

as pesticides. Because microbial pesticides are living organisms and have the potential to reproduce and spread on their own in the environment, they pose the potential for unique risks. Thus EPA's regulatory scheme for microbial pesticides is somewhat different from that for conventional chemical pesticides. Nevertheless, microbial pesticides are similar to conventional pesticides in that they are "applied to" crops, and thus they are in many ways regulated like traditional pesticides. In the past several years, however, EPA has been faced with a completely new class of genetically engineered pesticidal products that poses a new set of regulatory challenges. In the 1990s significant technological advances were made in altering plants to produce pesticidal substances. That rDNA technology has advanced to the point where researchers are able to more easily move genes from microorganisms, animals, or other plants into agricultural crop plants. For example, through these new rDNA technologies, plants can be made to produce toxins normally produced only by microorganisms such as the *Bacillus thuringiensis* (Bt) insecticidal delta-endotoxin. Bt acts by forming a protein crystal, referred to as the delta endotoxin, that becomes toxic upon ingestion by the insect. EPA considers the pesticidal substances produced by plants and the genetic material necessary to produce them to be "plant-pesticides." Although EPA does not yet have in place a final comprehensive regulatory scheme to address plant-pesticides, the biotechnology industry has advanced to the point where it is commercializing products, and thus, for the past several years, EPA has been regulating these products under a proposed regulatory scheme. EPA has issued several Experimental Use Permit (EUP) applications and registration applications for the Bt delta-endotoxin produced in various plants. (See section on FIFRA below.) EPA also has granted applications for tolerance exemptions for residues of pesticidal substances produced in plants as a result of genetic engineering.

Some of the agricultural crops that have been developed through this type of genetic modification in recent years include corn, cotton, and potato plants that have been genetically modified to contain a bacteria gene that leads to the production of the Bt insecticidal toxin, squash that has been genetically modified to contain a virus gene that leads to the production of a viral coat protein to make the plant resistant to infection by viruses, and cotton and soybean plants have been genetically modified to contain bacteria genes that cause the plants to tolerate herbicides that are applied to the plant. Recombinant DNA techniques also are being used to produce a new class of animal hormones, the somatotropins, such as the bovine somatotropin (BST) hormone, which was approved by the Food and Drug Administration (FDA) in 1993 for use in lactating dairy cows to produce more milk.

EPA received the first EUP application for the Bt toxin produced by a genetically engineered plant, cotton, in November 1991. In the years following the Bt in cotton EUP, EPA has granted EUPs for Bt in potatoes and corn and a number of registrations for Bt in cotton, potatoes, and corn. EPA has also granted exemptions from the requirement of a tolerance for a number of Bt plant-pesticides. In addition EPA has registered and granted a tolerance exemption for the potato leafroll virus resistance gene and has granted a number of tolerance exemptions for a variety of viral coat proteins in raw agricultural commodities. Currently EPA is reviewing a number of EUP registration, and tolerance exemption applications for other plant-pesticide products. It is anticipated that number of applications for EUPs, registrations, and tolerances for plant-pesticides will continue to grow at a rapid pace.

What is the Significance of Agricultural Biotechnology?

Risks. Many of the risk considerations for biotechnology pesticides are similar to, if not the same as, those for traditional chemical pesticides. In general, EPA has expressed a view that traditional chemical pesticides pose greater environmental risks than biochemical, microbial, or plant-pesticides (15). As with any pesticide risk assessment, the underlying considerations for analyzing risks posed by biotechnology pesticides are the potential for humans and other nontarget organisms to be exposed to the pesticide, and the hazard (usually toxicity) of the pesticide to nontarget organism, humans, and the environment. For biotechnology pesticides, as with other pesticides, hazard will be determined by the chemical and toxicological properties of the pesticidal substance. Exposure, on the other hand, will be determined somewhat differently for biotechology pesticides than for traditional chemical pesticides. For traditional pesticides the primary factors in determining exposure is the amount of chemical that is introduced into the environment and the likelihood that humans or other nontarget organisms will come into contact with the chemical. Because microbial and plant-pesticides are produced by living organisms, however, exposure issues are more complex for these substances and are dependent, in large part, on the biological characteristics of the organism itself. For example, exposure to a plant-pesticide could be determined by factors such as whether the production of the plant-pesticide is limited to particular plant parts (e.g., leaves, stems, fruit, or roots) and what organisms consume or are associated with those plant parts.

Moreover one of the most significant exposure considerations for microbial pesticides and plant-pesticides not seen for chemical pesticides is the potential for spread of the living organism or the organism's genetic material. For example, plants can reproduce sexually and/or asexually, and as a result the genetic material that was introduced into the plant and that enables the plant to produce plant-pesticides could spread through agricultural or natural ecosystems. Thus, if a plant that produces a plant-pesticide has the capacity to spread in the environment, or to spread its genetic material to other plants, there would be a greater potential for increased exposure to nontarget organisms than there would be for a plant-pesticide produced in a plant that can only grow in a limited geographic area or does not have the ability to cross-fertilize with other plants in the environment. This is a particular concern for plant-pesticides produced in plants that have wild relatives in the United States. If these wild relatives acquire the ability to produce the plant-pesticide, through

cross-fertilization, many additional nontarget organisms could potentially be exposed to the pesticide.

The potential for a genetically modified organism (GMO) or its genetic material to spread from one plant to another raises additional risk issues beyond those of exposure to humans and nontarget organisms. One potential risk of biotechnology products parallels the risk of the introduction of any nonnative species into a new environment (15). Small genetic manipulations can result in significant changes in an organism's ability to survive and flourish in a particular ecosystem (12). There are dozens of examples of the disastrous, but unpredicted, effects of the introduction of nonnative species into the environment displacing native species (13). Genetically modified organisms introduced into the environment could have similar impacts (13). The risks that appear to be most significant are that a genetically engineered plant might become a weed or pest itself or that it might outcross with related species to create new weeds or pests (13). Once released into the environment, the spread of a GMO may be difficult, if not impossible, to control (13). One of the most cited concerns about plant-pesticides is the concern over the potential for the development of "superweeds" through the outcrossing of plants producing plant-pesticides to wild relatives. If the ability to produce a plant-pesticide that, for example, makes a plant resistant to insect or viral pests is spread to a wild relative and passed on to subsequent generations of that relative, there is the potential that the wild relative, by virtue of its newly acquired ability to resist insects or viruses, could become a hardy weed. Development of such a weed has the potential to disrupt agricultural or natural ecosystems. For a transgenic plant to transfer its genes to related existing weed species, however, wild relatives of the transgenic plant must grow in the geographic areas where the transgenic plant is introduced (13). Most crops grown in the United States are of foreign origin. Thus the risk of hybridization between transgenic crops and wild relatives is unlikely in the United States. Many domestic crops including soybeans, corn, and wheat have been bred to the point where they have lost their ability to compete with wild species in the environment. Thus these crops are unlikely to become weeds when genetically altered (12).

Another issue that has received considerable attention is the potential for plant-pesticides in foods to pose a risk of allergenicity to humans. The primary concern appears to be that if a gene that leads to the production of a plant-pesticide is moved from one plant, for example, a peanut, into another plant, for example, corn, people who know they are allergic to peanuts will not know to avoid the corn plant. Thus, if the plant-pesticide derived from the peanut plant contains an allergen from the peanut plant, allergic consumers could be put at risk (14,15). Other areas of potential adverse effects on the environment center on specific plant-pesticides or categories of plant-pesticides. For example, some environmental organizations have expressed their concern that engineering plants to produce viral coat proteins has the potential to result in the develop of new unintended viruses.

In addition to the risk concerns described above, public interest organizations have articulated other concerns that are more philosophical, ethical, or religious in nature. For example, the movement of genes from animals to plants may be of concern to subpopulations of people with special dietary preferences such as vegetarians or persons who observe kosher (Jewish) or halal (Muslim) laws (17). Other philosophical issues that have been raised include a concern that the prospect of "human-made" organisms, even if they pose no risk to humans or the environment, may threaten the concepts of "wildness" and "wilderness" (18, p. 33). Some argue that while biotechnology pesticidal products may be environmentally preferable to traditional chemical pesticides, the focus on developing these products may be diverting attention from the more important goal of developing a system of sustainable agriculture (19, p. 67).

Probably the most significant concern with agricultural biotechnology stems from the fact that the risks of biotechnology are uncertain. Although the risk of a genetically modified organism released into the environment creating a new superweed or disrupting the balance of natural ecosystems may be small, the consequences could be disastrous and potentially irreversible (15). The precise nature and magnitude of the risk is difficult to predict because of the almost infinite variety of potential genetically modified organisms, the reproductive ability of GMO's, the complexity of the natural balance of ecosystems and the dearth of long-term data (15).

Benefits. To many, agricultural biotechnology products hold the promise of a less risky substitute for traditional chemical agricultural products. The use of rDNA technologies has enabled organisms, particularly plant varieties, to be developed that either could not have been developed through traditional plant breeding or could only be developed through traditional techniques with a great amount of time and difficulty. Chemical pesticides often are of relatively high toxicity. Many, but not all, traditional chemical pesticides are toxic to a broad range of organisms, including humans. In addition the manner in which traditional pesticides are applied — often sprayed over large areas — could result in significant exposure to nontarget organisms. Biotechnology pesticides, on the other hand, are generally of low toxicity, target-specific, and produced in relatively small quantities in the organism. Because plant-pesticides are generally produced in small amounts in the plant, nontarget organisms are not as likely to be exposed to these pesticides as they are to pesticides that are sprayed over large areas. Moreover, even if nontarget organisms are exposed to plant-pesticides, because these pesticides are often of low toxicity and are generally target specific, nontarget organisms are not as likely to be adversely affected by these pesticides as they are with pesticides that are more highly toxic or toxic to a broad spectrum of organisms. For example, the Bt. toxin is specific to specific groups of insects (e.g., Lepidoptera) and is not toxic to humans or other mammals.

One example of where a plant-pesticide is believed to have the potential for significant environmental benefits, is viral coat protein-mediated resistance. By genetically modifying plants to produce certain viral coat proteins, researchers have been able to produce plants that are resistant to infection by particular viruses. For viruses

spread by vectors such as insects, the most common agricultural practice for preventing viral attack is the use of chemical pesticides to control the insect vector that spreads the virus. It is believed that the use of viral coat protein-mediated resistance would reduce the need for these chemical pesticides. In addition to the environmental benefits of viral coat protein-mediated resistance, there is a high potential for significant economic benefits. Another potential environmental benefit is the reduction of runoff of agricultural chemicals such as pesticides and fertilizers, which can contaminate surface and ground water (11). For example, rDNA technique may be used to create plants with improved photosynthetic and nitrogen fixation capabilities, thereby reducing the need to apply fertilizers (11). Moreover, some of the most environmentally friendly herbicides with relatively low toxicity, low soil mobility, and rapid biodegradation are also the herbicides that are the most nonselective, and thus the most likely to kill crops plants along with the weed (11). Farmers are often forced to apply more selective and more toxic herbicides as a result (11). The use of genetically modified herbicide tolerant crop plants may benefit the environment by causing a reduction in the use of highly toxic herbicides. Environmental organizations have expressed concerns that plants that have been genetically modified to be tolerant to herbicides could actually result in an increase in herbicide use because herbicides would be able to be directly applied to crop plants without killing them (18). Industry groups, however, assert that these plants will enable farmers to reduce the number of herbicide applications by allowing farmers to target the timing of herbicide application to after the plant has emerged, when herbicides are most needed (20,21). Herbicide tolerant plants are not considered to produce plant-pesticides because the substances they produce are not intended to prevent, destroy, or repel pests. Thus these products would not be covered by the plant-pesticide policy or rules. EPA is planning to develop a separate policy for these plants. Other potential benefits of plant-pesticides may not yet be apparent. Nevertheless, many scientists believe that the technological advances in this area hold out great promise for the future.

Public Perception. The intensity of the public response to the 1992 FDA policy on foods derived from new plant varieties, as well as the public concerns surrounding FDA's approval of the BST milk, illustrates the important function that public perception will play in defining the role of agricultural biotechnology in the marketplace. As others have pointed out, while many new technologies will soon be commercially viable, they all will not automatically be put to use — consumers will be the ultimate judge of emerging technologies (22). Key to the success or failure of new biotechnology products will be the ability of the government agencies responsible for regulating these technologies, such as EPA, to effectively communicate to the public the risks and benefits of these products and the public's resulting acceptance or nonacceptance. Many people are skeptical of any new technology. This skepticism is even more pronounced with biotechnology, which could be difficult for the layperson to understand because it is surrounded by many uncertainties. A recent survey conducted to gather information on consumer attitudes about the use of biotechnology in agriculture and food production concluded that one of the most important factors influencing public perception of biotechnology is the perceived credibility of public policies and regulations. This survey found that while most consumers supported the use of biotechnology in agriculture and food production (22), they also favored an active role for government agencies in establishing biotechnology regulations that ensure environmental protection and food safety (22). Thus EPA must be mindful that the public will be looking to it, not only to evaluate the risks and benefits of biotechnology pesticides in order to develop a regulatory program that will protect humans and the environment but also to effectively communicate with the public on these issues. Possibly the most serious public concern over agricultural biotechnology is the use of the technology in the production of food crops (23).

HOW IS AGRICULTURAL BIOTECHNOLOGY REGULATED?

The Coordinated Framework

The U.S. government's first systematic attempt to address the regulation of biotechnology in a comprehensive fashion was with the publication of the 1984 document entitled *Proposal for a Coordinated Framework for Regulation of Biotechnology* (24). The purpose of this document was "to provide a concise index to U.S. laws related to biotechnology, to clarify the policies of the major regulatory agencies that will be involved in reviewing research and products of biotechnology, to describe a scientific advisory mechanisms for assessment of biotechnology issues, and to explain how the activities of the Federal agencies in biotechnology will be coordinated." In 1986 the Office of Science and Technology Policy (OSTP) published in the Federal Register a *Coordinated Framework for Regulation of Biotechnology; Announcement of Policy and Notice for Public Comment* (the Coordinated Framework) (25). This document made clear that the Executive Branch believed it could adequately regulate biotechnology under its existing authorities and did not intend to seek new legislation to address emerging technologies. The Coordinated Framework described in detail the roles of the five federal agencies with significant involvement in the regulation of biotechnology: FDA, the United States Department of Agriculture (USDA), EPA, the National Institutes of Health (NIH), and the Occupational Safety and Health Administration (OSHA). The Coordinated Framework was created to harmonize the regulation of biotechnology between several federal agencies and to address gaps and overlaps between and among agencies (15). The Coordinated Framework contains four major conclusions: (*1*) Existing federal statutes are sufficient to regulate biotechnology, (*2*) federal agencies should regulate "products" rather than the "process," (*3*) the safety of biotechnology products should be addressed on a case-by-case basis, and (*4*) the efforts of all agencies involved in regulating biotechnology should be coordinated (13). The Coordinated Framework gave

EPA the primary responsibility over the environmental regulation of biotechnology.

Under the Coordinated Framework, the regulatory approach taken by U.S. regulatory agencies, including EPA, has been to rely on existing statutes and to focus on the "product" rather than the "process" used to create the product (15). The thinking underlying this approach is that rDNA technology in itself does not create risk, and thus does not necessitate an entirely new regulatory system (15). Instead, certain types of products of biotechnology may pose risks that can be addressed in the same fashion as risks posed by traditional "chemical" products.

EPA's Statutory Authority

EPA's primary authority for regulating agricultural biotechnology products can be found in two statutes: the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), 7 U.S.C. §136–136y, and the Federal Food, Drug, and Cosmetic Act (FFDCA), 21 U.S.C. §301 *et seq*. Since EPA was created in 1970, it has had responsibility for the regulation of pesticides under both laws. Reorganization Plan No. 3 of 1970, 84 Stat. 2086. Under FIFRA, EPA is responsible for regulating the distribution, sale, use, and testing of pesticides to prevent unreasonable adverse effects to humans and the environment. In evaluating a pesticide, EPA balances the potential human and environmental risks against the potential benefits to society of using that pesticide. Under FFDCA, EPA has the authority to set tolerances for pesticide chemical residues in or on food. In establishing tolerances, EPA evaluates the impacts of human dietary exposure to the pesticide residues. EPA also regulates biologicals and biotechnology products that are not pesticides, food, or drugs under the Toxic Substances Control Act (TSCA), 15 U.S.C. §§2601–2692. TSCA grants EPA the authority to screen new chemical substances and impose controls to prevent unreasonable risks and, through rule making, to acquire information and impose restrictions to prevent unreasonable risks on existing chemical substances. Although some agricultural biotechnology products may fall within the purview of TSCA, the majority of agricultural biotechnology products regulated by EPA are considered pesticides under EPA's broad definition of the term, and thus are regulated under FIFRA and FFDCA.

FIFRA. Section 2(u) of FIFRA defines the term "pesticide" as "(*1*) any substance or mixture of substances intended for preventing, destroying, repelling, or mitigating any pest, (*2*) any substance or mixture of substances intended for use as a plant regulator, defoliator, or desiccant . . . ." This definition is very broad and can include living organisms and substances produced by living organisms as well as traditional chemical pesticides. The definition of "pesticide" in FIFRA does not depend on the process by which a particular pesticide is produced. EPA has interpreted this definition to include biological pesticides and genetically engineered pesticides.

Section 3 of FIFRA provides that no person may distribute or sell in the United States any pesticide that is not registered under the Act. FIFRA Section 3(c)(5) requires that before a pesticide can be registered, it must be shown that when used in accordance with widespread and commonly recognized practice, it will not generally cause "unreasonable adverse effects on the environment." The term "unreasonable adverse effects on the environment" is defined in FIFRA Section 2(bb) as any unreasonable risk to humans or the environment, taking into account the economic, social, and environmental costs and benefits of the use of any pesticide or human dietary risk resulting from pesticide residues in food inconsistent with the safety standard under Section 408 of FFDCA. Thus FIFRA involves a balancing of the risks presented by the use of the pesticide against the benefits associated with the use of that pesticide. The procedures governing the regulation of pesticides are set forth in 40 CFR Parts 152 through 172. One of the most important requirements is that the registrant or applicant submit data in support of registration. 40 CFR Part 158 sets forth data requirements for conventional pesticides and microbial pesticides (specifically at 40 CFR 158.740), and provides for the submission of comprehensive health and environmental effects data. EPA has not yet established specific data requirements for plant-pesticides. In addition to submitting required data, an applicant for registration must submit all proposed labeling with the registration application. FIFRA Section 2(p) defines the term "label" as the written, printed, or graphic matter on, or attached to, the pesticide. The term "labeling" under FIFRA includes the label as well as all other written, printed, or graphic matter that accompanies the pesticide or to which reference is made on the label. Registered pesticide products must bear a label or labeling that contains certain information, including precautionary statements, warnings, directions for use of the product, and an ingredient statement. FIFRA requires users of pesticides to follow all label directions. A product whose label or labeling does not contain the information required by EPA or that sets forth false or misleading information is misbranded pursuant to FIFRA Sections 2(q) and 12(a)(1)(E). For conventional pesticides, many risk reduction measures are achieved through labeling restrictions. As discussed below, however, many of these types of restrictions may not be appropriate for plant-pesticides.

FIFRA also provides EPA with a number of other regulatory tools beyond the registration authority. For example, large-scale field testing of pesticides is necessary to evaluate the efficacy of a potential product and to obtain data needed to support registration under FIFRA Section 3. This large-scale testing is regulated under Section 5 of FIFRA. Under this section, EPA is authorized to issue experimental use permits (EUPs) for limited use of an unregistered product for an unregistered use. Before an EUP is issued, EPA must determine that the field test will not cause an "unreasonable adverse effect" on the environment. For most new pesticides, EPA grants conditional registration while it continues to evaluate whether the pesticide product qualifies for full registration. Under Section 3(c)(7) of FIFRA, conditional registration is granted when EPA lacks sufficient data to make a final determination on a full registration. Finally, under FIFRA Section 25(b), EPA may exempt from some

or all requirements of FIFRA, by regulation, any pesticide determined to be (*1*) adequately regulated by another federal agency, or (*2*) of a character that is unnecessary to be subject to the Act in order to carry out the purposes of the Act.

FFDCA. EPA regulates pesticide residues in or on food under the authority of Section 408 of FFDCA. Under FFDCA Section 408, any pesticide chemical residue in or on food is deemed to be unsafe unless a tolerance, or an exemption from the requirement of a tolerance, for such pesticide is established and the pesticide is within the tolerance limits. The term "pesticide chemical" is defined in Section 201(q) of FFDCA as "any substance that is a pesticide within the meaning of [FIFRA] . . . ."

Thus pesticide chemicals subject to Section 408 of FFDCA are defined by reference to the definition of pesticide under FIFRA. Section 408(b) of FFDCA authorizes EPA to promulgate regulations to establish tolerances for pesticide chemical residues in or on food if EPA determines that the tolerance is "safe." This section goes on to provide that "safe" means that EPA has determined that there is a reasonable certainty that no harm will result from aggregate exposure to the pesticide chemical residue, including all anticipated dietary exposures and all other exposures for which there is reliable information. Thus unlike FIFRA, FFDCA only addresses human dietary risks. FFDCA Section 408(c) authorizes EPA to promulgate regulations exempting any pesticide chemical residue from the necessity of a tolerance when the exemption is safe.

In 1996 both FIFRA and FFDCA were amended by the Food Quality Protection Act (FQPA). FQPA amended FIFRA such that a registration cannot be issued for a pesticide to be used in or on food unless the residues of the pesticide in the food qualify for a tolerance or an exemption from tolerance. FQPA also modified FIFRA Section 2bb by incorporating the FFDCA safety standard into the test for determining whether a pesticide poses an unreasonable adverse effect. FQPA also amended Section 408 of FFDCA to require EPA to give special consideration to exposure of infants and children to pesticide residues.

EPA REGULATION OF AGRICULTURAL BIOTECHNOLOGY

Microbial Pesticides

EPA has regulated naturally occurring microbial pesticides, such as Bt, for many years. Microbial pesticides are regulated in much the same way as traditional pesticides at the large-scale testing and registration stages. For the past 10 years, however, EPA has been concerned about the potential for adverse effects associated with small-scale environmental testing of certain microbial pesticides, both naturally occurring and genetically engineered. Small-scale testing of most traditional pesticides generally is considered to pose very limited risks, and thus is not usually regulated by EPA. Because microbial pesticides are living organisms that have the potential to reproduce and spread in the environment, however, even small-scale testing has the potential to present unreasonable adverse effects on the environment.

Section 5 of FIFRA authorizes EPA to issue EUPs for the testing of new pesticides or new uses of existing pesticides. Under EPA's existing regulations at 40 CFR Part 172, EUPs are generally issued for large-scale testing of pesticides. A large-scale test under Part 172 includes any terrestrial application on a cumulative acreage of more than 10 acres of land or any aquatic application on more that 1 acre of surface water. EPA has generally presumed that tests conducted on 10 acres or less of land or 1 acre or less of water (small-scale tests) would not require EUPs. The Agency has determined, however, that small-scale tests conducted with certain naturally occurring and genetically engineered microbial pesticides may pose sufficiently different risk considerations from tests conducted with convention chemical pesticides so that a closer evaluation at the small-scale testing stage is warranted.

In October 1984, EPA published a policy statement entitled *Microbial Pesticides: Interim Policy on Small Scale Field Testing* (26). In June 1986, EPA reiterated the provisions of the Interim Policy Statement as part of the Office of Science and Technology Policy's Coordinated Framework for Regulation of Biotechnology. These policy statements described EPA's concern about the potential for adverse effects associated with small-scale environmental testing of certain microbial pesticides. To address this situation, these statements required that EPA be notified prior to initiation of small-scale testing of all nonindigenous and genetically engineered microbial pesticides. The purpose of the notification was to allow EPA to screen these small-scale tests and determine whether the tests should be carried out under an EUP that allows EPA oversight. In addition the 1986 Policy stated EPA's plan for future rule making in order to codify the interpretation set out in the policy.

After almost 10 years of deliberation and a series of EPA and federal government-wide policy statements that were made available to EPA's Scientific Advisory Panel (SAP) and the Biotechnology Science Advisory Committee (BSAC), on January 14, 1993, the EPA issued a proposed rule that was a somewhat revised version of the 1986 policy (27). The rule would codify the early screening procedure in the Coordinated Framework by requiring notification before the initiation of small-scale field testing of certain microbial pesticides in order to determine whether an EUP is necessary. Under the proposed rule, testing conducted in facilities designed and operated to adequately contain the microbial pesticide would not be subject to the notification requirements.

EPA received comments in response to the proposed rule making from trade associations, business firms, public interest groups, scientific researchers, and state and federal agencies. Perhaps the most controversial issue that arose during the lengthy development of this rule was the issue of what constitutes the appropriate scope of regulation. The proposal identified three options for defining the scope of genetically modified microbial pesticides subject to notification requirements. EPA's preferred option provided the most clear-cut scope of regulation — namely microbial pesticides whose pesticidal properties have been imparted or enhanced by the

introduction of genetic material that has been deliberately modified. EPA developed this option based on comments from the public in response to earlier Federal Register announcements, the SAP subpanel, the BSAC, and other agencies including USDA. The Agency preferred this option because it believed that it covers the appropriate microbial pesticides and has a high degree of regulatory utility. The majority of comments supported this option. The commenters generally agreed that EPA's preferred option was more clear-cut and that the decision of whether notification is necessary should not be left solely to the judgment of the researcher. In 1994 EPA issued the final microbial rule (28). The final rule included EPA's preferred option for the scope of regulation from the proposed rule. The final rule also included a mechanism to exempt, by rule making, additional microbial pesticides or categories of microbial pesticides from the requirement for notification as data and experience permit.

One other issue that was somewhat controversial was that of whether EPA should require notification for "nonindigenous" microbial pesticides. Under EPA's 1984 Policy Statement and the 1986 Coordinated Framework, EPA had been requiring notifications to be submitted for all small-scale testing of nonindigenous organisms. In all of the scope options presented in the proposal, EPA proposed to no longer require notifications for any nonindigenous microbial pesticides that have not been genetically modified. EPA based this decision on its belief that continued imposition of the notification requirement on these microbial pesticides would constitute duplicative oversight because the USDA's Animal and Plant Health Inspection Service (APHIS) already regulates small-scale testing of these organisms. Some commenters supported EPA's decision to exclude nonindigenous microbial pesticides from notification, while others believed that EPA should regulate any nonindigenous microbial pesticide that is not regulated by another federal agency. EPA responded to these comments by stating that it continues to believe that the vast majority, if not all, nonindigenous microbial pesticides are reviewed by APHIS. However, to address the concerns of some commenters that there might be a regulatory gap, EPA revised the language in the final rule to state that only those nonindigenous microbial pesticides that have not been acted upon by APHIS (i.e., either by issuing or denying a permit or determining that a permit is unnecessary; or when a permit is not pending with APHIS) are exempt from the notification requirement.

The final rule also contains several provisions that were not very controversial and were not changed significantly from what was proposed. In the final rule, testing conducted in facilities designed and operated to adequately contain the microbial pesticide would not be subject to the notification requirements. Records describing containment, however, would be required to be developed and maintained. The final rule also includes provisions that will enable EPA to address situations where small-scale testing results in unanticipated and untoward effects. Section 172.57 requires persons using microbial pesticides in small-scale tests to submit any information they obtain concerning the potential for unreasonable adverse effects from the microbial pesticide, and Section 172.59 enables EPA to take immediate actions to prevent use of a microbial pesticide if such use would create an imminent threat of substantial harm to health or the environment. Finally, the rule amends 40 CFR §172.3 to clarify its rationale for presuming that an EUP is not required prior to small-scale testing with most pesticides. As explained in the preamble to the final rule, Section 172.3 is modified to clarify that the determination of whether an EUP is required would be based on risk considerations, rather than on a definitional presumption about whether a substance is a pesticide. This clarification has general applicability to all pesticides and is not limited to microbial pesticides.

Plant-Pesticides

EPA's first attempt to describe its plans to regulate plant-pesticides was in early 1994. On January 21, 1994, EPA held a joint meeting of a subgroup of the Agency's SAP and BSAC to address certain scientific issues related to the regulation of pesticidal substances produced in plants. For the meeting, EPA made available to the public a draft proposal of a comprehensive policy and four draft proposed rules (together referred to as the "draft proposal") that were developed under FIFRA and FFDCA. On November 23, 1994, EPA published in the Federal Register somewhat modified versions of these draft documents (together referred to as "the proposal") (29–33). The proposal is intended to clarify the status of plant-pesticides under FIFRA and FFDCA and outline the scope of what types of plant-pesticides EPA believes warrant regulation based on risk–benefit considerations. Under the proposal many plant-pesticides would not be subject to regulation because they pose a low potential for risk to humans and/or the environment. Others would be subject to regulation but would be regulated somewhat differently than conventional pesticides because of the unique nature of plant-pesticides. The proposal outlines how EPA intends to assess plant-pesticides at different stages of environmental testing and at the sale and distribution stage. In developing this policy, EPA worked closely with two other federal agencies that also have regulatory jurisdiction over agricultural biotechnology products, to integrate the three agencies' regulatory programs and minimize duplicative regulation. Those agencies are APHIS, which regulates certain genetically modified plants, including plants that are modified to produce pesticidal substances, and FDA, which regulates nonpesticidal substances in food plants as food additives under FFDCA.

As described above, FIFRA defines the term "pesticide" very broadly, and under this definition both the "plant" and the pesticidal substances produced in the plant are considered to be "pesticides." However, in 1982 EPA promulgated a regulation under FIFRA Section 25(b) that exempted all biological control agents from the requirements of FIFRA, except for certain microorganisms (34). This exemption was promulgated because EPA found that macroorganisms used as biological control agents were adequately regulated by other federal agencies such as APHIS. Plants, as biological control agents, were implicitly exempted from regulation under FIFRA through

this exemption. EPA does not believe is it necessary to revoke this exemption for the plant itself. Instead EPA intends to focus on the pesticidal substance produced by the plant. This is consistent with EPA's past actions. For example, EPA does not regulate chrysanthemums, but it regulates the pesticidal substance pyrethrum that is produced by the chrysanthemum when it is extracted from the plant and applied onto other plants as an insecticide. However, prior to 1994, EPA had not clearly stated its policies for regulating pesticidal substances that are produced in living plants and not extracted from the plants (i.e., substances produced in plants naturally, or through genetic engineering or other technologies, that actually exert their pesticidal effect while still in the plant). It is these substances that EPA considers to be plant-pesticides and that are the subjects of the proposal.

One point that should be emphasized is that in the proposal, EPA has defined the pesticidal active ingredient as including not only the substance that is produced in the plant for the purpose of inducing the pesticidal effect but also the genetic material necessary for the production of that substance. To understand why EPA is including this genetic material in the definition of active ingredient, it is necessary to understand how a plant-pesticide is created. There are three primary steps involved in creating a plant-pesticide: (*1*) isolating the gene to be transferred from the source organism to the plant, (*2*) adding regulatory DNA sequences to the gene so that it will be properly expressed in the gene (these regulatory DNA sequences typically are derived from plant viruses), and (*3*) moving the gene to the plant. The last step can be accomplished by using physical methods such as microinjection and biolistic delivery (firing very small metal particles coated with DNA into plant tissues) or by using biological vectors such as the soil bacterium, *agrobacterium*. There are several reasons why EPA included the genetic material as part of the active ingredient. First, it is the genetic material that is actually added to the plant and that leads to the production of the substance that ultimately results in the pesticidal effect. Moreover EPA is not only concerned with the environmental risks associated with the pesticidal substance itself but also with potential environmental impacts associated with the spread of genetic material. Finally, from a practical standpoint, it may be easier to detect, for monitoring or enforcement purposes, the genetic material in a plant than the pesticidal substance itself.

Under EPA's definition of plant-pesticide, all substances produced by plants and intended for a pesticidal purpose are within EPA's jurisdiction, whether the plant is genetically engineered or not. However, just because a substance is considered to be a plant-pesticide, it does not necessarily mean that EPA will regulate it under FIFRA. The Agency believes there are many plant-pesticides that do not warrant any regulation under FIFRA because they pose a low probability of risk and will not cause unreasonable adverse effects on the environment. One category of plant-pesticides that EPA believes does not warrant regulation are those that will not cause new exposures to nontarget organisms. EPA is proposing to exempt from FIFRA regulation those plant pesticides that are not new to the plant (i.e., derived from closely related plants). Thus the Bt delta-endotoxin would not be exempt when it is produced in corn, for example, because the delta-endotoxin is derived from a bacterium rather than from a plant that is closely related to corn. A pesticidal substance that is naturally produced by a certain variety of corn and is introduced into another variety of corn, however, would be exempt. Another category that EPA is proposing to exempt are those plant-pesticides that would not be expected to adversely affect nontarget organisms because they are less likely to be directly toxic because of their mechanism of action. This category consists of plant-pesticides that act primarily by affecting the plant so that pests are inhibited from attaching to the plant, penetrating the plant's surface, or invading the plant's tissue. Thus a substance that acts by causing a structural barrier to pest penetration in the plant would be exempt. EPA also believes that coat proteins from viruses pose low risks and do not warrant regulation under FIFRA.

In addition to the low potential for risk associated with these categories of plant-pesticides, EPA believes that these plant-pesticides may have significant benefits associated with them because they could be used as alternatives to more toxic and persistent conventional pesticides.

Although EPA scientists and the members of the SAP and BSAC that have evaluated these exemptions believe that the plant-pesticides proposed for exemption pose low risks, many environmentalists are concerned that the exemptions are too broad. These concerns seem to stem, in large part, from the uncertainty surrounding many of the issues and the historical lack of experience with plant-pesticides. Some have suggested that EPA should require ongoing monitoring of exempt plant-pesticides. In response to this concern, EPA is considering proposing a regulation that would require reporting of adverse effects information for exempt plant-pesticides. This regulation would be similar to FIFRA Section 6(a)(2), which requires reporting of unreasonable adverse effects information for all registered pesticides. If EPA does impose such a requirement, the next issue to consider is how EPA will react if it finds that a particular plant-pesticide, or category of plant-pesticides, is riskier than EPA believed when it exempted it. Currently, under FIFRA Section 25(b), to exempt a pesticides, EPA must go through notice and comment rule making. It follows that to repeal an exemption, EPA also may be required to go through rule making. Rule making can be a lengthy process, particularly when coupled with the FIFRA requirement of submittal of all proposed and final regulations to the SAP and USDA for comment. A statutory amendment that would authorize EPA to repeal exemptions with a more abbreviated process would enable EPA to more quickly gain regulatory control over plant-pesticides found to pose unreasonable adverse effects.

Under the proposal, once it is determined that a substance is a plant-pesticide subject to FIFRA regulation, the regulatory process is similar to, with some modification, the regulatory process for all pesticides. Prior to sale or distribution, if a crop is to be used as food or feed at any test acreage, an EUP would be required. For crops that will not be used as food or feed, and if

subject to the authority of the Plant Pest Act, an EUP would be required when environmental testing will be on greater than 10 acres of land or greater than one surface acre of water. Currently, for all pesticides, the 10-acres requirement is triggered when the cumulative acreage of environmental tests reaches ten acres. In the proposal, EPA indicates that it is considering changing this requirement for plant-pesticides so that an EUP is required when a single environmental test exceeds 10 acres. EPA is also considering a number of other options for EUP triggers. One option is to utilize APHIS's determination that a plant is no longer a regulated article as the point at which regulatory responsibility is handed off from APHIS to EPA. If a plant-pesticide is not subject to the authority of the Plant Pest Act, an EUP would be required at first introduction into the environment regardless of acreage. If a producer has been granted an exemption by APHIS from permitting requirements under the Plant Pest Act, an EUP would be required at the time the exemption is granted.

Before sale or distribution of a plant-pesticide, a producer must obtain a registration under FIFRA Section 3 if the plant-pesticide is not otherwise exempt. Where there is food or feed use at sale or distribution, the potential registrant must further fulfill the necessary FFDCA obligations. FIFRA Section 3 requires that all registered pesticides be labeled. Labeling includes both the written, printed, or graphic material on, or attached to, the pesticide or any of its containers or wrapper and all other written, printed or graphic material accompanying the pesticide at any time. An improperly labeled pesticide is considered to be misbranded and in violation of FIFRA. As noted earlier, EPA generally relies on labeling requirements to impose risk reduction measures on the use of traditional pesticide products. For example, EPA regulations at 40 CFR 156.10 contain extensive labeling requirements dealing with, among other things, warnings and precautionary statements and directions for use. Other labeling restrictions are imposed, case by case, through the registration process. Restrictive labeling may include anything from requirements that personal protective equipment such as gloves and respirators be used to reduce the risk to pesticide users, to the requirement that a buffer zone be provided around fields to prevent risks to bystanders from spray-drift, to geographic restrictions on the use of certain pesticides to reduce the risk to endangered species or other beneficial organisms that occur in a limited geographical area. These labeling restrictions are translated into use restrictions via FIFRA Section 12(a)(2)(G), which provides that it is unlawful for any person to use any registered pesticide in a manner inconsistent with its labeling. EPA has stated that it recognizes that many types of restrictive labeling that it relies on to regulate traditional chemical pesticides may not be appropriate for plant-pesticides. For example, geographical limitations on the use of the plant-pesticide may not be meaningful if the plant that produces the pesticide can reproduce and spread in the environment beyond those geographical limits. Similarly other use restrictions (e.g., prohibiting use within 50 feet of a stream, river, or lake) may not be effective if seeds from plants that produce plant-pesticides are saved and planted during subsequent growing seasons. Such seeds would not be labeled, and the farmers using these seeds might not even be aware that the seeds were from plants that had been engineered to produce a plant-pesticide. Although EPA recognizes that the more typical labeling restrictions may not be meaningful for plant-pesticides, it is not yet clear how EPA will adapt its regulatory practice to these new forms of pesticides. The success of EPA's plant-pesticide program will depend, in large part, on EPA's ability to diverge from its historical reliance on labeling restrictions to achieve risk reduction. Because traditional restrictive labeling is not likely to result in true risk reduction for plant-pesticides, EPA will need to consider whether registrations should not be granted for plant-pesticides that pose significant risks in the absence of meaningful risk reduction. Despite the problems with traditional risk reduction labeling, EPA recognizes that other forms of labeling may be useful for plant-pesticides. Specifically, EPA is considering requiring labeling on bags of seeds containing plant-pesticides that inform farmers or other users of the type of pesticide that the plants will produce and against which pest it is active. This information could help prevent unnecessary application of additional pesticides to the plants that already produce plant-pesticides.

If a plant-pesticide is being used in food or feed, EPA has two options in its regulation under FFDCA: It can set a tolerance for the plant-pesticide, or it can exempt the plant-pesticide from the requirements of a tolerance. FIFRA and FFDCA are independent statutes: A plant-pesticide that is exempt from regulation under the proposed scope for FIFRA is not necessarily exempt from regulation under FFDCA. Moreover, the two Acts have different, but overlapping, purposes: Under FIFRA, EPA considers *all* environmental and human health risks, whereas, under FFDCA, EPA focuses on the risks posed by human dietary consumption. In the proposal, under FFDCA Section 408(c), EPA would exempt certain categories of plant-pesticides from the requirement of a tolerance. The plant-pesticides that EPA believes warrant review, and thus would not be exempted, are those that are most likely to result in new or different dietary exposures. The proposal would exempt the following:

1. Plant-pesticides produced in food and derived from closely related food or nonfood plants.
2. Plant-pesticides produced in food and derived from food plants that are not closely related to the recipient food plant and would not result in significantly different dietary exposure when produced in the recipient food plant. "Results in significantly different dietary exposure" can be interpreted in a number of ways:
   a. The pesticidal substance is produced in inedible portions of the source food plant, but in the recipient plant, the pesticidal substance is present in the plant's edible portions.
   b. The pesticidal substance is produced in the immature but not in the mature edible portions of the source food plant, but in the recipient plant,

the pesticidal substance is present in the mature edible portions.

c. The pesticidal substance is from a source food plant normally cooked or processed and is produced in a recipient plant that is not normally cooked or processed prior to consumption.

d. The pesticidal substance is derived from a source food plant that is not a major crop for human dietary consumption (i.e, not wheat, corn, soybeans, potatoes, oranges, tomatoes, grapes, apples, peanuts, rice, and beans or any other crop that EPA has determined is a major crop for human dietary consumption) and is introduced into a recipient plant that is a major crop for human dietary consumption.

EPA also is proposing to exempt from the requirement of a tolerance the coat proteins from plant viruses and nucleic acids. EPA believes that tolerances are not necessary for coat proteins from viruses because virus-infected plants have always been a part of the human diet without any known adverse human health effects. It is necessary for EPA to address nucleic acids under FFDCA because they are considered part of the pesticidal active ingredient. EPA plans to exempt these substances from the requirement of a tolerance, however, because nucleic acids are present in the cells of every living organism and thus are ubiquitous in the food supply. Because of their ubiquity in the food supply and because they lack any toxicity when consumed in food, EPA does not believe tolerances for nucleic acids are necessary to protect the public health.

During the five years since it first published its plant-pesticide proposal, EPA has issued a number of registrations and granted several tolerance exemptions for a variety of plant-pesticides. EPA has not, however, completed or published a set of final regulations governing plant-pesticides. The reason for the delay most likely stems from the controversies surrounding the plant-pesticide debate. One of the most significant controversies involves the strongly opposing views on whether genetically engineered food should be required to be labeled. Many, particularly in the European Community, believe that all genetically modified foods should be labeled so that consumers are fully informed. Thus EPA's position on whether to require labeling may have serious implications also with regard to international trade. In addition serious concerns have arisen regarding the risk that plants producing pesticidal substances such as the Bt toxin on a continual basis may hasten the development of pest resistance to these beneficial pesticides. This is a very difficult issue that EPA has not yet come to grips with. These and other issues will have to be resolved before EPA's plant-pesticide program will be fully in place. Perhaps only time and experience will tell how to address these difficult and uncertain issues.

ACKNOWLEDGMENT

Portions of this article are reprinted from the previously published article by the author, *Genetically Engineered Plant Pesticides: Recent Developments in the EPA's Regulation of Biotechnology*, 7 UFJL & Pub. Pol'y 257 (1996), with the permission of the University of Florida Journal of Law and Public Policy.

BIBLIOGRAPHY

1. R.A. Bohrer, *U. Pitt. Law Rev.* **55**, 607–610 (1994).
2. K. Bosselmann, *Colo. J. Int. Environ. Law Policy* **7**, 111–148 (1996).
3. D.L. Burk, *U. Pitt. Law Rev.* **55**, 611–633 (1994).
4. 7 U.S.C. §136–136y (1988); Supp. (1988).
5. 21 U.S.C. §301 et seq. (1988).
6. R. Carson, *Silent Spring*, 1962.
7. Pimmentel et al., *Bioscience* **41**, 402 (1991).
8. S. Matten, W. Schneider, B. Slutsky, and E. Milewski, in L. Kim, ed., *Advanced Engineered Pesticides*, 1993.
9. 57 Fed. Reg. 6753 (1992).
10. S.H. Mantell, J.A. Matthews, and R.A. McKee, *Principles of Plant Biotechnology: An Introduction to Genetic Engineering in Plants*, 5, Blackwell Scientific, Oxford, UK, 1985.
11. P.J. Goss, *Calif. Law Rev.* **84**, 1395–1435 (1996).
12. D.J. Earp, *Environ. Law* **24**, 1633 (1994).
13. J.J. Kim, *Fordham Int. Law J.* **16**, 1160 (1993).
14. J.E. Beach, *Food Drug Law J.* **53**, 181–191 (1998).
15. C.M. Steen, *Ariz. Law Rev.* **38**, 763–791 (1996).
16. U.S. General Accounting Office, Food Safety and Quality, *Innovative Strategies May be Needed to Regulate New Food Technologies*, U.S. Government Printing Office, Washington, DC, 1993.
17. Environmental Defense Fund, A Mutable Feast: Assuring food safety in the era of genetic engineering, *New York*, October 1, 1991.
18. M. Mellon, *Biotechnology and the Environment: A Primer on the Environmental Implications of Genetic Engineering*, National Wildlife Federation, Washington, DC, 1988.
19. M. Mellon, in *Agricultural Biotechnology at the Crossroads: Biological and Social Institutional Concerns*, National Agricultural Biotechnology Council Report No. 3, Washington, DC, 1991.
20. W. Fehr, R. Goldberg, and J. Fearn, in *Agricultural Biotechnology at the Crossroads: Biological and Social Institutional Concerns*, National Agricultural Biotechnology Council Report No. 3, Washington, DC, 1991.
21. W. Fehr, in *Agricultural Biotechnology at the Crossroads: Biological and Social Institutional Concerns*, National Agricultural Biotechnology Council Report No. 3, Washington, DC, 1991.
22. T. Hoban and P. Kendall, in *Agricultural Biotechnology: A Public Conversation about Risk* (National Agricultural Biotechnology Council Report No. 5, Washington, DC, 1993 (citing Office of Technology Assessment, A New Technological Era for American Agriculture, Summary Report 1992).
23. R.A. Bohrer, *U. Pitt. Law Rev.* **55**, 653–679 (1994).
24. 49 Fed. Reg. 50856 (1984).
25. 51 Fed. Reg. 23302 (1986).
26. 49 Fed. Reg. 40,659 (1984).
27. 58 Fed. Reg. 5878 (1993).
28. 59 Fed. Reg. 45,600 (1994).
29. 59 Fed. Reg. 60,496 (1994).
30. 59 Fed. Reg. 60,519 (1994).
31. 59 Fed. Reg. 60,535 (1994).

32. 59 Fed. Reg. 60,542 (1994).
33. 59 Fed. Reg. 60,545 (1994).
34. 40 CFR §152.20.

See other entries AGRICULTURAL BIOTECHNOLOGY; ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE; FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, CONGRESS, EPA.

AGRICULTURAL BIOTECHNOLOGY, LAW, AND FOOD BIOTECHNOLOGY REGULATION

HENRY I. MILLER
Stanford University
Stanford, California

OUTLINE

INTRODUCTION

Governmental oversight of the foods made with the techniques of the new biotechnology offers examples of the spectrum of possible approaches and their ripple effects. FDA's official risk-based policy has served the public interest well, but the agency appears to be retreating to a more politically correct, defensive posture that is closer to the European approach. The European experience provides striking illustrations of what can happen when regulatory policy is built on a foundation of invalid scientific assumptions, gratuitous controversy, and political and ideological goals. The outcome is expensive, expansive, and irrational regulation — which leads, in turn, to narrower application of the technology, and fewer benefits. The arguments that have been marshaled during the policy "debate" over biotechnology regulation are revealing. Those who would encourage unnecessary regulation sometimes argue that, in the face of uncertainty, it is only prudent to "err on the side of safety," to avoid taking any chances, and act instead on the basis of the worst-case scenario — the "precautionary principle." A related argument is that even if only a handful of adverse events (e.g., toxicity caused by the consumption of genetically engineered plants) are prevented by government oversight regimens, not to act would be unconscionable and would amount to putting a price on human life. But the principles of "erring on the side of safety" and the pricelessness of life do not withstand rigorous scrutiny. What appears to be the "safe" choice may, upon analysis, actually pose greater risk. One must consider the risks of various alternative courses of action; forgoing new technology can put lesser theoretical risks ahead of known, palpable, existing ones.

Despite their many advantages, gene-spliced organisms — sometimes called genetically modified (GM) — are controversial in some parts of the world. In Europe, for instance, there has been widespread public opposition to importing gene-spliced corn and soybeans. Foods produced through gene splicing must be labeled as such, and most major supermarket chains and food producers have said they will not sell them. Threats by antitechnology activists of boycotts and hostile publicity have induced several companies doing business in the United States to reject gene-spliced ingredients used to make their products. For instance, the Japanese breweries Kirin and Sapporo, whose beer is popular in the United States, have announced that they will phase out their use of gene-spliced corn. Two of the largest U.S. producers of baby food, Gerber and Heinz, have promised not to use gene-spliced materials — even if what they use instead is nutritionally inferior or less safe. As an example, they will reject materials from corn plants modified so that they do not need to be sprayed with toxic chemical insecticides.

Scientists around the world agree that introducing genes from other organisms does not make plants less safe either to the environment or for humans to eat. Dozens of new plant varieties produced through hybridization and other traditional methods of genetic modification enter the marketplace each year, without special scientific review or labeling for consumers. Moreover many of these foods on the market are from "wide crosses," hybridizations in which genes are moved from one species or one genus to another to create a variety of plant that does not exist in nature. While such changes may sound dramatic, the results are as mundane as a tomato that is more resistant to disease, or that has a thicker skin that won't be damaged during mechanical picking. Plants that have undergone those slight but important alterations have been an integral part of European and American diets for decades. However, these scientific and commercial realities have often been lost in the nether world of regulatory politics.

FOOD BIOTECH'S VENERABLE PAST AND PROMISING FUTURE

The almost unimaginably wide spectrum of foods consumed throughout the world owes its existence to both the ingenuity of history's cooks and the practitioners of biotechnology. Microorganisms were creating and improving humans' food and drink long before anyone knew that microorganisms existed. In time — but still without knowing what was happening biologically — early practitioners of biotechnology learned to exploit the fermentative action of microorganisms to produce such things as cheese, bread, and alcoholic beverages. Still later, food producers began to isolate favored microbial cultures with highly descriptive names such as *Penicillium roqueforti* (used to make

Roquefort cheese) and *Lactobacillus san francisco* (for San Francisco-style sourdough bread). It is no coincidence that some of the past century's most sophisticated microbiology applied to beverage production has been performed in the laboratories of companies like Guinness, Carlsberg, Kirin, and Bass. Vastly popular regional foods produced by fermentation include milk products (yoghurt, sour cream, buttermilk, kefir), preserved vegetables (cabbage, olives), tempeh, sufu, tofu, soy sauce, and natto.

The modification of crop plants has been performed ever since ancient agriculturists selected and cross-bred plants with desirable traits, often creating domesticated relatives of wild species. The rediscovery in 1900 of Mendel's concepts of inheritance ushered in the scientific application of genetic principles to crop improvement. Since then, each scientific advance has increased our ability to improve predictably the *genotype* (and more important for the farmer, food manufacturer, and consumer, the *phenotype*). Currently a combination of several techniques is routinely used to improve plants. For example, an existing plant might have been modified by many generations of classical breeding and selection, and more recently by techniques developed during the past half-century, including somaclonal variation and wide crosses with embryo rescue. These plants are being further improved by the newer molecular techniques such as recombinant DNA (rDNA), or "gene splicing," and can then be reintroduced into a classical breeding program from which its descendants eventually will be released into commerce.

In this century plant breeders have increasingly used hybridization to transfer genes from certain noncultivated plant species to a variety of a different (but closely related) species. These "interspecific" transfers of traits from wild species to domesticated relatives in the same genus stimulated attempts at even wider crosses, including those between members of different genera. These "wide crosses," that transcend natural barriers to mating, have been facilitated by "embryo rescue" or culture techniques in which a sexual cross yielding a viable embryo but abnormal endosperm is "rescued" by culturing the embryo. This is done by providing the hybrid embryo with the life support normally supplied in the early stages of development by maternal tissue and the endosperm. A number of plants resulting from wide crosses have been used in further breeding, extensively field tested, and marketed in the United States and elsewhere. These plants include commonly available varieties of tomatoes, potatoes, corn, oats, sugarbeets, rice, and bread and durum wheat (1).

The use of molecular techniques for the genetic manipulation of plants enables scientists to direct the movement of specific and useful segments of genetic material readily between unrelated organisms. These techniques offer several advantages and complement existing breeding efforts by increasing the diversity of genes and germ plasm available for incorporation into crops. The numerous molecular techniques for genetic manipulation of plants can be divided into two main types — vectored and nonvectored. Vectored modifications rely on the use of biologically active agents, such as plasmids and viruses, to facilitate the entry of the foreign gene into the plant cell. Nonvectored modifications rely on the foreign genes being physically inserted into the plant cell by such methods as electroporation, microinjection, or particle guns. In both kinds of approaches, new DNA enters the plant's genome and is stably maintained and expressed. A landmark report from the U.S. National Research Council concluded that:

> Recombinant DNA methodology makes it possible to introduce pieces of DNA, consisting of either single or multiple genes, that can be defined in function and even in nucleotide sequence. With classical techniques of gene transfer, a variable number of genes can be transferred, the number depending on the mechanism of transfer; but predicting the precise number or the traits that have been transferred is difficult, and we cannot always predict the phenotypic expression that will result. With organisms modified by molecular methods, we are in a better, if not perfect, position to predict the phenotypic expression (2, p. 13).
>
> Crops modified by molecular and cellular methods should pose risks no different from those modified by classical genetic methods for similar traits. As the molecular methods are more specific, users of these methods will be more certain about the traits they introduce into the plants (2, p. 3).

Far from eliciting concern, techniques that yield a better-characterized and more predictable plant variety should be welcomed as a means for improving food. The new biotechnology lowers even further the already minimal risk associated with introducing new plant varieties into the field and the food supply. The use of the latest biotechnology techniques makes the final product even safer, as it is now possible to introduce pieces of DNA that contain one or a few well-characterized genes. In contrast, the older genetic techniques transferred a variable number of genes haphazardly. Users of the new techniques can be more certain about the traits they introduce into the plants and about the presence of unwanted, deleterious genetic changes. Thousands of products from plant varieties crafted with the older techniques have entered the marketplace in the last three or four decades, and only three products (two squash varieties and one potato type) had unsafe levels of toxins; in addition one celery variety caused allergic skin reactions in some farm and supermarket workers. But today's more precise gene-splicing techniques mitigate against any repetition. A group of chefs who announced a boycott of biotechnology-produced foods in 1990, lacked perspective on the new products' pedigree — that is, on the continuum between conventional and new biotechnology. They were against the use of plants engineered with the newest, most precise, and sophisticated techniques, while they lacked any scruples about using the mutant peaches we call "nectarines," the genetic hybrid (of tangerine and grapefruit) known as "tangelos," or the genetically improved oats, rice, and other plants that have resulted from wide crosses (3).

Many of the improvements introduced by gene-splicing enable plants to grow with less agricultural chemicals such as pesticides and herbicides, and in regions that have salty or other low-quality water. They offer increased yields with lower inputs, and diminish the runoff of chemicals

into waterways, so they are favorable to the environment. It is difficult, however, to pass on inflated regulatory costs for these kinds of improvements whose value is obscure to consumers.

FDA'S OVERSIGHT OF NEW CROP VARIETIES

FDA, which is responsible for the safety and wholesomeness of the nation's food supply (excepting only most meats, which are regulated by the U.S. Department of Agriculture, USDA), monitors the continuing progress of new biotechnology-derived products. The agency's policy on foods derived from new plant varieties — whether crafted with conventional or new biotechnology or other techniques — was published in 1992. It set out a carefully considered, scientific, and "transparent" — that is, clear and predictable — regulatory approach.

The 1992 Policy

FDA's approach to safety assessment of new varieties of crops developed by both traditional and newer methods of genetic modification (4) (Fig. 1) is based on the agency's long-standing oversight of "old" biotech-derived varieties commonly introduced into the U.S. marketplace (5). Foods derived from new plant varieties are not routinely subjected to extensive scientific tests for safety, although there are exceptions. The usual practices employed by plant breeders — such as chemical and visual analyses and taste testing — are generally recognized as adequate for ensuring food safety. Additional tests, however, are performed when required by the product's history of use or scientific judgment. For example, potatoes are tested for the glycoalkaloid solanine, because this natural toxicant has been present at toxic levels in some new potato varieties.

FDA's 1992 regulatory approach identifies scientific and regulatory issues related to characteristics of foods that raise safety questions and that elicit a higher level of FDA review. Such characteristics, which are further discussed below, include the presence in the new variety of a substance that is completely new to the food supply, the presence of an allergen in an unusual or unexpected milieu, or increased levels of toxins that are normally found in foods. Consistent with scientific consensus about recombinant DNA techniques, the use of any particular technique(s) of genetic manipulation does not in itself determine the need for or the level of governmental review.

The "Guidance to Industry" section of the 1992 policy statement instructs developers to consider initially the characteristics of the host plant that is being modified, the donor organism that is contributing genetic information,

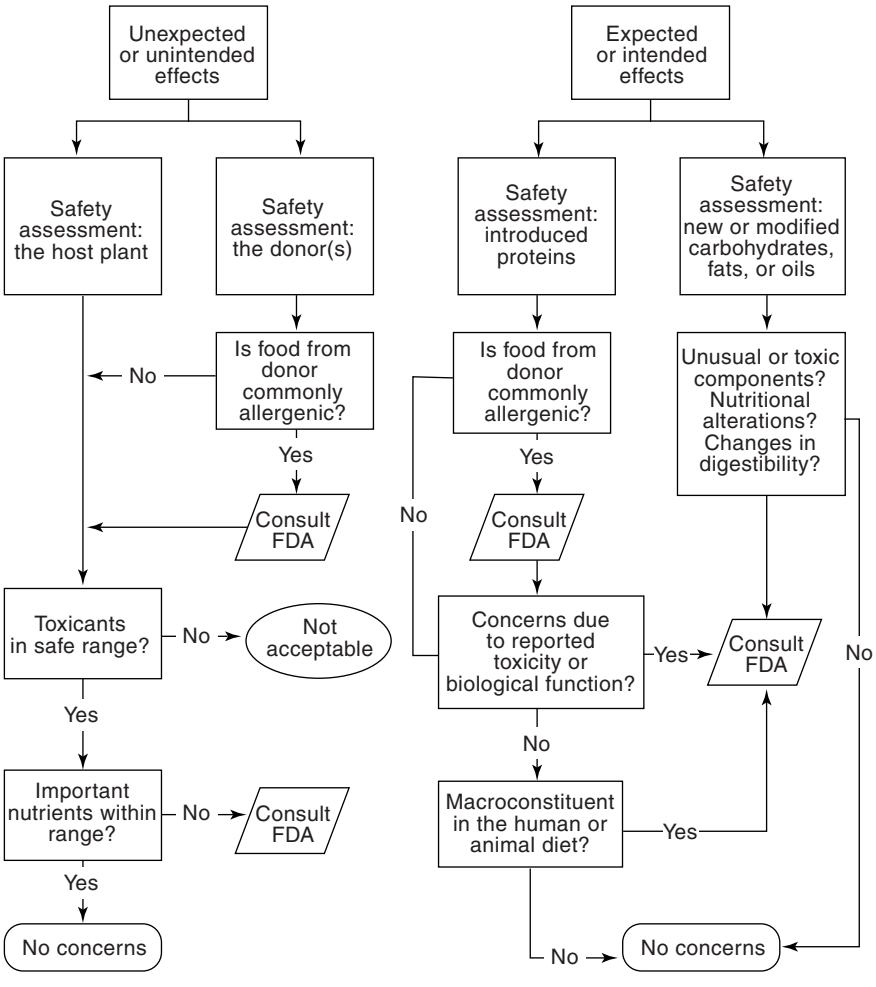

Figure 1. Food and Drug Administration (FDA) safety assessment of new varieties of crops.

and the genetic material and substances being introduced or modified. The guidance section also provides criteria that enable developers to determine whether a substance intentionally introduced or altered by genetic modification will require premarket approval as a food additive. In general, premarket review is not required for introduced or modified proteins of known function if they are derived from food sources or are substantially the same as existing food substances, are not known to be toxic or to raise food safety concerns, and will not be a major constituent of the diet. New carbohydrates with unusual structural or functional groups or oils that contain new, unusual fatty acids may require premarket approval as food additives.

The question whether biotech-derived foods should be labeled as such has received a disproportionate amount of attention. The primary reason is not ambiguity in either the science or law; rather, it is that the issue has been a focus of activists' attention. Their attention to this issue is yet another manifestation of anti-biotechnology activists' attempts to use overregulation to retard the progress of every stage of biotechnology R&D. They have lobbied against laboratory research, then against scale-up and against testing and commercial uses of products. Finally, thwarted in their desire to get the FDA to require clinical trials or case by case evaluation of biotech-derived foods, activists have retreated to demanding labeling that would reveal when new biotechnology techniques were used in a food's manufacture. The ostensible rationale for such a requirement is that information is power and that consumers can never know too much about the products they buy. Especially for foods, the more information the better, goes the mantra. But that's not necessarily true. A message can mislead and confuse consumers if it is irrelevant, unintelligible, or crafted to tell only part of the truth. Moreover a requirement for labeling carries added production expenses and raises costs to both producers and consumers that can constitute a barrier to the development of and access to new products.

To serve the consumer best, regulation should focus on genuine risks and should require only disclosure of information about a food's origin or use that is relevant to safety and that supports informed choice. Mandatory labeling of all biotech foods would achieve none of this.

Labeling has become a key issue for gene-splicing applied to food production. Labels that warn of gene-spliced ingredients in food are fundamentally different from the labels currently on food — that list calories, fat content, and so on — because these represent a modest, one-time expense for food producers. By contrast, gene-spliced fruits, vegetables, and grains would have to be continually segregated through all phases of production — planting, harvesting, processing, and distribution — adding costs and compromising economies of scale. The need to segregate gene-spliced foods, especially the thousands of processed foods that contain small amounts of derivatives of corn or soybeans, would raise production costs and pose a particular disadvantage to products in this competitive market with low profit margins.

FDA's long-standing commonsense approach to food labeling has been that label information must be both accurate and "material." FDA does not require a "product of biotechnology" or "genetically engineered" label for foods from plants or animals that have been improved with rDNA techniques. In the 1992 food policy statement, the FDA said that labeling is required "if a food derived from a new plant variety differs from its traditional counterpart such that the common or usual name no longer applies to the new food, or if a safety or usage issue exists to which consumers must be alerted." The statement of policy also emphasized that, as for other foods derived from new plant varieties, no premarket review or approval is required unless the characteristics of the new biotechnology products raise explicit safety issues. It noted that these safety issues could be raised by food from new plant varieties however they were created. The safety issues include the introduction of a substance that is new to the food supply (and, hence, lacks a history of safe use), increased levels of a natural toxicant, changes in the levels of a major dietary nutrient, and transfer of an allergen to a milieu where a consumer would not expect to find it (e.g., a peanut protein transferred to a potato).

FDA clarified that if a new food raises any of these safety issues, it could be subject to FDA regulations for premarket testing, product labeling, or removal from the marketplace. FDA cited the example of new allergens in a food as a possible material fact whose omission could make a label misleading. The agency reiterated that the genetic method used in the development of a new plant variety is *not* considered to be material information because there is no evidence that new biotech foods are different from other foods in ways related to safety. Therefore FDA said that product labeling will not be required to include the method of development of a new plant variety. Biotech foods would not be required to be labeled as such.

The 1992 FDA policy statement has already been tested and validated. A 1996 report in the *New England Journal of Medicine* reported that allergenicity common to Brazil nut proteins was transferred into soybeans by genetic engineering and was readily identified by routine procedures (6). The plant breeder, Pioneer Hi-Bred International, was required to and did consult with the FDA during product development. During the course of consultation and subsequent analysis, the allergenicity was identified. Confronted with the dual prospects of potential product liability and the costs of labeling all products derived from the new plant variety, the company abandoned all plans for using the new soybeans in consumer products. Not a single consumer was exposed to or injured by the newly allergenic soybeans. In what might be considered a "positive control," the system worked.

The approach taken by FDA in its 1992 policy statement is consistent with scientific consensus that the risks associated with new biotechnology-derived products are fundamentally the same as for other products (2). Dozens of new plant varieties modified with traditional genetic techniques (e.g., hybridization and mutagenesis) enter the marketplace every year without premarketing regulatory review or special labeling (7). As discussed above, many of these products are from "wide crosses" in which genes have been moved across natural breeding barriers (without rDNA techniques). None of these plants exists in nature. None requires or gets a premarket review by a government

agency. (Safety tests by plant breeders are primarily taste and appearance and, in the case of plants with high levels of known intrinsic toxicants — e.g., tomato and potato — levels of certain alkaloids.) Nonetheless, they have become an integral, familiar, and safe part of our diet: wheat, corn, rice, oats, black currants, pumpkins, tomatoes, and potatoes.

There are other reasons why special regulations and labeling requirements are often not in the best interest of consumers. As food producers know well, requiring a label can add significantly to the production costs of certain foods, particularly those that are produced from pooled fresh fruits and vegetables. To maintain the accuracy of labels, recombinant DNA-modified fruits and vegetables would have to be segregated through all phases of production — planting, harvesting, processing, and distribution — which adds costs and eliminates economies of scale. Added production costs, in turn, raise consumer prices and disadvantage products in the highly competitive, low profit-margin marketplace of processed foods.

Superfluous labeling requirements for new biotech products would constitute, in effect, an unwarranted and punitive tax on the use of a new, superior technology. The requirement would exact excess costs and reduce profits to plant breeders, farmers, food processors, grocers, and others in the distribution pathway. The power of regulatory disincentives is such that this burden could virtually eliminate new biotechnology tools from food research, development and production. For example, as required in the European Union, the United Kingdom introduced mandatory labeling of gene-spliced foods in 1998, which Britain's agriculture minister called "a triumph for consumer rights to better information," and which the country's additives and novel-foods chief regulator characterized as "a question of choice, of consumer choice." But as a direct result of the labeling law, there's hardly any choice there now at all: The new law sparked a stampede by manufacturers, retailers, and restaurant chains to rid their products of any genetically modified ingredients so that they would not have to alter their labels and risk losing sales.

It is unclear how far these scientifically dubious food label requirements will extend. Will special labels be required for foods such as pizza or burritos containing cheese made with new biotech-produced chymosin (rennin), for chickens raised on feed from new biotech-manipulated corn, and for cattle vaccinated with a new biotech vaccine? Will labels be required if highly sensitive analytical techniques detect one part per million of recombinant DNA in a food?

An analysis of the economic impacts of a labeling requirement for new-biotech foods by the California Department of Consumer Affairs (CDCA) predicted that the additional costs would be "substantial," and that "while the American food processing industry is large, it is doubtful that it would be either willing or able to absorb most of the additional costs associated with labeling biotech foods" (7). The analysis concluded that "there is cause for concern that consumers will be unwilling to pay even the increased price for biotech foods necessary to cover biotechnology research and development, much less the additional price increases necessary to cover the costs associated with labeling biotech food."

The CDCA assessment implies another outcome of unwarranted but compulsory labeling. Overregulation will reduce competition and, therefore, increase prices, and overpriced biotech products would be limited to upscale, higher-income markets. Wealthier consumers would be able to pay more for the "boutique" products, while the less affluent would simply do without them.

It is noteworthy that FDA's current approach to labeling was upheld indirectly by a federal appeals court, which found in a pivotal 1996 decision regarding another product of biotechnology that food labeling cannot be compelled just because some consumers wish to have the information. That case, *International Dairy Foods Association v. Amestoy* (92 F. 3rd 67; 2nd Cir. 1996), involved a Vermont state law requiring labeling of dairy products from cows treated with a gene-spliced protein that increases the productivity of dairy cows. In overturning the law, the appeals court found that such regulation merely to satisfy the public's alleged "right to know" is a constitutional violation of commercial free speech. "Were consumer interest alone sufficient, there is no end to the information that states could require manufacturers to disclose about their production methods," the court wrote.

Although FDA's 1992 food biotech policy is scientifically defensible and favors the public interest, as discussed below, the agency has shown a willingness to ignore scientific consensus, bow to political pressure, and accommodate activists' whims.

FDA's Volte-face on Biotech Food Policy

In 1993, only a year after publishing the progressive and scientific food biotech policy described above, FDA informally announced plans to require registration of all new biotechnology foods. FDA never published a proposal, but various agency officials announced repeatedly that one was being prepared. The new policy would have directly contradicted the widely praised 1992 policy statement that specified that new biotechnology foods would be treated in the same manner as other, similar foods. The ostensible rationale for this *volte-face* in policy was the gratuitous "controversy" over biotech foods in Europe and the United States, fueled by activists who are ideologically opposed to the new biotechnology (8).

The actions of antitechnology activists have shown that their agenda is neither a good-faith attempt to air issues of technological risk, nor an attempt to offer innovators and consumers a greater spectrum of choices. Rather, they wish to control what research is performed, what techniques are permitted, and what products are brought to market. Academic freedom, industrial innovation, free markets, and consumer choice are among their victims. The controversies about the new biotechnology are only a microcosm of that struggle. Activists' minds will not be changed by scientifically reasonable arguments, by assertions of the primacy of empirical evidence and the scientific method, or by invoking the benefits to the public of new products and choices. The activists' modest success at discouraging prospective end-users of new

biotechnology products from accepting them is worrisome, for if end-users such as food producers and consumers do not want the products, plant biologists, breeders, and farmers will stop developing and growing them, and the technology will no longer be widely used. The major exceptions will be in which the inflated price of the final product can offset the cost of making it. Inevitably, with lowered demand for and acceptance of gene-spliced products, there will be less interest in and resources for basic research on plants for food, fiber, and medicines.

Overregulation, in the form of additional risk assessment (e.g., case-by-case review) or risk management (e.g., labeling) is interpreted by the public as a warning of potential risk. This is evident from the observations of Barbara Keating-Edh, representing the consumer group *Consumer Alert*, before the National Biotechnology Policy board, a group established by Congress and located within the NIH, September 20, 1991:

> For obvious reasons, the consumer views the technologies that are *most* regulated to be the *least* safe ones. Heavy involvement by government, no matter how well intended, inevitably sends the wrong signals. Rather than ensuring confidence, it raises suspicion and doubt (9, emphasis in original).

Since 1993 FDA has generally been receptive to activists' demands. The agency has seemed more concerned about placating activists than ensuring access by consumers to the fullest array of improved foods and providing opportunities to American companies for product innovation.

At the direction of Vice President Gore's domestic policy staff, in 1993 FDA announced a policy that would require selected foods to be registered with the agency before being sold to consumers (10). Extra government scrutiny certainly makes sense when there is uncertainty about health risks or a reason to suspect a problem such as the presence of toxins or allergenic components. In this proposal, however, the FDA decided to require registration only of those foods made with the most precise state-of-the-art biotechnology techniques. While in the first instance there was to be only a requirement for *registration* of new products at FDA's Center for Food Safety and Nutrition, officials could, of course, request additional information and testing for individual products. Such a policy would confer no public-health advantage but would discourage research that could produce foods of better quality and greater variety. The inevitable result of such regulation would have been more responsibility and bureaucracy for FDA but more limited choices for consumers at the supermarket or greengrocer.

Scientific and professional groups — including the Chicago-based Institute of Food Technologists and the University of California's Systemwide Biotechnology Program — objected strenuously to the unscientific basis of the proposal. Only after the Republican-controlled 104th Congress made pointed inquiries about the plan did FDA officially withdraw the proposal — deleting it in early 1995 from the agency's regular report to OMB of regulations in preparation. The proposal was gone but not forgotten.

In July 1996 FDA began, in a peculiar and surreptitious way, to circulate the news of new requirements. A seven-page document, "Foods Derived from New Plant Varieties: Consultation Procedures," went out to state officials (11). In it, FDA adopts a pretense that the new policy is applicable to all new plant varieties and not just those produced with the new biotechnology. The agency's intent is transparent, however: Oversight of the consultation process rests with the *Biotechnology* Evaluation Team, and the degree of detail requested from the plant's developer would only be available for those crafted with new biotechnology.

Again signaling apparent willingness to retreat from its scientifically sound 1992 policy, in late 1999 the agency held a series of public meetings around the country to inform the public "about current FDA policy for assuring the safety of bioengineered foods [and to ask] whether this policy should be modified," according to the agency's press release. This was a thinly veiled invitation for anti-biotech activists to stuff the ballot box and demand more stringent regulation — and that is exactly what happened at the meetings, held in Chicago, Washington, DC, and Oakland.

In Chicago, three hundred prospective speakers showed up, the vast majority of them from radical environmental groups, to denounce gene-spliced food as, variously, unproven, dangerous, worthless, unnatural, and anti-religious. Interestingly many members of these anti-biotech organizations registered merely as "consumer," presumably so their groups could in effect have multiple representatives. There were also two panels (whose members were selected by FDA), with long-time antagonists of biotechnology heavily represented. Outside the meeting, many of the activists mugged for the cameras, staging mini-morality-plays in which, for example, children costumed as monarch butterflies fled in mock terror from a figure dressed as a huge ear of gene-spliced corn.

These most recent attempts by FDA to regulate biotech foods in a discriminatory way reverse the agency's 15-year-old guiding principle for the oversight of biotechnology: Regulation should focus on real risks and should not turn on the use of one technique or another. These tenets have provided effective oversight for thousands of new biotechnology products, including drugs, vaccines, diagnostic tests, and foods. Ironically, as discussed above, as recently as May 1992 FDA formally reiterated this policy for foods, affirming that new biotechnology foods would be treated no differently from those produced with other techniques and that oversight would be risk-based.

In several ways the FDA's new policy will discourage the application of biotechnology to foods. The data requirements are substantial; FDA lists nine categories of obligatory information:

1. Name of the bioengineered food and the crop from which it was derived.
2. Description of the various intended uses of the bioengineered food, including animal feed uses.
3. Information concerning the sources and functions of introduced genetic material.
4. Information on the purpose or intended technical effects of the modification, and its expected effects on the composition or characteristics of the food or feed.

5. Information about the identity and function of the newly introduced genetic material and new gene-expression products, including an estimate of the concentration of any expression product in the bioengineered crop and the food derived from it.
6. Comparison of the composition and characteristics of the bioengineered food with the food derived from the parental variety or other commonly consumed varieties.
7. Information about the identity and levels of toxicants that occur naturally in the food.
8. Discussion of the available information concerning the potential for altered allergenicity (ability to elicit an allergic reaction) in the bioengineered food.
9. Any other information relevant to the safety and nutritional assessment of the bioengineered food.

The detail is far greater than would be required for food products made with less-precise, less-sophisticated techniques; if applied to traditionally crafted plants these new, draconian requirements would spell the end of new varieties of apples, pears, strawberries, or wheat, for example. Imagine trying to determine the function of poorly characterized genes situated on whole chromosomes and newly introduced into a new cultivar of wheat by the wide cross-hybridization of wheat and a wild grass to which it is distantly related.

FDA's new policy will entail significant costs for the government and industry, and by extension, the public, in those few instances where food producers actually decide to apply biotechnology to foods and attempt to negotiate the new regulatory hoops. According to FDA's description of the new regulatory scheme the Biotechnology Evaluation Team will always consist of no fewer than *six* FDA staff, drawn from different parts of the agency. There will be endless and conflicting demands for information about each product, causing delay and uncertainty among manufacturers. Another disadvantage of the new regulatory regime is that every biotechnology product will be placed squarely in the sights of anti-biotechnology activists: according to FDA, the results of consultations with industry will be available on the Internet.

The policy is intentionally murky about whether developers of new biotech foods are *required* to consult with the agency, although in private conversations Dr. James Maryanski, the biotechnology coordinator in FDA's Center for Food Safety and Nutrition, has protested that these "requirements" are "only suggestions." However, the reality is unequivocal: A "suggestion" from the nation's most ubiquitous and draconian regulator is akin to an armed mugger "suggesting" you turn over your wallet. In practice, this is mandatory premarket regulation — applied uniquely to biotech-derived foods. And it is extra-legal regulation, in the sense that it is not the product of the rule making required by law.

The bottom line is that the policy will, in effect, impose a tax on the use of biotechnology for food production. This discriminatory treatment will discourage research on more varied, appetizing, and nutritious foods — research that has given us low-saturated-fat oils, seedless grapes, tangelos, and the like. American farmers and food processors will be less competitive, consumers will be deprived of new choices and the price of biotech-derived foods will be inflated.

OECD'S BIOTECH FOOD POLICY

The Paris-based Organization for Economic Cooperation and Development (OECD) has endorsed a policy for biotech foods similar to FDA's 1992 approach. In its 1993 publication, *Safety Evaluation of Foods Derived by Modern Biotechnology*, OECD invoked the concept of "substantial equivalence" (borrowed from the U.S. FDA's medical device regulations), the crux of which is that new foods that are "substantially equivalent" to other varieties should be regulated "in the same manner as their analogous conventional counterparts" (12). In other words, no additional regulatory requirements, such as notification, review, or labeling, should be required. *Safety Evaluation of Foods Derived by Modern Biotechnology* spells out clearly the rationale, theory, and practice for applying the principle of "substantial equivalence":

> Historically, foods prepared and used in traditional ways have been considered to be safe on the basis of long-term experience, even though they may have contained natural toxicants or anti-nutritional substances. In principle, food has been presumed to be safe unless a significant hazard was identified.
>
> Modern biotechnology broadens the scope of the genetic changes that can be made in food organisms, and broadens the scope of possible sources of foods. This does not inherently lead to foods that are less safe than those developed by conventional techniques. Therefore, evaluation of foods and food components obtained from organisms developed by the application of the newer techniques does not necessitate a fundamental change in established principles, nor does it require a different standard of safety.
>
> [T]he precision inherent in the use of certain molecular techniques for developing organisms for use as food should enable direct and focused assessment of safety where such assessment is desired. Knowledge obtained using these methods might also be used to approach safety assessment of new foods or food components from organisms developed by traditional methods.
>
> For foods and food components from organisms developed by the application of modern biotechnology, the most practical approach to the determination of safety is to consider whether they are *substantially equivalent* to analogous conventional food product(s), if such exist. . . .The concept of substantial equivalence embodies the idea that existing organisms used as food, or as a source of food, can be used as the basis for comparison when assessing the safety of human consumption of a food or food component that has been modified or is new.
>
> If one considers a modified traditional food about which there is extensive knowledge on the range of possible toxicants, critical nutrients or other relevant characteristics, the new product can be compared with the old in simple ways. These ways can include, *inter alia*, appropriate traditionally performed analytical measurements (for example, alkaloid levels in potatoes, cucurbatin in vegetable squash cultivars, and psoralens in celery) or crop-specific markers, for comparative purposes. The situation becomes more complex as the origins/composition/exposure experience decreases, or

if the new products lack similarity to old established products or, in fact, have no conventional counterpart.

A demonstration of substantial equivalence takes into consideration a number of factors, such as:

- knowledge of the composition and characteristics of the traditional or parental product or organism;
- knowledge of the characteristics of the new component(s) or trait(s) derived, as appropriate, from information concerning: the component(s) or traits(s) as expressed in the precursor(s) or parental organism(s); transformation techniques (as related to understanding the characteristics of the product) including the vector(s) and any marker genes used; possible secondary effects of the modification; and the characterization of the component(s) or trait(s) as expressed in the new organism; and
- knowledge of the new product/organism with the new component(s) or trait(s), including the characteristics and composition [i.e. the amount of the component(s) or the range(s) of expression(s) of the new trait(s)] as compared with the conventional counterpart(s) (i.e. the existing food or food component).

The OECD document elaborated "principles for the application of substantial equivalence to the assessment of organisms developed by the application of biotechnology":

- If the new or modified food or food component is determined to be substantially equivalent to an existing food, then further safety or nutritional concerns are expected to be insignificant.
- Such foods, once substantial equivalence has been established, are treated in the same manner as their analogous conventional counterparts.
- Where new foods or classes of new foods or food components are less well-known, the concept of substantial equivalence is more difficult to apply; such new foods or food components are evaluated taking into account the experience gained in the evaluation of similar materials (for example, whole foods or food components such as proteins, fats or carbohydrates).
- Where a product is determined not to be substantially equivalent, the identified differences should be the focus of further evaluations.
- Where there is no basis for comparison of a new food or food component, that is, where no counterpart or similar materials have been previously consumed as food, then the new food or food component should be evaluated on the basis of its own composition and properties.

The consideration of safety may include the need to evaluate possible effects occurring through cooking or other processing. For example, trypsin inhibitors from certain leguminous plants, such as the cowpea trypsin inhibitor, have a long history of safe consumption when properly cooked. However, if the cowpea trypsin inhibitor is expressed in other plants, the safety question relates to whether the normal use of these plants as food involves cooking sufficient for its inactivation.

Another consideration [related to whether a new food is substantially equivalent to another] is the influence of the newly introduced modification(s) on the nutritional value of the food or food components(s). For the majority of modifications being carried out, such changes are unlikely. Nonetheless, when modifications are directed at metabolic pathways of key macro- or micro-nutrients, the possibility of an impact on nutritional value is increased. Such impacts are of potential significance in cases where the modified food or food component may become a major dietary source of the nutrient affected.

It is obvious from the foregoing that in order to apply substantial equivalence generally, as well as to specific cases, judgments by regulators are necessary. And therein lies what has become in practice an anomaly: Contrary to the concept as conceived at the OECD, national regulators and others have often defined virtually *any* change wrought by molecular techniques as yielding a food or food component that falls outside the realm of substantial equivalence and that, therefore, requires more extensive review and evaluation. Although FDA does not apply the term "substantial equivalence" to its oversight of food, the concept is implicit. As described above, FDA's 1992 official policy defines certain safety-related characteristics of new foods that, if present, define "non-substantial-equivalence" and require greater scrutiny by the agency. These include the presence of a substance that is completely new to the food supply, an allergen presented in an unusual or unexpected way (e.g., a peanut protein transferred to a potato), changes in the levels of major dietary nutrients, and increased levels of toxins normally found in foods. (The absence of such characteristics, in effect, defines foods that are substantially equivalent to antecedent products.) Foods lacking characteristics that raise these safety issues need not be subject to premarket FDA review.

OTHER INTERNATIONAL AND SUPRANATIONAL APPROACHES

The European Union

The European Union (EU) announced controversial rules for the labeling and sale of new biotechnology-derived foods in December 1996, after months of acrimonious debate (13). The now-mandatory labels will add significantly to the costs of processed foods made from fresh fruits and vegetables. The precise costs will vary according to the product. But a company using a gene-spliced, higher-solids, less-watery tomato (more favorable for processing), for example, must bear the additional costs of segregating the product at all levels of planting, harvesting, shipping, processing and distribution. Labels must appear on minestrone soup, indicating the presence of gene-spliced tomato, potato or other products (at least any amount above an arbitrary one percent threshold). The added production costs are a particular disadvantage to products in this competitive, low profit-margin market segment, and at best, will likely relegate many gene-spliced products to the status of expensive "boutique" foods, out of the reach of less affluent consumers (15). As discussed above, labeling requirements have virtually eliminated biotech foods from the shelves of European retailers.

The EU compromise was reached after five years of negotiations by a joint committee of the European

Parliament and the EU Council of Ministers (which represents the 15 states). The Council and the European Commission had preferred to require labeling only when the new food or ingredient was "significantly different" from its predecessors; but the European Parliament had its way, and labeling is required for "live" genetically modified products — those that could, in theory, grow if put in soil, such as tomatoes or potatoes. The compromise has not mollified the radicals. A new "technical amendment" to the regulation adopted by the European Commission on April 2, 1997, would require the labeling of seed products that will give rise to transgenic plants. The European Novel Food Regulation, with or without the April 1997 amendment, is irrelevant to public health. A label that says "genetically modified" provides no useful or material information to consumers — but at significant economic and societal cost.

Quite apart from gene-splicing considerations, other parts of the regulation also fail to take scientific principles and precedents into account. For example, new varieties of wheat improved by the introduction of genes from hardy grasses (a common plant-improvement strategy) might be deemed "different in comparison with a conventional food or food ingredient." Under which circumstances, says the regulation, the varieties are "no longer equivalent" to preexisting foods or ingredients, and would require special — and costly — labeling. (Consider, also, that using sophisticated analytical techniques, the chemical composition of English potatoes can easily be distinguished from those grown in Italy; under the regulation, the two varieties — even if the same species and cultivar — would arguably be nonequivalent and need to be distinguished by labeling.)

As often happens with political compromises by poorly informed, paternalistic politicians, the citizenry are compromised by an outcome that makes neither scientific nor economic sense. The unnecessary and arbitrary novel food regulation constitutes, in effect, a punitive "tax" on regulated products or activities, which, in turn, creates a potent disincentive to product development and use. Finally, the EU's regulation and its tax are incompatible with the U.S. policies and may well precipitate trade conflicts or even a trade war — corollaries of the law of unintended consequences.

Japan

The Japanese government has made no pretense of adopting policies that are consistent with the scientific consensus that the new biotechnology is an extension, or refinement, of older genetic techniques, or with the spirit of the OECD's "substantial equivalence." Rather, the Japanese Ministry of Health and Welfare (MHW) has imposed a strict regulatory regime specific to foods and food additives manufactured with rDNA techniques (16). This regime captures virtually all rDNA-derived products for case by case review and subjects them to extraordinarily stringent standards for manufacture, documentation, record keeping, characterization of the source organism and the actual food products, and so on — a far higher standard than any other class of food, except perhaps for the preparation of *Fugu*, a fish much favored in Japan that contains a potent and potentially fatal neurotoxin. Because food products' profit margins are low, discriminatory and unnecessary regulation is a strong disincentive to using a new technology. Predictably, regulatory disincentives have prevented Japan from exploiting its experience and traditional strengths in agriculture and the production of fermented foods and beverages.

In 1986 MHW promulgated guidelines concerning the manufacture of new-biotechnology products, loosely based on the OECD 1986 report, "Recombinant DNA Safety Considerations" (18), but failed to incorporate the spirit of "substantial equivalence." MHW also has regulatory responsibility for food additives, and in 1991 issued a policy statement, "Basic Principles on Safety Assurance for Foods and Food Additives Produced by Biotechnology" (17). A year later the MHW issued two guidelines for the new biotechnology used in food production: a Manufacturing Guideline (GMP, or Good Manufacturing Practice) and a Safety Assessment Guideline. In December 1999 MHW announced that beginning in April 2001 mandatory tests on the potential health risks of genetically modified (GM) foods would replace voluntary testing, and also that products approved and considered "safe" would be identified by labels as having been produced by the new biotechnology (18). However, as of March 2000, there remained uncertainty about how this would be accomplished and whether the labels would, indeed, indicate "safety." The confusion lies in the existence of conflicting (August 1999) draft regulations of the Ministry of Agriculture, Forestry, and Fisheries, which require 30 food ingredients containing gene-spliced ingredients to be labeled as such, and that products with a mixture of gene-spliced and non-gene-spliced ingredients be labeled as "undifferentiated" (18).

In theory, the Japanese adhere to the concept of "substantial equivalence" as articulated by the OECD (*vide supra*), but in practice their primary regulatory trigger is process based — that is, the use of recombinant DNA techniques — and there are unprecedented and irrational admonitions that "recombinants themselves are not to be consumed" (19).

BIBLIOGRAPHY

1. R.M. Goodman et al., *Science* **236**, 48–54 (1987).
2. U.S. National Research Council, *Field Testing Genetically Modified Organisms: Framework for Decisions*, National Academy Press, Washington, DC, 1989.
3. H.I. Miller, *San Francisco Chronicle*, May 31, A17 (1994).
4. Fed. Regist. **57**, 22984–23005 (May 29, 1992).
5. H.I. Miller, *J. Am. Med. Assoc.* **269**, 910–914 (1993).
6. J.A. Nordler et al., *N. Engl. J. Med.* **334**, 688–699 (1996).
7. V.J. Taylor, Memorandum to the Labeling Subcommittee, California Interagency Biotechnology Task Force, April 15, 1994.
8. H.I. Miller, *Policy Controversy in Biotechnology: An Insider's View*, R.G. Landes, Austin, TX, 1997, pp. 43–51.
9. Statement of Barbara Keating-Edh, President of Consumer Alert, before the National Biotechnology Policy Board, September 20, 1991.
10. Anon., *Food Chemical News*, January 31, p. 59 (1994).

11. R. Arbaugh, FDA Office of Regulatory Affairs, Memorandum to State Feed Control Officials, July 1, 1996.
12. Organization for Economic Cooperation and Development, *Safety Evaluation of Foods Derived by Modern Biotechnology*, OECD, Paris, 1993.
13. D. Butler, *Nature* **384**, 502–503 (1996).
14. H.I. Miller, *Financial Times (London)*, April 22 (1997).
15. H.I. Miller, *Financial Times (London)*, November 29 (1996).
16. Japanese Ministry of Health and Welfare, *Guidelines for Foods and Food Additives Produced by Recombinant DNA Techniques*, Tokyo, Japan, 1992.
17. Organization for Economic Cooperation and Development, *Recombinant DNA Safety Considerations*, OECD, Paris, 1986.
18. Y. Uozum, *Nature* **402**, 846 (1999).
19. Y. Suzuki, in *Proc. Japan–USA Workshop on Scientific Political and Social Aspects of Recombinant DNA*, Japanese Science and Technology Agency, Tokyo.

See other entries AGRICULTURAL BIOTECHNOLOGY; ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE.

AGRICULTURAL BIOTECHNOLOGY, LAW, AND SOCIAL IMPACTS OF AGRICULTURAL BIOTECHNOLOGY

WILLIAM P. BROWNE
Central Michigan University
Mount Pleasant, Michigan

OUTLINE

INTRODUCTION

Biotechnology in agriculture has gained as much attention for its social and political controversies as it has for its science and its promise for food and fiber production. Those controversies as they pertain to different types of scientific products and processes are the main subject of this article. One polar view holds that agricultural biotechnology is the primary key to providing for future world food needs. The opposite view is that biotechnology products subject agriculture, people, and food availability all to potentially grave risks — too much so as some see them. These opposing positions as well as myriad centrist views have led to legal, regulatory, and other political activity and conflicts that are described here. In the process of engendering conflict, they also have led to a complex and growing maze of statutes, regulations, and international agreements that govern agricultural biotechnology to the satisfaction of almost no one. Those institutional rules and dissatisfactions are also covered here.

LEGAL FRAMEWORK

Two types of genetically engineered products have received the greatest attention in agriculture. The first is recombinant products such as hormones that are added to living organisms through application or injection. The second is transgenic manipulation, in which genes of one organism are engineered into another, often of a very different species. Since the two types of products are quite distinct in technology and in approach, they have evoked unique interests within the governing process in recent years. Considerable irony exists in these unique interests because, at least in the United States, the essence of legislation actually proceeded the emergence of agricultural biotechnology products. Precedents were set long before the new technology arrived.

U.S. Domestic Legislation: Protecting Property

The generic legal framework for biotechnology plant products is the Plant Patent Act (PPA) of 1930 (1). It was essentially a matter of protecting profits for those building an industry on new products. So was subsequent legislation. The intention was to provide protection of intellectual property developed primarily by agribusiness companies (2). The Act allowed patent-style protection for new plant innovations, which sets those important legal precedents. PPA, however, had limited utility since it only covered asexually reproducing plants, which included grafting but not new seeds. That obviously failed to cover much of the U.S. plant industry, as became startlingly clear in the 1940s with the introduction of corn hybridization.

As a corrective, the Plant Variety Protection Act (PVPA) of 1970 extended important features of the 1930 legislation. Specifically, it gave intellectual property protection for 20 years to sexually reproducing plants that hold over time their uniform characteristics. Far more products were able to get certificates of protection. PVPA, however, had significant exemptions, for research efforts to develop new products and for crop growing purposes, or farmer-to-farmer sales. These exemptions meant that plant industries still had relatively limited product control as well as incentives to litigate. In 1994 these two problems with PVPA were marginally narrowed by Congress, but not anywhere to the satisfaction of industry.

Much of the logic for plant protection was touched on but not actually extended to livestock and other animals

in the Animal Patent Act of 1986 (3). It, like a subsequent act in 1988, failed. Nonetheless, arguments made in the congressional hearings combined with federal genetics privacy legislation which protects DNA, provided the eventual basis for certifying animal innovations as well as plant species (4). But federal administrators rather than Congress were to become the actual arbitrators.

Evolution: Litigation and Regulation in the United States

The congressional emphasis on property protection rather than the direct topic of biotechnology products is not surprising. Historically U.S. federal legislation develops only in the vaguest sense and with minimal explicitness. Then bureaucracy takes over. Many issues simply get avoided for political reasons, although administrators are often forced to eventually rule as part of their legal duties. Either that or the courts rule. Those reasons lie in part with the complex and shared governing powers within the policy-making process. Powers are both divided and separated as well as shared from one institution to the next. This brings very different and more compromised results than the far more explicit and integrated policy processes that characterize parliamentary forms of national government, which dominant in most other democratic regimes. With more executive control in keeping agencies accountable, governing in parliamentary nations is far less fragmented.

Within the United States, Congress disposes of policy initiatives half-heartedly. The presence of both active federal courts and of technical administrative agencies explains why. With their involvement Congress understands full well that most of its policy decisions will be refined further through judicial litigation and through agency regulation when enabling authority exists. Thus Congress has rarely tried to resolve all relevant issues in statutory law. That, of course, is why congressional initiatives on protecting plant and animal properties were necessarily to be evolving rather than dealing with all possible new products and resulting legal concerns (5).

It was clear from legislative hearings on all of the major acts that property protection was not the only issue up for contest. Most of those who testified as critics in the hearings addressed side issues such as animal well-being, environmental impact, human health as linked to new products, and such economic and social consequences of biotechnology as resulting financial concentration in farming (6). Some of those concerns led to the aforementioned plant variety exemptions for research and for growers.

But for the most part these things were left to be at least partially resolved in other, mostly nonpatent federal institutions. So they evolved in those places to set legal case law. Federal courts were involved at first, and again most explicitly in interpreting property protection since that had been the emphasis of legislation.

While litigation has been prolific, only three cases merit specific comment for their precedent setting nature for biotechnology. A California state case, *Moore v. Regents of the University of California*, ruled that biotechnology industries can utilize genetic raw materials freely and to their own purposes, regardless of donor circumstances. Even more landmark, however, was the earlier U.S. Supreme Court ruling in *Diamond v. Chakabarty*. The Chakabarty decision affirmed that traditional patent law is indeed still patent law, that there are no differences in law between the animate and inanimate innovation. The legal distinction is rather between what nature produces and what humans engineer. This case not only supported the biotechnology industry, it also dealt with living bacterium and thus extended plant protections to animals. The federal patent office moved under this authority in 1988 to certify the famous Harvard mouse. *Ex parte Hibberb* then moved to clarify which products could receive utility patents. The courts ruled that when Congress did not explicitly exclude plants from such patents, it left intact all forms of procedural plant protection. Armed with that interpretation, and confident that it is also extended to animals, industry firms moved successfully to prosecute numerous patent violations as well as to patent more and more products (7). There, however, still remained a downside for industry: Patent protection for plants and animals is hard to obtain because of the nature of both these products and the utility patent process. To gain protection, inventors must prove novel and useful application, describe the invention well enough to let others with necessary skills recreate it, and demonstrate true innovation as opposed to a simply logical extension of past ideas. Protecting trade secrets is allowed within that legal context.

The regulatory arena has been no less active than the courts in promulgating legal precedents that shape the law for biotechnology, indeed far more so. Using and interpreting patenting and other regulatory processes shaped by Congress through enabling legislation, administrators have become prominent and consequential public policymakers. Both the U.S. Patent and Trademark Office (PTO) within the Department of Commerce and the U.S. Department of Agriculture (USDA) share specific patent responsibilities which lead to frequent interpretive differences. USDA in the 1990s had very explicit opposition, for example, to patenting an entire species, which its officials argue cannot represent a new discovery. PTO disagrees, leaving resolution to continue into the future.

Points of administrative confusion and conflict are far more than just over patent interpretations since myriad other nonpatent agencies have agricultural biotechnology jurisdictions or at least interests. This opens a whole Pandora's Box of American public policymaking for product review. Most notable of the agencies are the federal Food and Drug Administration (FDA) and the Environmental Protection Agency (EPA). USDA's traditional agency in this area of food safety also is important but inclined to be less critical of biotechnology. The Animal, Plant, Health Inspection Service (APHIS) nonetheless adds to the regulatory muddle through PPA. Under the long existing Public Health Service Act and other lesser statutes, FDA has clear authority to regulate and ensure the safety of foods derived from new plant varieties and, therefore, new techniques of production. This entails what can be an extensive product safety review. EPA has contrasting and less clear authority to mandate and review environmental

impact statements. It has been at this juncture where transgenic manipulation has been separated in impact from recombinant technologies. EPA has been concerned that new products with variable gene structures may contaminate existing and especially wild or natural settings. This could threaten existing strains of plants and animals ranging from commercial canola to migratory salmon. To respond to this problem, EPA has moved to regulate and test all new agricultural varieties which promise plant protection under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA).

EPA's response infuriated most of the active U.S. scientific community, including eliciting a direct and angry critique from 11 scientific societies (8). This was, they argued, neither EPA's job nor was it a scientifically sound conclusion (9). The case nonetheless shows a great deal about the effect American politics has on a fragmented and barely integrated regulatory process that opens its official doors to nearly any and all interest groups and citizen complaints. If complainants fail to be content with decisions in one place, they simply look for other agencies with which to complain. When APHIS says no to a complaint, for example, the complaintant goes on to EPA — and looks and hopes for a different response on different grounds. In effect, the United States regulates agricultural biotechnology to ease public fears more than to bring forward sound and lasting scientific decisions. Dissenting social forces move, for example, from PTO to USDA patent officials to FDA to EPA to APHIS. Animal Patent Law hearings, as noted, were places where many nongermane issues were raised.

A good example of cultivating the public can be seen in another task of USDA's Office of Agricultural Biotechnology (10). Since 1987 this office has taken a neutral stance on product innovations and on biotechnology itself. Yet it further opens up decision-making units within government to intense lobbying and media coverage from any dissenting social faction. Federalism, or the separation of American governments into three levels of government, does even more of the same. When no federal agencies respond, agricultural biotechnology critics become politically active within the American states and even localities. Chicago provides a superb illustration with its decision to insist on consumer labeling for any food product that may be adulterated by agricultural biotechnology, either through recombinant or transgenic processes. Quite clearly, those who wish to promote this biotechnology industry in the United States face a costly, cumbersome, and uncertain structure of governance. Yet biotechnology critics and opponents love that structure for the many obstacles it provides.

Agricultural Biotechnology in Other Countries

As the above makes abundantly clear, the United States can hardly have a systematic and supportive public policy to promote agricultural biotechnology. Things within governing structures are just too cumbersome for that to be the case. Yet, on a comparative basis, U.S. governments have been generally friendlier to agricultural biotechnology than have those of almost all foreign governments (11). There are some exceptions. Germany, Australia, and Japan actively support the industry, because in large part both agriculture and high levels of industrialization matter to their political economies. That support, however, has not quelled loud and extensive public criticism in those countries. China with its major worries over food supplies and Brazil with its aspirations as a leading world food producer tend also to be supportive yet not particularly activist in their biotechnology political agendas.

Much of Europe as well as countries of the Southern Hemisphere are far more skeptical. France as a recent convert to conservation policy links agricultural biotechnology to those concerns, particularly in protecting indigenous germ plasm. Industry is also not advantaged in the United Kingdom, where both government and public skeptics about technology's effects on a prominent farming sector prevail. With such splits as between France and Germany, and with family farming revered throughout most of Europe, no common laws on anything of substance have passed. The European Parliament has, despite the splits, passed a nonbinding resolution against animal patenting, seen the introduction of several resolutions attacking plant patents, failed to come up with a proposal for a European initiative supporting and regulating agricultural biotechnology, and condemned the international European Patent Office for actually approving an animal patent as it did in 1992 (12). Europe, as a consequence, lags several years behind the United States in actual public policy supports for the industry. Moreover its federated European Union (EU) governing structure is proving to be even as cumbersome and fragmented as the United States on biotechnology matters in agriculture, only just not *within* the individual countries.

Despite serious food and trade needs as well as rapidly growing human populations, countries of the Southern Hemisphere also act skeptically. Culture plays a major role in structuring negative responses as does the importance of farming in each nation's politics. The greatest reluctance, however, comes from fears of further domination by the United States and other highly developed industrial nations. Exploitation has long been a dramatic political issue in Asia, Africa, and Latin America. Its importance has only escalated as industry firms have moved into underdeveloped countries in major initiatives to research as well as to market biotechnology products. Pioneer Hi-Bred International, with its offices in over 30 countries, has become for many an archetype of the fearsome multinational corporation. Thus, while governments around the world could quite easily promote agricultural biotechnology because of their more integrated parliamentary structures, these countries have demonstrated the ease with which they can also react reluctantly.

International Cooperation

Agriculture has long been a policy area marked by substantial efforts at international cooperation. In part, this is because agriculture and now agricultural biotechnology are subject to agreements between so many participant nations within several international agencies. From world food needs to nutrition problems of the poor, the United

Nations has played an activist role in fostering various agreements. That has worked in part because there has been much to agree on in promoting solutions to world farm, food, and fiber problems. Humanitarianism is a common goal, one easily favored. Agricultural biotechnology, however, brings no such consensus, as those nation-by-nation differences discussed above would obviously suggest. Disagreements abound and limit any efforts to think that global cooperation on this subject will ever take on either a voluntary or a legal status. There was passed in 1992, though, a UN Convention on Biological Diversity. While the United States did not ratify the convention, it still participates actively in its frequent and ongoing conferences. The biggest controversy surrounding the convention has been over the issue of a legally binding protocol governing the release of genetically modified organisms in developing countries. While an earlier UN report by 15 experts saw no need for such a protocol, it still passed but has not yet been adopted. Opposition of major agricultural biotechnology producer countries favored a voluntary protocol and numerous future meetings were planned and held to facilitate discussions (13). No substance has yet emerged. That both irritates friends and foes of the binding protocol and reveals again the lack of authority to be found in international cooperation.

The UN, however, is not the sole international agency to be involved with or split by controversy. Global standards on health and environmental protection are developed by the World Trade Organization (WTO). On those grounds WTO voted to delay action on rBGH, which meant that the EU's ban on the product remained in place. The World Health Organization (WHO) advocated strategies for assessing foods produced by biotechnology. And the Organization for Economic Cooperation and Development (OECD) has directed its mostly European member countries to carefully scrutinize agricultural biotechnology products. As a member of all three organizations, the United States stood as a biotechnology proponent and faced formidable opposition from many developing nations and several European countries. Once again, none of these three organizations have actually changed global production or trade conditions. They have, however, increased negative attention to biotechnology as applied to food products.

Much of that negative attention goes back to the UN as well as its relatively independent suborganizations. Two strong suborganizations have along with central UN deliberations had the strongest impact. The UN Food and Agriculture Organization has both recognized the promise of agricultural biotechnology and issued conditions under which its application should continue. FAO wants that biotechnology to be used only for highly planned special circumstances and specified problems, to be adapted to local infrastructure needs, to wrestle with the complexity and equality of property rights issues, and to ensure food and environmental safety along with fostering biodiversity. FAO clearly holds that what it calls "novel foods" should be subject to use with extreme caution. Such concerns impose further checks on the actions of agricultural biotechnology producer countries, further limiting and restricting their market inroads.

The same is true of the UN Environmental Programme (UNEP). Along with FAO, UNEP produces extensive research and information. But unlike FAO, UNEP takes a more activist and indeed interventionist position as well, especially through its Environmental Law Programme (ELP). The ELP promotes the development of international legal instruments, develops international environmental law aimed at sustainable development, provides technical and legal assistance to countries with developing and transitional economies, and exercises leadership in implementing environmental law both internationally and nationally. It has become clear from recent actions of the 1990s that ELP/UNEP has concerned itself extensively and provocatively with food biotechnology products and practices. Its partner in all such instances has been OECD. That partnership, like OECD more generally, has focused more negative than positive attention on such agricultural issues. Thus the balanced result of voluntary member organizations has proved more negative than positive for the whole of worldwide agricultural biotechnology.

SOCIAL PROMISE OF AGRICULTURAL BIOTECHNOLOGY

There is no mystery why agricultural biotechnology has been subject to so much social and political concern. Two related reasons explain it. World politics has long been influenced by Malthusian fears. Economist Thomas Malthus determined in the early 1830s that food production was increasing at a far slower rate than was the rate for the human population. Yet industrialization of agriculture occurred in the twentieth century, and the food production rate of growth outpaced the population rate of growth by 3 to 2 between 1950 and century's end. Still, however, fears of food shortages remain real, for a very good reason. The world's population of 6 billion people is expected to peak at nearly 11 billion by 2050 (14).

These contrasting observations have affected politics and social values at two levels. First, Malthusian fears have led at least indirectly to massive government intervention in agriculture over time. Nations have supported agricultural expansion, policies to bring about a more educated and productive farm population, and research that has hastened farm industrialization and modernization (15). All that has been especially true of the United States. But the U.S. model of agricultural education, research, and outreach or extension assistance to farmers has been exported to other nations worldwide. Thus proactive and interventionist government in agriculture has prevailed. No country trusts agriculture only to the private market. The contrasting belief that the world will not experience long-term food shortages has produced a rather different political scenario in more recent years. While food availability fears still exist, governments have become more skeptical of their own agricultural expenditures and involvement. A view exists that Malthusians have always been wrong and will continue to be wrong in the future. As a consequence fear no longer drives, though it does still in tandem influence, agricultural policy initiatives. The private sector, or market, also gets much of the credit so far for having avoided food shortages. Some trust in business

has then emerged from some parts of governments. This enhances the influence of agribusinesses. However, since much of the social base of fears of a food-short world are gone, more stern critics of agriculture also have emerged. These new critics are much less inclined than previous generations of critics to support commercial agriculture and research at any social and economic cost. Indeed, the new critics want instead to closely scrutinize any costs of a further developing agriculture, especially one led by market considerations which may not take into account human safety, cultural values, or essentials of nutrition.

Those contrasting views of advocates and skeptics both get played out daily in agricultural policy making and in the media. Nowhere is that more true than for transgenic agricultural biotechnology. Without the fear factor of Malthus, agricultural biotechnology has been the subject of intense calls for regulation, as can be surmised from the previous section. Government therefore intervenes in what is often less than a supportive fashion for industry. At the same time, few interests want agricultural biotechnology to truly fail — but rather just to be safe (16). Exceptions do exist, mostly from those widely labeled to be modern Luddites. The general view of both industry and most critics of agricultural biotechnology, though, is much the same: Support for or opposition to the social promises and fearful circumstances that emerging technologies seem possibly to produce. Agricultural technology then is the subject of both severe complaints as well as valued for what it may offer in the near- and long-time future. This makes for conditions that fail to bring about extensive and comprehensive public policy. No politicians wish to decide.

That social promise, as well as the diverse fears that come with it, need to be specifically examined in order to make sense of this rather dizzying politics, for the promises are immense and of far more than marginal social value. So too are the fears. Promised contributions of agricultural biotechnology include the obvious one of finalizing the closing of the gap between food availability and food needs. What began with commercial industrialization of agriculture and moved from a green revolution of hybridization into molecular biology now can be turned toward engineering plants and animals whose genetic characteristics make for more food. By identifying those genes, by marking them, and by transferring them to host organisms, better products can result. "Better," however, does not just mean more world food to those doing agricultural research. "Better" also means foods that are, for instance, more nutritious, build in pesticide control genes, and foster a resulting ecological improvement from reduced agricultural chemical use. That value-laden term of the advocates of "better" also implies social gains as new biotechnology products use nutrients more efficiently and so lead to higher yields that also promise to limit destruction of old growth wildernesses for farming and ranching (5). Even with the increased financial costs of the technologies, they might bring economies of production scale as well and therefore at least some lower food costs. "Better" also means food products that are less prone to spoilage, have longer shelf lives, and are less subject to bacterial contamination. The social concern over wasted foodstuffs may thus be addressed. Perhaps "better" also means foods that have more appealing taste and can find usable markets where only limited ones existed before (17). That has compounding importance when these more tasty and now marketable products can substitute for foodstuffs grown in short supply on environmentally fragile lands or by using high levels of degrading chemical inputs.

The "better" arguments are countered by fears of worsening food conditions. To opponents, "worse" also means many things. The range of possible difficulties go from the creation of pesticide resistant bacteria and other pests to the loss of genetically proven and strong seed and animal strains, to the contervailing view that food will be less nuitrious over time. A major disagreement over "worse" owes to what will be the nature of a specific food product. What are the religious implications of a Muslim or Jew eating plants expressing pig genes? What is the likelihood of a fish allergic consumer being ravaged by eating a tomato expressing trout genes? Both the public and the media pay widespread attention to such fears for quite obvious reasons.

With these conditions in mind, U.S. government regulators have approved the following sorts of products, which in total still number less than two dozen by FDA. A few more are approved by other regulatory agencies depending on their jurisdiction. There are far fewer approvals worldwide. Recombinant bovine somatotrophin (rBST) came first. It falls into the category of more product, or milk, with more efficient nutrient use by cows. Adoption in the United States has become relatively high but plagued with significant milk producer expense. European countries still await approval, which is hardly surprising since they have — for trade reasons — long banned imported meat raised with growth hormones. Beyond rBST, there are varieties of marketed biotechnology substances. Calgene brings a transgenically altered and more lasting and transportable tomato. Large firms, mostly Monsanto, have introduced products to deal with agricultural chemicals and use, both for plant survival and to reduce applications. There are on the market soybeans and canola that can withstand herbicide application, canola that produces improved oils, and corn that resists some insects. Despite these innovative products, however, there is little to conclude from them about the future successes of agricultural biotechnology. These are simply too few of these innovations yet on the market to comment on and assume success of the technology in attaining its social promises. Yet from the above illustration two fears about "worse" have come to life: the example of a tomato expressing trout genes and the death of monarch butterflies eating too near pesticide resistant crops.

All that anticipated value then is still but the visionaries' promise of agricultural biotechnology: more food, less future starvation, a more secure environment, and more choices of affordable products. Small wonder that these new food technologies have been widely championed, especially with some fears coming real. The scientific and industrial advocacy, however, is enhanced further by the accompanying advent of new mechanisms for business and profit. There exists substantial room for firms to make money by developing new products

with all those desirable characteristics (18). Some are large firms such as Monsanto and Pioneer that have set up extraordinarily large scientific enterprises to foster the new technologies. Monsanto, for example, has invested more than four billion dollars in agricultural biotechnology. Others are far smaller firms set up by entrepreneurs who organize around a single product, often with state government backing to enhance area economic development. Both types have in general found appeal among finance capitalists and stock market investors, which further drives advocacy of the technology. Industry and scientific jobs and technically sound advancement are, of course, powerful social motivators for investors and for governments. Combined with all the visionary promises, profits and their additional potential add real social luster to agricultural biotechnology and lend it greater social legitimacy. The technology is therefore a dramatic social and political force. Yet that economic promise is still to be proved. Monsanto supports its agricultural biotechnology science by selling more herbicides, many corporate stock prices have experienced declines, some firms have failed, and some states have divested their investments as far too costly to state budgets.

BUT THERE ARE THOSE LOOMING CRITICS

These broad-ranging promises, however, are only predictions premised on the good guesses of a very able but always limited scientific community. So too are social fears. No one can be exact as to what will actually occur, either as a benefit or as a consequence of agricultural biotechnology. That uncertainty produces, first of all, considerable perceived risk and, second, opens up political and economic doors to every advocate who holds a contrary position. To the great frustration of public policy makers, no one can give either outright assurances or absolute reasons for resistance to biotechnology. Thus, for scientists and industry, there are a great many opposing and competing views from formidable critics that especially plague them. Some common and recurring critical assessments that filter down to and move the national publics at large need highlighting.

These can be typed and analyzed in five general but not mutually exclusive categories: the philosophy-and-ethics-of-risk problem, the agricultural-sustainability problem, the farm problem, the Jeremy Rifkin problem, and the opposing-coalition problem. Each will be explained in full, not because of the importance of any one of them. Rather the explanations are sequenced according to the formal logic and indeed science of their advocates' opposition. The most scientifically logical problems come earliest, Then, in descending order, the problems become even more purely political and driven by the competing social values of and myths held out by the opponents. To an important extent, however, a degree of both logic and politics can be found within each type of critical disagreement because there are so few hard, or perfect, pieces of evidence.

The Philosophy-and-Ethics-of-Risk Problem

At the center of the debate for those critics are the growing relationships between scientists, the universities that employ them, and the industries that offer grants and contracts to university budgets and researchers (19). The role that government plays in encouraging applied research, cutting university budgets, and supporting economic growth and development also enters in according to these critics. The basic skeptism results from beliefs that as these forces became more interdependent, scientific inquiry lost its former neutral ethics of research. Ethical values, for reasons of scientific self-preservation, had to give way to science's move to the center of social and economic service (20). As a consequence scientists lost much of their traditional culture of critical and objective inquiry. This, of course, is a harsh view.

The harshest charge of those who see this problem goes beyond that perversion-of-science charge. It emphasizes that in pursuing social and economic relationships, scientists fail to adequately consider even the possibilities of the negative consequences of their actions. What they learn might hurt. In drawing an analogy that simplifies the problem, one critic raised the specter of Dr. Frankenstein. Frankenstein was not scientifically wrong for what he did in creating the monster. He was wrong in not anticipating what might result (21). Thus he failed to take any action to prevent disasters that may occur. Scientists at the center of society are thus myopic, inclined not to analyze everything that should be subject to inquiry. Finding likely consequences may be a threat. At least some of that myopia is tied to budget constraints but even more is linked to doing whatever best serves the new scientific culture of interdependent rather than independent relationships with those who fund research.

Beyond this generic concern there are few agreements as to what are the specific risks of biotechnology and what should be done about them. An extreme position is that science should ensure that any genetic creations that escape to the environment should quickly die and never become ecological contaminants (22). Doing so certainly makes avoiding risk the highest priority of agricultural biotechnology. Many such proponents favor no-risk tolerance laws. These critics look to China as the object lesson for this concern over preserving natural or indigenous germ plasm, species, and plant varieties. Due to massive manipulation of plants and animals, China has not seen the semblance of a normal environment for over 1000 years. Major loss of varieties has occurred.

Others, however, have less comprehensive and restrictive concerns. Some will even accept low levels of risk. Biotechnology, according to others, should not be injurious to animals as a basic ethical standard (23). An ironic expression of risk is the likelihood of producing too much food and destroying in the process social and institutional stability in the world order (24). On an entirely different level, communication specialists argue for getting citizens involved in assessing risks (25). With increased public involvement of citizens, risk-minimizing benefits of two kinds accrue. First, the likelihood of public opinion having too much faith in that culturally changing science is minimized. Increased social scrutiny results. A competing concern comes from the came critics. They note that despite perhaps misplaced faith in science, the public also has simultaneously developed a highly critical view of government and other social institutions. Should those views

be moved further by food system or environmental failure laid to agricultural biotechnology, public opinion would become irrationally cynical of these new technologies and lead to hostile and even backward controls.

Thus, with a philosophical framework intact of minimizing risk, ethics and developing standards still mean several different things to several different critics. That does not imply, however, that these critics have not made their mark. The EPA takes a strongly risk-averse approach, even in confronting organized scientists. USDA's Office of Agricultural Biotechnology publicly and politically holds a neutral position to ensure confidence and avoid charges of product advocacy over that of the public interest. Finally, USDA's APHIS has also taken an innovative stand by establishing standards for releasing information and encouraging a dialogue with various segments of society (10). Avoiding risks of various kinds has certainly penetrated the public policy process and the attendant debate over regulatory standards. That, of course, complicates U.S. national politics, especially in bringing hypothetical fears to a cautious public that might well revolt. European politics is hardly immune to the same problems.

Agricultural-Sustainability Problem

While discussions of agricultural biotechnology have been a mix of effusiveness, critical thought, and social attention, other innovative approaches to the Malthusian fear of world food shortage have been in evidence. A prominent one is sustainable agriculture, the search for methods of production and food products that will not lead either gradually or even in a crisis to production disasters. Lester Brown's Worldwatch argues first and foremost for long-term sustainability over biotechnical advancement as do Rodale industries, both two organizations that reach countless Americans and a great many residents of other countries worldwide. The sustainability debate in agriculture is often linked to better and more productive agricultural biotechnology (26). This linkage is but more of the idea that biotechnology offers food and production needs almost limitless social value. The logic, given earlier, is obvious but unproved.

Not all observers agree with the positive linkage, though, including many in USDA's original office of sustainable agriculture. The reasons for the disagreement are complex and not as intuitively obvious as ones that in a positive way link biotechnology and sustainability (27). Part of the reason for obfuscation of problems lies in the origins of sustainable agriculture. Although not all of its advocates agree, sustainability has its origins in organic farming. As such, much of both its rhetoric and science are anti-chemical input with a championing of healthy ecological conditions. As sustainability developed its own goals, though, it moved its position from no chemicals to reduced chemicals. Sustainability advocates, accordingly, mostly want food that is healthier to consume and healthier for long-term soil and water resources. Thus not all these advocates are extremists in their beliefs. They also want food that is less expensive for producers to grow as the amount of expensive agrochemical inputs are decreased. The latter point has gained these advocates a growing number of supporters among farmers and within several federal agencies in the United States. Research support is established and considerable. Sustainability has become popular enough that most USDA agencies try to promote it at least at the margins. The Extension Service promotes it more aggressively through its farmer assistance outreach in some regions of the country. The U.S. South is especially sustainability conscious. Sustainable advocates in general are also more impressed with precision farming innovations rather than biotechnology, which holds up another roadblock for many.

Opposition from some sustainability segments nonetheless charges agricultural biotechnology interests as bringing forth several production negatives. They tend to argue the following: that biotechnology is owned and controlled by large industries that have records of environmental abuse rather than of contribution, that biotechnology in agriculture can have no life without this corporate control, that diversity in agriculture will further give way worldwide to specialized and large-scale crop and animal production, that all of this takes decision making away from local farmers who now often creatively address their own sustainability and environmental needs, and finally, that the sum total effect is to distance people and producers from what is grown and how it is grown.

Stewardship over natural resources, as a consequence, is likely to disappear as a human value. These are another set of very harsh and especially cynical views. The cynicism and the targets of attack make them less potent than are the arguments of advocates against risk. Despite what seems the marginal status of the most extreme sustainability critics, these views have enough attention in public, farm, and policy-making circles worldwide to be a considerable factor in evaluating policy options. The 1999 move of the Henry Wallace Institute for Alternative Agriculture, with its sustainability emphasis, to the worldly acclaimed Winrock International has enhanced that research identification. The anti-business concerns are nonetheless particularly persuasive in bringing opposition in developing nations to the policy wants of heavily industrialized, biotechnology-producing agricultural nations.

The Farm Problem

It seems unlikely at first thought that farmers, producers, and growers would present agricultural biotechnology with an obstacle. The farm problem, however, exists as a major one for industry and science. Production agriculture, even through the present, is treated both in the United States and in much of the world as a unique economic sector deserving special government assistance (28). Even as farming loses its economic uniqueness, a raft of institutional structures can still be found that protect producers against market failure, natural disaster, and even economic loss. The reason owes to another irony of production agriculture: Just after Malthus was spreading his fears, the United States entered a period through today and into the foreseeable future that has brought food supplies that are too large, farm prices that are too low, and accompanying failures and farm losses among producers. Technological innovations kept that spiral in place (29).

In a nation that was settling its frontiers and attracting massive development, failures of this magnitude among so many national agricultural institutions were found to be politically unacceptable.

First, considerable social and political investment had been made in increasing the number of producers, getting them to frontiers, keeping them productive, and even giving them land. Nations as diverse as Japan and Mexico passed legal structures either to keep farm prices artificially high and away from a worldwide market or simply to give expropriated lands to peasants. The belief was and still remains the simple one that producers as providers of food and stewards of both natural resources and national security deserve special treatment. As Jeffersonian Democracy spread in the United States, a myth emerged that family farmers were important and should be preserved for reasons of protecting basic social values (30). This played out as the widespread social belief that family farming structures of production deserved protecting against large industrializing farm and ranch holdings (31). Other nations were not immune to this notion of farming as a basic social value, and protectionism became a common worldwide practice.

Farm failures were unacceptable from a second perspective as well. Not only was national development important and existing on a powerful base of social values and myths, farmers were also acquiring vast political importance in agrarian and in evolving nations. Farm power was institutionalized in the United States in its own structures within Congress and in administrative agencies. Farmers felt entitled to assistance, and government officials treated them as if they indeed were. Homestead laws were passed, railroad transportation was regulated, a huge agricultural establishment of research and education was put in place, marketing supports were advanced, and, as farm failures continued, direct government payments to farmers were made. Farmers and organized farm interests zealously protected all this. Congress proved the best vehicle for protection because congressional members from rural districts and states were anxious to serve local constituents, especially where they could effectively destabilize regional politics (32). Similar ventures worked worldwide to play off and protect myths of agrarians as special and often unassailable interests.

Thus, when agribusiness industry took on economic and political significance, farmers won concessions frequently and with a political flair. In the United States, patent laws allowed farmers to sell certified products to others as these were left over from individual production efforts. Until patent law cases were more effective, farmer actions were often abusive. Yet only in 1994 did the courts limit farmers against industry. In *Asgrow v. Winterboer*, the U.S. Supreme Court ruled that it was quite excessive for a single family farm to stockpile and sell enough soybean seed to plant 400,000 acres. As long as farmer interests work to compromise with those other social and economic interests in their nation, however, farmers still hold potent influence. Environmental and farm trades among otherwise competing interests are a good example, one directed harshly and in tandem against biotechnology industries.

The emergence of opposition to agricultural biotechnology shows that influence with clarity, at least relative to countering industry and science. Dairy farmers delayed the opposition of rBST for years. They argued, through to the present, that the cost of the technology was too high for smaller-scale farmers to bear. Economic benefits could go primarily to larger-sized producers who enjoyed economies of scale in distributing costs throughout large herds. This smaller-scale producer fear argued that rBST would further increase escalating farm size and disrupt the widely valued structure of family farm production. Several other biotechnology products have been opposed for similar reasons of costs and farm structure. These controversial products range from strawberries that can withstand lower temperatures and not freeze as well as transgenic innovations that at considerable expense added preservation qualities and had longer shelf lives. Opposition also was strong against adding transgenic innovations in farm animals. In all these cases farmers wanted protection from changes and competitive circumstances in their own operations. They wanted to preserve their own current animal breeding plans, keep personal costs low, resist greater agribusiness dependency, and continue to produce with well-understood and familiar practices. As a result much of the farm problem over biotechnology has been directed against agribusiness corporations and production structures that are far from the traditions of family farming. This same farm problem perspective has also been directed against other than biotechnology producers, for instance, against those proposing irradiation of food for longer shelf life. This rather populist rhetoric finds considerable public support as well, and it has gained a number of limited-technology champions within Congress and numerous state legislatures.

Agricultural biotechnology thus comes constantly against prevailing social sentiments that are protective of family farming. The tendency is for liberal, progressive, and populist elements in agriculture to generate much of this daily opposition. A long and enduring tradition of ideological farm protest has always split farm sectors in many nations, including the United States (33). The farm problem, however, is not restricted only to populist opposition. Many mainstream general and commodity farm interest groups have cooperated in challenging some biotechnology products and attacking agribusiness projects. What to support and what to oppose is in many of these cases a matter of primarily what farmers like and what they do not. Farmers, for obvious reasons, like and support new pesticide resistant crops. Yet they dislike many other products because these disrupt existing and widely accepted production and even marketing practices. That opposition results from all types of producers, regardless of political ideology and farm size. It gets played out in all of the institutions of government that have long provided farm services. This indicates that the farm problem will continue to plague at least some biotechnology innovations, as well as others, into the more than immediate future.

The Jeremy Rifkin Problem

It may seem unfair to personalize a critical problem with the name of a single human being. Yet Jeremy Rifkin and his unrelenting and vitriolic opposition to agricultural

biotechnology deserve special note for personalizing and spreading this dissenting style of politics. Rifkin is a guru, the one who founded the word of warning. His works have spread internationally and he would probably welcome the guru label, since he has personally and continuously contested genetic research since the mid-1970s. Labeling the problem after Rifkin also calls attention to the widespread growth of public interest groups that lobby against the unintended externalities of agricultural biotechnology and for alternatives to genetic experimentation. For example, Greenpeace has been a Rifkin disciple organization in Europe, bringing this same type of opposition and same rhetoric to nearly the whole of that continent's extreme Green politics.

In this and other issue areas, individuals have a strong tendency to be better identifiable than their own organizations. That is, Rifkin matters more than Greenpeace, or even more than his own group. As entrepreneurs, such people have founded numerous organized political interests, sought grants and donations to get and keep them going, made extensive contacts in policy-making and media circles, and lobbied as near free agents with their own personal agendas (34). Their organizations are often quite small, sometimes with no members. In opposing agricultural biotechnology, scientists and industry have often been confounded by the importance of these personalities, ranging from the aforementioned Lester Brown to the Land Institute's Wes Jackson and including the agrarian poet from Kentucky, Wendell Berry. As well-recognized entities, these individuals keep the attention of much of society because they personally make claims that followers enthusiastically disseminate. They matter less because of what they say than how and to whom they routinely say things.

Rifkin epitomizes this approach and came to command ongoing media coverage after the publication of his highly critical and controversial book, *Algeny* (35). Politically, even as a distinct outsider from government, Rifkin remains a major contact or at least reference for those who wish to investigate what might be the seamy side of agricultural biotechnology. Having only very small financial resources compared to the largest agribusiness firms, the Rifkins of international politics raise enough money to mount obstacles through law suits, frightening scenarios released to the press, and ongoing exposure to electronic media hosts who include them in news reports. With political and social institutions around the world both skeptical of new technologies and lacking integregation, Rifkin and other public interest entrepreneurs have proved to be influential in both starting and formulating debates. By articulating the semblance of public policy plans and by always having a story that reporters enjoy and pick up, public interest groups like that Rifkin follower Greenpeace often appear far better at basic politics than do advocates of emerging agricultural biotechnology. That explains why in the United States these people have won permanent friends in such agencies as FDA, EPA, and USDA's Food Safety Inspection Service. The same is true in the UN. In Europe they have won even more policy-making friends and have been very influential in delaying product introductions. This is true even in countries that are strong in biotechnology research. Yet even some developing nations are homes to the critic entrepreneurs, especially by using international grants. Others from developed countries also influence domestic policies in such developing countries. Robert Rodale of Rodale Press gained frequent attention for being the worldwide exporter of organic farming. And his successors have recently used Rodale's publications to link organic farming to biotechnology skeptism.

The Coalition Problem

Closely tied to the Jeremy Rifkin problem is that of coalition politics. Political coalitions are formed when different organized interests and social segments cooperate on public policy questions (36). The basic objective of a coalition is bringing together various interests so that they can share resources of money, political and media contacts, and even skill in order to win their agreed-upon common goals. Proponents of causes related to agricultural biotechnology have formed coalitions in order to gain government cooperation. Critics of biotechnology have organized coalitions in order to delay or prevent product introduction and regulatory approval. The coalition problem results because critics are more advantaged by such cooperation than is science and industry. That, for instance, makes Rodale Press opposition to biotechnology especially formidable when it is linked to the idea of organic production.

The explanation as to the opponents' advantage is quite simple. First, critics aim to stop something. Second, stopping or halting a product or technology is but a matter of creating doubts. Third, critics need not share all of the same values as reasons for their common opposition. They only need to agree to cooperatively bring impediments into the policy process. Thus minimal cooperation is necessary, and no plans for the details of future policy need be formulated. Their politics is really only about saying "no" in a unified voice. The more who say "no" together, the stronger is the coalition.

Proponents of biotechnology have contrary disadvantages. Given the uniqueness of products and technology, scientists and agribusiness firms must decide whether or not any cooperation is worthwhile. Cooperation on securing a patent or on approving a single pesticide is a task for the stockholder firm, not the entire industry. Moreover, too much cooperation between advocates of agricultural biotechnology may release trade secrets and give away firm advantages in the marketplace. That explains why most universities, research facilities, and corporations lobby on their own. And it explains why they also support more generically active common trade associations, the Biotechnology Industry Organization or the National Agricultural Biotechnology Council, when obvious policy questions such as binding international protocols unite the diverse proponents. This quite clearly is neither as simple or as easily agreed upon a cooperation as that of the opponents collectively and loudly saying "no" (37). Far more strategic options need be considered by proponents in order to pursue collective ends or even merely share information. As a consequence opposing coalitions tend to

grow to the largest possible numbers while coalitions of proponents tend to be kept far smaller, and therefore more internally conciliatory and politically weakened.

The net result is that coalitions tend then to be especially favored by Rifkin-style public interest groups. When Rifkin's Foundation on Economic Trends worked to stop the introduction of rBST, coalition politics was the obvious and favored route. All of the other three types of critical opponents were brought together to force greater study and long and costly industry delays. Sustainable agriculture specialists raised their questions under Rifkin's direction. Populist, smaller-scale dairy producers engaged with Rifkin in active social protest in favor of family farming. Academics who studied philosophy of science and policy ethics willingly shared their skeptical information (38). On the other side, as more industries that were once hoping to produce BST disassociated themselves from the innovation, Monsanto was left nearly alone to lobby for regulatory approval and against legislative roadblocks. In a more recent and publicly inspiring effort to capture attention, Rifkin helped organize the Pure Food Campaign. Nationally and internationally prominent chefs with important local followings gained Rifkin's views extensive publicity, even on cable TV's *The Food Channel*. The Beyond Beef Coalition was similarly conceived and attended.

The conclusion as to the impact of the coalition problem on science and industry is easy to draw: When it comes time to challenge agricultural biotechnology in public policy making, the interests are easily merged of those who represent each of the opposition problems. The critics do not exist in political isolation nor have secrets to necessarily withhold from one another. Thus critics of agricultural biotechnology may each be relatively small and resourceless and anti-science, but they still exist in cooperation with one another, with a resulting considerable political influence over corporate and research innovations.

SOCIAL IMPACT ON AGRICULTURAL BIOTECHNOLOGY

The social impact of agricultural biotechnology is full of both promise and fear, but for the foreseeable future, that impact is problematic. Sound sciences and capable industries make projections that must be far from reliable forecasts. Intelligent critics raise questions that are logical and rationally derived. As a consequence these emerging technologies may yield great benefits but they might also offer several social and political problems. No one knows. All extrapolations are only guesses as to probable impact.

The more immediate question for the early decades of the twenty-first century is on what social and political impact affects agricultural biotechnology. Answers exist for this query. First, any world order for governing agricultural technology's future will be fragmented and full of inconsistencies nation to nation. Second, what people believe even mythically in each society will exert strong political pressures that influence product and technology innovations. Rational science will not preordain adoption outcomes because near perfect information is never possible to have. National politics, as affected by international events, will prove determinate. Third, existing evidence indicates that a skeptical public, responsive public officials, and numerous critical positions on this technology will put distinctive and either unfortunate or fortunate limits on its ability to offer the social values it otherwise might. There exists a general suspicion both of new technologies and of corporate enterprises, especially in developing countries (39). That explains why public interest groups, as well as many academics, so often predict food scares that seldom, but certainly could, bring about the anticipated disasters (40). The rhetoric seems believable. The unfortunate plight of the monarch butterfly as it feeds close to genetically altered crops is an example that adds to the believability of critical commentaries. So too is the Muslim who confronts a tomato having a pig gene. People and politicians worldwide like change to be slow; and they can from that value position be easily influenced and mobilized by social and political critics to impose impediments from an anti-science perspective. Agricultural biotechnologies for these reasons will not soon escape either its controversies or its difficult politics.

BIBLIOGRAPHY

1. N.J. Seay, *Aipla Q. J.* **16**, 418–422 (1955).
2. P.J. Goss, *California Law R.* **84**, 1395–1836 (1996).
3. W.H. Lesser, *Animal Patents: The Legal, Economic, and Social Issues*, Stockton Press, New York, 1989.
4. M.M.J. Lin, *Am. J. Law Med.* **22**, 109–134 (1996).
5. J. Doyle, *Altered Harvests: Agriculture, Genetics, and the Fate of the World's Food Supply*, Viking Penguin, New York, 1985.
6. P.B. Thompson, *Agricul. Hum. Values* **14**, 11–27 (1997).
7. N.D. Hamilton, *Tulsa Law J.* **28**, 587–596 (1993).
8. S. Schuff, *Agri-Pulse* **13**, 1–2 (1997).
9. J.A. Nettleton, *Food Tech.* **50**, 28 (1996).
10. A.L. Young and M.R. Asner, *J. Nutr.* **126**, 1010S–1012S (1996).
11. L. Busch et al., *Making Nature, Shaping Cultures: Plant Biodiversity in Global Context*, University of Nebraska Press, Lincoln, NE, 1995.
12. E.I. Tsevdos, R.A. Chadwick, and G. Matthews, *Nat. Law J.* **19**, C1, C27–C31 (1997).
13. B. Hileman, *Chem. Eng. News* **73**, 8–17 (1995).
14. *The Economist* **341**, 21–23 (1996).
15. D.F. Hadwiger and W.P. Browne, in D.F. Hadwiger and W.P. Browne, eds., *Public Policy and Agricultural Technology: Adversity Despite Achievement*, St. Martin's Press, New York, 1987, pp. 1–13.
16. D. Avery, *Biodiversity: Saving Species with Biotechnology*, Hudson Institute, Indianapolis, IN, 1993.
17. D.D. Jones and A.L. Young, in G.R. Takeoka et al., eds., *Biotechnology for Improved Foods and Flavors*, American Chemical Society, Washington, DC, 1996, pp. 1–13.
18. L. Grant, *Forbes* **135**, 116–118 (1997).
19. W.B. Lacy, L.R. Lacy, and L. Busch, *Agricul. Hum. Values* **5**, 3–14 (1988).
20. K.P. Ruscio, *Sci. Tech. Hum. Values* **19**, 205–222 (1994).
21. B.E. Rollin, in S.M. Gendel et al., eds., *Agricultural Bioethics: Implications of Agricultural Biotechnology*, Iowa State University Press, Ames, IA, 1985, pp. 292–308.

22. M. Sagoff, *Agricul. Hum. Values* **3**, 26–35 (1988).
23. A. Gore, Jr., *Gen. Eng. News* **7**, 4 (1987).
24. G. Comstock, *Agricul. Hum. Values* **3**, 36–52 (1988).
25. S.H. Priest, *J. Comm.* **45**, 39–54 (1995).
26. H.A. Schneiderman and W.D. Carpenter, *Env. Sci. Tech.* **24**, 472 (1990).
27. J. Kloppenburg, Jr. and B. Burrows, *The Ecologist* **26**, 61–67 (1996).
28. D. Paarlberg, *Farm and Food Policy: Issues of the 1980s*, University of Nebraska Press, Lincoln, NE, 1980, pp. 5–13.
29. W.W. Cochrane, *The Development of American Agriculture: A Historical Analysis*, University of Minnesota Press, Minneapolis, MN, 1979, pp. 378–395.
30. W.P. Browne et al., *Sacred Cows and Hot Potatoes: Agrarian Myths in Agricultural Policy*, Westview Press, Boulder, CO, 1992, pp. 5–16.
31. J. Burkhardt, *Agricul. Hum. Values* **3**, 53–60 (1988).
32. W.P. Browne, *Cultivating Congress: Constituents, Issues, and Interests in Agricultural Policymaking*, University Press of Kansas, Lawrence, KA, 1995, pp. 22–39.
33. P.H. Mooney and T.J. Majka, *Farmers' and Farm Movements: Social Protest in American Agriculture*, Twayne Publishers, New York, 1995.
34. W.P. Browne, *Private Interests, Public Policy, and American Agriculture*, University Press of Kansas, Lawrence, KA, 1998.
35. J. Rifkin, *Algeny*, Viking Press, New York, 1983.
36. K. Hula, in A.J. Cigler and B.A. Loomis, eds., *Interest Group Politics*, 4th ed., Congressional Quarterly Press, Washington, DC, 1995, pp. 239–258.
37. W.H. Riker, *The Theory of Political Coalitions*, Yale University Press, New Haven, CT, 1962.
38. W.P. Browne and L.G. Hamm, *Policy Studies J.* **17**, 181–192 1988.
39. A. Wildavsky, *But Is It True? A Citizen's Guide to Environmental Health and Safety Issues*, Harvard University Press, Cambridge, MA, 1995.
40. S.C. Phillips, *CQ Researcher* **4**, 675–691 (1994).

See other entries AGRICULTURAL BIOTECHNOLOGY; ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE.

AGRICULTURAL BIOTECHNOLOGY, LAW, APHIS REGULATION

DAVID R. MACKENZIE
University of Maryland
College Park, Maryland

OUTLINE

INTRODUCTION

Agriculture in the United States is extremely productive today because of extraordinary natural resources and technological efficiency. The present capacity of the U.S. agricultural production system (from field to fork) is a result of scientific contributions made by both the private and public sectors. It rests on a foundation of basic and applied research outcomes delivered to the intended users through multiple technology transfer mechanisms.

Science in general, and U.S. agricultural research in particular, has been for the most part exempted from regulatory oversight, with some notable exceptions. Research scientists have for some time been expected to comply with federal regulations regarding the handling of radiological materials, and they are required to obtain institutional permission to conduct research with human subjects. Federal law also sets strict standards for the care of research animals. In recent times research scientists have been required by federal law to obtain certification to handle registered pesticides, although many types of hazardous chemicals (including experimental pesticides) have long been exempted from federal regulation in small-scale tests.

One notable exception to this regulatory pattern for research activities is the strict federal requirement on the interstate shipment of plant pest, pathogens and noxious weeds. To move a "regulated article," from one state to another, or to import a "regulated article" requires a permit under the Federal Plant Pest Act. The Plant Pest Act is administered by the United States Department of Agriculture (USDA) through the Animal and Plant Health Inspection Service (APHIS), a division of the USDA's Food Safety and Inspection Service (FSIS) (1).

Thus, for the most part, during the first three-quarters of the twentieth century agricultural researchers in the public and private sectors were mostly free to manipulate the genetics of animals and plants, and even microorganisms, with little government oversight or regulatory attention. The freedom to investigate all types of organisms for

applications to crop and livestock improvement came to be a common expectation. New technologies emerged and flourished, as did U.S. agricultural production.

In the United States this virtual absence of scientific regulation seemed to work, and when something untoward happened, solutions were modest. This point is perhaps best typified by the federal government's response to the 1969 release of a potato clone named Lenape. This cultivar was soon discovered to have a significant concentration of poisonous glycoalkaloids in the tubers, under most commercial growing conditions. Lenape was the product of a conventional breeding program that had used some wild species as parents to obtain superior potato chip processing quality. It has been presumed that the wild parentage brought to the progeny high glycoalkaloid concentrations that later required its withdrawal from commercial production. The federal government's response to public concerns was to establish, through the U.S. Food and Drug Administration (FDA), requirements for establishing the safety of newly released cultivars. The agency's Generally Regarded As Safe (GRAS) guidelines asked plant breeders to give self-assurances that what was about to be released was at least as safe as the cultivar(s) to be replaced. This "self-policing" approach to the problem was generally well received within the scientific community. And it was fairly typical of the federal government's hands-off approach to research-related concerns for environmental, public health, and food safety issues. The GRAS guidelines are still in place, but there is no GRAS police force.

AGRICULTURAL BIOTECHNOLOGY

In the early 1970s, the relationship between society and the scientific community began to change as a result of research breakthroughs in the ability to manipulate the cells, tissues, and the genetic code of plants and animals (2). These collective technologies were subsequently called *biotechnology*. Understanding the technology's components is necessary for gaining an understanding of the shift that occurred in the regulation of biotechnology in general, and agricultural biotechnology research in particular, by the federal government.

Tissue and Cell Culture

The culture of living tissues, and subsequently of individual living cells, has long been available to research scientists. Plant tissue culture dates back six decades. But in the 1970s scientists discovered how to chemically dissolve cell walls and grow single protoplasts (naked plant cells). This represented a technological advancement of significance to plant researchers. Similarly, in the animal science research community, the ability to culture animal tissue and individual cells opened up new areas of investigation. At the time, none of these discoveries were considered to represent a risk worthy of federal regulation.

Regeneration

Subsequent research discoveries allowed the regeneration of individual cells and tissues into once again whole organisms, through a phenomenon called *totipotency* (which is actually very poorly understood). This remarkable biological characteristic means that certain types of individual living cells have all of the genetic information necessary to become a complex organism, and can do so when given the right culture conditions.

Regeneration of plants and animals also was not considered worthy of federal regulatory attention. This was true until recently when the cloning of some higher animals (notably with the cloning of the sheep named Dolly) raised public concerns regarding the ethics of a technology that might lead to the cloning of humans. Still, no federal regulations for sheep cloning are being seriously considered.

Transformation

The third emerging technology is genetic transformation which was first described in bacteria. This discovery had obvious applications to higher organisms, once it was understood that complex organisms could be reduced to a single cell (or small group of cells), transformed, and then regenerated. Taken together, tissue or cell culture, genetic transformation, and regeneration have become the tools of genetic engineering (a.k.a. biotechnology).

PLANT GENETIC ENGINEERING

The genetic engineering of plants first occurred in 1982 at the University of Wisconsin. Researchers there genetically engineered a protein from sunflower into a bean plant. This "genetic transformation" technology later became the standard for genetically engineering plants. Involved in genetic transformation processes are the following steps.

- Identification of a specific *genetic sequence*
- Isolation of the sequence by the use of *restriction enzymes*
- Matching the result to a *promoter sequence*
- Application of a gene *vector*
- Insertion into a host plant's *genome*
- *Regeneration*.

Both the vector and the promoter DNA sequences are essential to the transformation process. A *marker gene* (e.g., a gene for antibiotic resistance) is also desirable in the *genetic construct* to more easily identify the successfully transformed cells or tissues.

The terminology associated with genetic engineering requires some explanation. Restriction enzymes were first named because when they are applied to cultures of a bacterophage, the plague's (i.e., infected bacteria's) growth is "restricted." Subsequent research showed that the restricted infection was due to the presence of enzymes, but it was not known why they restricted plague growth. In the end it turned out to be the result of ribonucleases (RNA enzymes). These enzymes were cutting the DNA at specific sites in the DNA sequence, thus restricting infectivity. The proteins came to be termed "restriction enzyme" (because the bacteria's growth was restricted). Technically they are more accurately called endonucleases, since they cut the DNA sequences internally (endo).

Gene vectors come in many forms, but the most commonly used (and highly efficient) vector in plants was found in the plant pathogen *Agrobacterium tumafasciens*. This pathogen has a wide host range and induces a cancer like growth near the soil line, which accounts for the common name for the disease, crown gall (3).

Significant research funding in the 1970s from the National Institutes of Health (NIH) for investigation of crown gall was likely a consequence of their interest in the cancer-like growth promoting characteristic of the pathogen. Thus, as a benefit of cancer research, plant pathologists were able to discover that a plasmid (an extra-chromosomal ring of DNA) harbored a sequence for tumor induction (commonly referred to as the TI plasmid) (4). The plasmid had the capacity to insert itself into host DNA, so it could serve as a vector for attached DNA. The attachment of cloned DNA to the TI plasmid came to be called a *construct* and targeted sequences could be used to transform plants that were susceptible to *Agrobacterium tumafasciens*.

It was soon discovered that in order to get the DNA sequences to express themselves effectively, a promoter sequence had to precede the gene of interest. The commonly used promoter sequence during the early years of plant research was also from a plant pathogen, the cauliflower mosaic virus.

All of this is important to subsequent regulation of biotechnology by the Animal and Plant Health Inspection Service (APHIS) of the U.S. Department of Agriculture (USDA) under the Plant Pest Act, as we will see later.

HISTORY OF BIOTECHNOLOGY REGULATION

In 1974 a group of about 300 concerned scientists met at the Asilomar Conference Center in Monterey, California, to discuss their concerns for the safety of research with recombinant (a.k.a. rDNA) organisms (5–7). At that time virtually all genetic engineering was being done with bacteria. There were no biosafety standards for the handling of these biological materials in research laboratories. Moreover much of the biotechnology research at the time was medically related, often involving human pathogens. Fear of an epidemiological disaster that might result from an unintended release of a genetically engineered human pathogen prompted the Asilomar conference. From that conference emerged a set of voluntary guidelines that set levels of containment for the handling of recombinant organisms, primarily based on human disease hazards (8,9). That is, the disease risk of the research organisms being handled helped determine the levels of containment, which increased with the level of concern for human pathogenicity. The guidelines described handling and containment protocols for various types of research microorganisms.

NIH Guidelines

The Recombinant DNA Guidelines were to be administered by NIH, and they were to be voluntary (10–12). In response, the NIH created the Office of Recombinant DNA Activities (ORDA). ORDA relied on a panel of experts to review application and make recommendations to the director of NIH on proposed research protocols.

The NIH guidelines were more than self-policing. The awarding of federal funds for research projects required compliance with the guidelines. It was commonly asserted at the time that voluntary compliance by the private sector and nonfederally sponsored public sector research was virtually 100 percent. Some individuals concluded that the threat of civil penalties for negligence induced many public and private laboratories to adopt the NIH rDNA guidelines, even though technically they were not required to do so. Although never tested in court, it is presumed that with NIH approval the researcher was doing what a reasonable and prudent scientist would have done. Negligence would then be hard to prove. Thus the endorsement by the NIH's Recombinant DNA Advisory Committee (RAC) became a standard for research protocols for new and novel recombinant organisms, but only for contained laboratory research.

By the early 1980s it became obvious that biotechnology would have many applications in agricultural science (13). What was not known at the time was how to provide biosafety assurances for field-testing (i.e., tests to be conducted outside of a contained laboratory or greenhouse) with organisms of recombinant parentage (14–17). The NIH guidelines were strictly for controlled laboratory facilities, and inasmuch as they were primarily based on human pathogenic traits, the 1982 announcement of transgenic sunbeans drew attention to biosafety issues, rightly or wrongly. Many felt that something had to be done by the USDA to ensure the safety of organisms being handled in research activities outside of containment (at least for the projects if funded). Thus began USDA's two-track approach to biosafety assurance.

USDA's Two Track Approach

The USDA is a complex structure with responsibilities that required it both to promote and to regulate scientific activities. At the time (early 1980s) the Assistant Secretary for Science and Education, Dr. Orville Bentley, foresaw a need to extend the NIH recombinant DNA guidelines to cover field experimentation with plants and animals. This would be done in ways to provide public assurances of the safety of research funded through his office. At the same time, NIH was expressing little interest in assuming responsibility for safety assurances for the environmental release of recombinant DNA organisms. And thus was born the idea for an Agricultural Biotechnology Research Advisory Committee (ABRAC), patterned after the NIH RAC. (To be factually accurate, the Committee on Biotechnology of the National Association of State Universities and Land-Grant Colleges first proposed to the USDA that it should use the NIH rDNA guidelines model for the oversight of field tests with recombinant organisms.) Additionally ABRAC was given staff support by the Office of Agricultural Biotechnology (OAB), which was obviously patterned after the NIH ORDA. But ABRAC/OAB was to address the biosafety (i.e., the environmental and public health) questions arising from research tests conducted outside of containment facilities.

Simultaneously, the regulatory arm of the USDA (which houses APHIS) began exploring its authority under the Federal Plant Pest Act to regulate genetically engineered plants as plant pests. This created a dual track situation and set up an interesting dynamic between the research side and the regulatory side of the same federal department.

Meanwhile, other federal regulatory agencies were giving thought to asserting their authorities to regulate biotechnology in ways that would impact agricultural research. The Food and Drug Administration (FDA) initiated notable regulatory activities, under the U.S. Food, Drug, and Cosmetic Act. Also the U.S. Environmental Protection Agency (EPA) became active under the Federal Insecticide, Fungicide, and Rodenticide Act and the Toxic Substances Control Act.

Coordinated Framework

By 1983 it was becoming obvious that a plan was needed for assuring the safety of biotechnology (18). Some interest groups argued that regulation was needed at the federal level, to coordinate better regulatory actions, and perhaps to preclude state-by-state, or county-by-county, or even city-by-city regulation of biotechnology research and commercial development. At the time one of the greatest fears of the technology's champions was a patchwork quilt of federal, state, and local government regulations that would make product commercialization not feasible.

In response to this need the White House Office of Science and Technology Policy (OSTP) called together federal regulatory and research agencies to map out a coordinated plan for the regulation of biotechnology. In June 1984 the OSTP published in the Federal Register a proposed Federal Coordinated Framework for Regulation of Biotechnology (herein after referred to as the Federal Coordinated Framework) (19,20). Intense public debate ensued, and the debate still has not subsided.

There were two fundamental principles of the Federal Coordinated Framework:

- No new laws were needed to regulate biotechnology, because existing laws were adequate.
- All regulations would be based on the product, not the process, of biotechnology (21).

Both points need some explanation.

Leading proponents of biotechnology, particularly the pharmaceutical industry, argued strongly at the time that sufficient regulatory authority already existed to ensure the safety of biotechnology (research and commercialization) and that no new laws were needed. It was argued, particularly strongly, that the pharmaceutical, drug, and medical device industries were already subject to extensive clinical trials and product registration. Adding another layer of regulatory oversight would unfairly and unnecessarily slow the development of the commercial products from this new and exciting technology.

Opponents of the Federal Coordinated Framework came from many sectors, including the scientific community that was, as noted above, unaccustomed to regulatory oversight for research activities heretofore not regulated. Environmental interest groups saw the Coordinated Framework as inadequate, and they registered their concerns during the public comment period. Interestingly the U.S. Department of Interior, which historically has responsibilities for aquatic organisms through the Fish and Wildlife Service, was not invited by OSTP to participate in the first rounds of the Federal Coordinated Framework's development. This later turned out to be a major oversight, as aquatic species, especially genetically engineered fish, soon became one of the major environmental safety concerns. No provision for regulatory oversight of fish or shellfish, or even informal research guidelines, were proposed through the Federal Coordinated Framework. Additionally the USDA's Food Safety and Inspection Service (FSIS), and FDA have long sought for themselves regulatory authority over fish and shellfish. Thus, with no existing law and no place at the table, a biosafety void was created.

Following a substantial period of public criticism, the Federal Coordinated Framework was formally announced in June of 1986 as a plan for biosafety assurance (22). Lead agencies were identified with specific responsibilities. Regulations were to be built on existing legal authorities, and with a focus on the *product*, not the process of biotechnology.

The second point of consideration became an issue of contention when defining the scope of regulatory authority, especially for obtaining White House approval to implement the regulations (23). Regulation of the products (not the processes) required that any regulatory wording could not single out biotechnology as a process. The regulations had to identify the product that was being regulated, and only if it specifically represented a biosafety hazard requiring federal regulation. This seemingly logical approach to regulation fit well with the needs of the pharmaceutical industry that had long been required to verify product safety, quality, and efficacy. Under the new rules this assurance was irrespective of whether the product was recombinant or conventional. Policy makers argued that it does not matter whether insulin was to be provided as a recombinant product or not. The process was irrelevant. It was, they said, the product that was important for making biosafety assurances.

As one might guess, significant problems occurred in several areas of scientific research from this policy since heretofore many research activities, for the most part, were not regulated. As noted above, this was particularly true for agricultural science. Many scientists shuddered to think that a federal regulatory apparatus would be imposed on their activities, on an experiment-by-experiment basis.

In the worse case scenarios, some critics lamented, a new, perhaps unforeseen hazardous consequence might occur in field tests with rDNA organisms. Would the transferred genes be stable? Would the traits be expressed in ways that heretofore were not seen? Would pleiotropic effects be expressed? Would recombinant organisms have a superior ecological advantage or greater fitness over natural types? Would unknown or unanticipated characteristics cause environmental, public health, or safety problems?

An additional biosafety issue that needed to be addressed was the capacity of biological organisms to reproduce and disseminate, once released into the environment. Would the Genie, once out of the bottle, become an unretrievable problem? Could experiments be designed to contain or mitigate an organism released into the environment? And who would do the biological monitoring?

To implement the Federal Coordinated Framework, USDA decided to move ahead with its two-track approach (24). This led to some interesting consequences.

APHIS began, in quick succession, its proposed rule making, followed 30 days later with final rules to regulate the "products" of plant genetic engineering, using its Federal Plant Pest Act authority (25). The final rule clearly states "genetic engineering" as the regulatory trigger, thus seemingly violating the product-not-the-process terms of the Federal Coordinated Framework. The intended regulatory targets were the plant pathogen-derived vectors and promoters of genetically engineered plants. Through this regulatory strategy the plants themselves would become the regulated articles as "plant pests," under the Federal Plant Pest Act. The White House policy reviewers allowed this curious strategy. This way APHIS regulatory authority was put in place for plant biotechnology.

Simultaneously USDA's Science and Education Office drafted guidelines for field testing recombinant organisms (26). The guidelines met with mixed results.

First, the ABRAC guidelines attempted to prescribe physical and biological "confinement" practices in anticipation of field experiments yet to be proposed. That task proved to be too daunting. The deployment of recombinant DNA confinement methods on so many forms of organisms was evolving faster than the advisory committee could come up with procedures that could be accepted as reasonably safe. Notably, drafts of the ABRAC guidelines met with severe criticism, leading to more revisions and considerable frustration by ABRAC members.

Second, federally funded scientists were required to get both a permit from APHIS and an ABRAC review of their research protocol. The science community saw this as double jeopardy (27).

Also applicants expressed frustration with the delays in obtaining their reviews. Some ABRAC members began to question the validity of their own decision-making process, and the legitimacy of using peer review for biosafety risk assessments (28).

Meanwhile, USDA's Science and Education office created the National Biological Impact Assessment Program (NBIAP) with responsibility for facilitating safe agricultural and environmental biotechnology research (29). NBIAP was to monitor progress, foster biosafety communication, and focus research activities on priority biosafety issues. As an independent office, the NBIAP could work across agency lines and with institutions external to USDA to identify emerging biosafety issues and expedite solutions. During a very short period of time, and with limited funding, the USDA's Cooperative State Research Service, working with several Land-Grant University partners, developed a protocol:

- A biotechnology information bulletin board (a forerunner World Wide Web)
- A compilation of all APHIS generated Environmental Assessments for field testing permits as a CD-ROM
- An expert system for assisting scientists with completing an application for an APHIS permit
- An annual international conference, co-sponsored with U.S. EPA and Environment Canada, on biosafety research results
- A $1.7 million competitive grants program in biotechnology risk assessment research

To this last point Section 1664 of the 1990 Farm Bill established a competitive grants program in biotechnology risk assessment to begin answering some of the risk assessment questions brought forward by the critics of agricultural biotechnology.

NBIAP continues today, renamed as the Information Systems for Biotechnology (ISB), with modest USDA funding to Virginia Polytechnic Institute and State University (Virginia Tech).

THE APHIS REGULATORY PROCESS

As noted earlier, under the requirements of the Federal Coordinated Framework, APHIS could not seek new legal authority. APHIS had to use the provisions of the Plant Pest Act to write regulations to assure the safety of recombinant plants being tested outside of containment. This required some significant reinterpretations of the Plant Pest Act, including promulgating provisions to regulate the movement of articles intrastate as well as interstate, and the use of on-site inspections to verify the conditions stated in the application for a permit. Some individuals said that the statutory authorities that APHIS had to claim for these activities exceeded those assigned in the Plant Pest Act, including the regulation of genetically engineered plants as a "plant pest."

The National Environmental Policy Act (NEPA) requires a federal agency, when making a decision that may have a significant impact on the environment, to conduct an Environmental Assessment (EA) of the various options considered, and to post the reasons for the final determinations. Although some said it was unnecessary, APHIS decided to comply with the NEPA by requiring each permit application, and the agency's own determination of plant pest status, to undergo its own EA, on a permit-by-permit basis. This decision was no doubt driven by the experience of other federal agencies (e.g., the Department of Health and Human Services, DHHS) challenged in federal court by the Foundation on Economic trends on procedural missteps.

Implementing the provisions in the NEPA caused APHIS to hire or reassign considerable staffing to prepare the lengthy documents assessing the environmental consequences of a proposed experiment with a genetically engineered plant. This in turn required that the applications for an APHIS permit had to provide information sufficient for the agency to make its determination and complete an EA. From this emerged

the NBIAP idea of designing an expert system that would build on the experiences of past permit applications and the associated EAs. The expert system was to ensure the design of safe experiments with genetically engineered organisms and to facilitate the drafting of similar permit applications. Secondarily, the expert system was to facilitate, to some extent, compliance with ABRAC guidelines. NBIAP software was developed and distributed in 1988, and updated versions were periodically made available free of charge to scientists upon request. Sample paragraphs were offered by the expert system for adoption, revisions, or technical correction, based on previously successful applications. Applicants made responses to a series of organism-relevant questions relating to the permit decision-making process. Key biosafety questions were developed to identify high-risk situations. Responses to some of the critical expert questions resulted in advice to the applicant that under no circumstances would APHIS be likely to issue them a permit. The NBIAP expert system software was well received by some, but it was not extensively used by the scientific community, for reasons that were never well understood.

Following the preparation of EA by APHIS, one of two determinations was made, each requiring more documentation by APHIS. If there was a Finding of No Significant Impact (FONSI), the document was filed with the EA, and the availability of the EA and FONSI documents was then announced in the Federal Register. When more than 300 Environmental Assessment/FONSIs had accumulated, NBIAP assembled the documentation into a searchable database and issued it on a CD-ROM, as a service to the research community and biotechnology interest groups.

If the EA made a finding of a significant impact, the agency would have been required, under NEPA, to prepare an Environmental Impact Statement (EIS). An EIS is often an enormously large document and always requires considerable technical detail, large amounts of information, and sophisticated analysis. Some critics note that to their surprise, APHIS never made a finding of significant impact for a proposed field test with a recombinant plant, and thus was never required to write an EIS. In response to this criticism, APHIS countered that if a permit application was submitted that would have led to a conclusion of a significant environmental impact, the applicant was so notified and was provided an opportunity either to redesign the experiment or to withdraw the application. Inasmuch as the application process was not a public record, little information exists beyond FONSIs.

In the 10 years following the initiation of the APHIS permit process, more than 3000 field trials were conducted in the United States, and more than 30 products were commercialized. This high level of success is related to a number of legal, policy, and political issues that needed to be resolved to make the APHIS permit process work.

LEGAL, POLICY, AND POLITICAL ISSUES

APHIS had a number of issues that needed to be resolved for the successful implementation of the agency's field-testing permit system. In the end APHIS was successful in resolving most of these issues, and this required them to expend a lot of time and attention on their resolution.

Conflicting Legal Authorities

In 1990 Congress proposed to hold hearings on the Omnibus Biotechnology Bill (OBB) that would have standardized the regulation of the process of biotechnology research and the commercialization of the resulting products through one federal agency (30,31). Although not specifically named in the bill, it was presumed at the time that the U.S. Environmental Protection Agency would be assigned the authority to administer the OBB.

The OBB had both opponents and proponents who were equally outspoken. A congressional hearing placed those issues on the table and the divisions were clearly evident. Proponents liked the idea of "one-stop shopping" through a centralized regulatory process that would end "shopping around" and would promote the consistent application of biotechnology regulation with an environmental perspective.

Opponents of the proposed bill (primarily the biotechnology industry) were very content with their regulatory experience, particularly with the APHIS office that issued the field-testing permits, because they understood the procedure. Under the careful guidance of attorney Terry Medley, the unit gave careful and courteous attention to their "customers" (in the sense of Ed Deming's Total Quality Management). Although criticized as being too helpful, the unit became known as an office that returned its phone calls and answered correspondence in a timely manner. As a consequence permit applicants became supporters and strong defenders of the existing APHIS regulatory process, and the services provided. (An interesting historical comparison could be made with the EPA regulatory office. EPA was viewed at the time, as an adversary of the technology and less than helpful to their applicants.)

Thus, when Congress presented the OBB for public comment, strong industry support for the existing regulatory procedures defended the status quo, and the Federal Coordinated Framework continued as originally devised.

Policy Conflicts

Regulatory policies that were in conflict with other agencies were addressed by APHIS through ongoing conversations, staff exchange, and high-level consultations. This open dialogue with other agencies avoided conflicts (and the consequent interference) with APHIS's regulatory decisions. When questions arose regarding lead-agency responsibilities, APHIS was quick to move to a satisfactory resolution. For example, when EPA had overlapping authority with APHIS on plant pathogenic microorganisms with pesticidal properties, they conferenced to work out the differences. As a result of these patterns, it became relatively easy for APHIS staff to understand their authorities vis-à-vis other regulatory agencies, and to act decisively.

Political Issues

APHIS was particularly effective in providing information through congressional hearings during the uncertain

early years of regulating agricultural biotechnology. Even during highly charged hearings, APHIS presented carefully thought out arguments for why it was necessary to apply the Plant Pest Act for the regulation of plant biotechnology products.

It is important to note that the political milieu of the federal government during this period was a Republican White House and a Democratically controlled House of Representatives. Traditionally it is said that Republicans favor industry and business, while Democrats favor environmental stewardship, equity, and other issues that may be at odds with commerce. This placed APHIS in an awkward situation, seemingly in need of representing, through the USDA, a pro-industry approach to biosafety assurance. It provided opportunity for special interest groups to prod members of Congress to hold hearings, some of which were designed to embarrass the administration. Thus a spotted congressional record exists on APHIS's accomplishment, probably reflecting the political situation of the time more than the merits of the accomplishments of APHIS.

Special Interest Groups

The Environmental Defense Fund, the National Wildlife Federation, and the Foundation on Economic Trends provided a continuing challenge to the APHIS regulatory process (32,33). This effort met with mixed results (34).

Retrospectively, these special interest groups probably provided an important public service as watchdogs to the biotechnology regulatory process. They no doubt deserve credit for serving as a public conscience for agricultural biotechnology. And they served well as an information funnel for APHIS. Many of their arguments were founded on solid science and careful investigation. However, more often their arguments showed an absence of any scientific knowledge, one way or the other. This left many observers to question the validity of regulatory decisions being made by APHIS. But this was an important contribution as well, as it helped identify areas needing more research to uncover the scientific facts required for better regulatory decision making. And it was a constant reminder that not all things that matter can be quantified.

Section 1664 of the 1990 Farm Bill established a Biotechnology Risk Assessment Research program, as a 1 percent set aside of USDA's outlays for biotechnology research (yielding about $1.7 million annually). It was understood at the time that this provision in the Farm Bill was placed into law at the insistence of environmental interest groups. The funding allowed establishment of a targeted competitive grants program to begin answering some of the risk assessment questions brought forward by both critics and the regulators of agricultural biotechnology. This competitive grants program, although opposed by some influential leaders in the scientific community (who saw it as an unnecessary admission of biotechnology's risks), was able to resolve some of the questions raised by outspoken critics of agricultural biotechnology. The program continues today, but with less than enthusiastic support from either the USDA or the agricultural research community.

Trade Groups

The Biotechnology Industry Organization (BIO) and the National Association of State Universities and Land-Grant Colleges' Committee on Biotechnology were the primary trade groups interacting with APHIS during the early years of agricultural biotechnology regulation (35,36). The trade groups tended to operate as a counterbalance to the special interest groups by seeking to voice the concerns of the public and private sectors over unneeded and unnecessary regulatory burdens. APHIS was responsive to this perspective by periodically reviewing its permit application procedures, by disseminating information on permits issued and eventually, by the mid-1990s, converting to a notification procedure for six crops (and later, additional ones) that eliminated the need to apply for a permit under fairly broad circumstances. The APHIS notification procedure greatly facilitated field experimentation in prescribed areas, which was justified by the Agency's accumulated experience and knowledge derived by that time from issuing hundreds of permits for field tests, which were all carried out safely.

An additional influential trade group was the National Agricultural Biotechnology Council (NABC) that performed a different service for APHIS (37). NABC was a coalition of universities that annually sponsored a forum for dialogue primarily focused on the biosafety issues of agricultural biotechnology. The proceedings of their annual meetings clearly depict a pattern of an evolving consensus on how to approach the biosafety questions of agricultural biotechnology research.

Science Community

The scientific community became divided on the issues of biosafety, mostly along scientific discipline lines. Plant pathologists gave arguments that the Federal Plant Pest Act was, scientifically, not a legitimate legal authority for regulating the products of agricultural biotechnology (38). They used contemporary scientific information to question the supposition that a disarmed TI vector or a promoter sequence from a virus could in any way lead to a plant becoming a plant pest. To many plant pathologists this supposition was an absurdity (39,40).

To the molecular biologists, the APHIS regulatory approach seemed overly heavy-handed, unnecessary and probably an impediment to the agricultural applications of biotechnology (41,42). Molecular biologists appeared at the time to be giving little attention to the environmental consequences of field releases with genetically modified organisms, even though problems might occur in very low frequency.

Ecologists had a very different perspective. They foresaw that severe consequences could result from "releases" of recombinant DNA, based on other "environmental disasters," such as the gypsy moth, Dutch elm disease, and the kudzu vine (43,44). Several serious and scholarly treatments of the issues of biosafety were published during this period. The failure to resolve the differences among the plant pathologists, molecular biologists, and ecologists stemmed from the absence of a factual basis for decision making.

It must be noted that APHIS never asserted that its regulatory authority was based on risk. It was merely a determination of whether or not the organism was a plant pest, and therefore a regulated article. At the time many individuals in the scientific community seemed to have been arguing for a risk-based determination of the safety of recombinant organisms being tested in the environment. This would have been a major challenge for APHIS, since much of the necessary information was missing (and still is) for a thorough risk assessment of proposed field tests with recombinant DNA organisms. Moreover the fundamental risk analysis paradigm, as presented in the 1983 National Research Council's "Red Book," proposed a conceptual separation between risk assessment and risk management (45). Risk assessment is a science-based, stepwise process, for which the first step is hazard identification. Part of the biosafety controversy was over arguments that biotechnology in and of itself represents a hazard, and therefore the whole process needed to be regulated. Others argued that biotechnology was not in and of itself a hazard, and thus not a risk. Out of this paradigm difference came the product versus the process debate. The Federal Coordinated Framework sided with the second view (i.e., regulating the product). But the argument continues today as to whether biotechnology itself represents a hazard, and thus should be subject to risk assessment (46–48). APHIS redirected this argument by asking whether or not the organism was a plant pest, and thus a regulated article.

Another complication regarding the application of the 1984 Red Book approach to risk analysis was the identification of the appropriate authority for risk management. According to the Red Book, a firewall needs to be created between risk assessment and risk management. The first activity is a scientific process used to derive a science-based recommendation, while the second process is policy implementation that takes into account scientific fact and public consideration. The scientific community in general has long been resistant to the notion that anything other than scientific fact should be used to decide the safety of an organism (49). However, in a democracy, sometimes the best scientific evidence is not sufficient to gain social or market acceptance.

The APHIS strategy to focus on a determination of pest status completely avoided the risk assessment/risk management question, but much of this finesse was lost on the scientific community. The science-oriented critics of APHIS were focused on an all together different set of questions.

Industry and Academe

The transactional cost of complying with the federal regulatory permit application process segmented industry from academe. Many of the larger companies that were investing heavily in agriculture biotechnology (e.g., Monsanto, Calgene, CIBA-Geigy) could afford staff assistants and legal support to help in the preparation of lengthy permit applications for field tests. Universities, on the other hand, argued that they could not afford to hire such staff. Thus the regulatory burden fell upon the shoulders of the individual faculty member wishing to conduct the field research. During the early years of APHIS permit issuance, the ratio of private to public sector was approximately 9 to 1. This caused considerable alarm, as it appeared that the regulatory apparatus was interfering with the normal flow of research activity from the laboratory to the field, at least at the public institutions and federal research laboratories.

It was this point of concern that caused the NBIAP to conduct a national survey in 1990 on the impact of regulating agricultural and environmental biotechnology research in both the public and private sectors (50). The survey method was face-to-face interviews with open-ended questions. The questions were designed to determine the degree of interference with the research process being caused by regulatory requirements, with a particular focus on APHIS.

The results of the NBIAP survey were surprising, from two perspectives. First it was discovered that there was not a big backlog of research results awaiting field testing for lack of permits. Interviews with university scientists mostly noted their apprehensions, but little specific biological material could be cited as ready but not yet tested. Moreover a significant portion of the private sector's permits had been issued by APHIS to private sector and university scientist working in partnerships to jointly conduct field tests with recombinant plants. But, inasmuch as the permit was issued to the private company, the participation of the public institution was lost in the calculated ratio of private to public sector permits.

Second, when asked about the transactional costs for university scientists, an APHIS officer described a permit request they received from a university scientist that had been handwritten on yellow paper, with an attached hand-drawn diagram of the genetic construct and a crude plot map of the proposed test site. To APHIS's credit, they began to work with the scientist to develop an adequate permit application that had the required information. APHIS eventually issued a field test permit to the institution. This test became the university's first sanctioned field test of a recombinant plant. In defense of APHIS, and contrary to alleged burdens of the permit application process, it is to their credit that they did not summarily reject the application but instead worked with the scientist to assemble a proper application for a permit.

The results of the NPIAP survey clearly demonstrated high levels of satisfaction with APHIS's permit application process, services, and the professionalism of the agency's staff. As noted earlier, this was very likely the consequence of their customer-focused attention, in the sense of Ed Deming's Total Quality Management, a then-contemporary model for the federal government. Ironically this also led to criticism that APHIs was too friendly with its permit applicants. One then needs to ask, what is the proper balance between government agency responsiveness and regulatory adversity?

Trading Partners

U.S. agriculture has maintained for decades a positive trade balance that is derived from its reliability as a source of many types of commodities. Among these

commodities are the basic grain and oil crops that are traded internationally as a very profitable business.

Early in the process of identifying the benefits of agricultural biotechnology, it became apparent that the acceptance of recombinant commodities by U.S. trading partners would become a key issue, if U.S. agricultural biotechnology-derived commodities were to be accepted in global commerce. Recombinant corn, soybean, and cotton are just some of the commodities that are now a point of contention in international trade markets. This pattern will undoubtedly continue to grow unless some significant change comes about. What is not known is the market receptivity and consumer acceptance of those recombinantly derived commodities. One could conclude that the eventual consumer receptivity of the products of agricultural biotechnology will depend more on the regulation of the process (as Europe has done), rather than regulation of the products (as the United States has done).

Early on, APHIS saw the need to begin discussions with our economic trading partners. This was done through the Paris-based Organization of Economic Cooperation and Development (OECD). Through this forum APHIS began the process of finding common ground and agreement on biotechnology regulations impacting agricultural products. The OECD was in a good position to provide a forum for the discussion of how to harmonize biosafety assurances. Working through the U.S. Department of State, and by coordinating USDA activities, APHIS gave early leadership to the development of documents that set out concepts and expectations for regulatory requirements. This turned out to be difficult in Europe, which is noted for its resistance to the application of biotechnology to food products. After 15 years of discussion and consensus building, there are no formal agreements on the acceptance of the products of U.S. agriculture biotechnology. Some farm organizations are now calling for farmers not to plant seed of biotechnology-derived cultivars. They say the prospects for the loss of international markets are too great of a risk.

State and Local Governments

APHIS actively sought the cooperation and compliance of state and local governments that were expressing an interest in creating competing regulatory mechanisms under their own jurisdiction. Concerns that a "patchwork quilt" of regulatory requirements could emerge no doubt motivated APHIS to give extra attention to these governmental units. This was done through listening sessions, workshops, and shared documentation. APHIS contained earlier proposals to create subfederal-level regulatory requirements for biotechnology. Had this conflict not been resolved, the pace of U.S. biotechnology commercialization would undoubtedly have been slowed.

THE ROAD MAP

APHIS co-sponsored with several other federal regulatory agencies a forum for discussions on the public and private sector's development of a road map to bring the products of biotechnology forward to commerce. This discussion proved enlightening, as it soon became clear that the conflicting requirements of the various agencies were going to slow the process of product commercialization. Out of these discussions came interagency agreements for better coordination of the product commercialization process.

A particularly difficult issue for APHIS was the process of no longer regulating the products of agricultural biotechnology as they entered commercialization. This was approached by APHIS in several ways, including the eventual certification that a trait was not a plant pest (51). This certification allowed an owner to take a product to commerce, certified by APHIS as an unregulated article.

APHIS also made the determination that it was the genetic trait, not the specific cultivar, that would be certified as not a plant pest. This became important to plant-breeding programs. If a transformation trait could be subsequently moved by conventional plant-breeding methods, once certified by APHIS as not a plant pest, it would not require another round of permit applications and certification. That is, if a recombinant potato plant carrying the Bt endotoxin were to be subsequently crossed by conventional plant-breeding methods to another potato cultivator, the progeny would not be an APHIS-regulated article. This opened significant opportunity to exploit plant genetic engineering, relatively free from federal regulatory oversight, using conventional plant-breeding methods.

A RETROSPECTIVE VIEW

During the first 15 years of APHIS regulating the products of biotechnology, the following significant outcomes have occurred:

- Plant genetic engineering in both the public and private sector is now mostly conducted under a notification process that represents minimal burden to the scientific community.
- More than 3000 field tests have been conducted without a significant incident.
- Thirty commercial products of agricultural biotechnology research are now a market reality. These include genetically engineered corn, soybeans, potatoes, tomatoes, squash, cotton, tobacco, and papaya, which carry in various combinations resistance to plant viruses, tolerance to stress, herbicide tolerance, improved product quality, insect resistance, and male sterility (for making hybrid seed).

Meanwhile ABRAC was decommissioned and the Office of Agricultural Biotechnology no longer exists (52,53). The National Biological Impact Assessment Program's functions have been distributed within the USDA, although a modest special research grant to Virginia Tech continues to provide biotechnology information services to the public and private sector scientific communities. The NASULGC Committee on Biotechnology was disbanded in 1997, after a 15-year history of institutional services and information exchange (54).

Two distinct areas remain in need of resolution. First, the Federal Coordinated Framework's approach to regulating the products (not the process) of biotechnology

has left a regulatory incongruency that will not be easily resolved with our European trading partners. The European Community (EC) has shown a preference for regulating the process of biotechnology apart from other regulatory requirement for the products of commerce. This distinction rose to white-hot intensity in the early 1990s when our European trading partners wanted to establish a "fourth criterion" for registering the products of biotechnology. Heretofore the three standard criteria for registering a product of commerce were safety, efficacy, and quality. Europeans at the time were proposing the fourth criterion that would answer the question, "Do we need it?" In the United States there is no such regulatory authority, since the marketplace is expected to determine whether or not the product has market value or meets a social need. The EC, on the other hand, through its fourth criterion, was proposing that a social-need standard should be imposed on the products of biotechnology. This proposal still seems to be floating in the gap between the U.S. approach to regulating biotechnology and that of our European trading partners.

Second, U.S. policy has established that the products of biotechnology will not be required to be labeled unless the product has a distinctly identified allergen, such as a peanut protein. Our European trading partners, on the other hand, are approaching a consensus that all the products of biotechnology must be labeled. This represents a major challenge for U.S. agriculture production inasmuch as many of our grains and oil seeds are commonly blended from multiple sources prior to shipping to Europe. Maintaining product identity seems, at least to the export/shipping industry, not to be feasible. Labeling everything as "may contain recombinant DNA" is simply uninformative.

LESSONS LEARNED

In the years that have passed since the development of the Federal Coordinated Framework a few biotechnology regulation-lessons have been learned. We now know that:

- Assigning regulatory functions based on existing authorities (as was done by the Federal Coordinated Framework) probably hastened the commercialization of the products of biotechnology. This is the conclusion drawn in a study for the OECD that compared the progress of agricultural biotechnology in those countries with a specifically implemented biotechnology regulatory authority. In these countries (e.g., Germany, the Netherlands, and Belgium) commercialization of the products of agricultural biotechnology seems to be going slower (55).
- Divergent views on the safety of biotechnology had to be expressed through public dialogue, and NABC met this need very well.
- Communication of perspectives was essential for establishing regulatory positions and aligning public policies that allowed safe field testing with recombinant organisms.
- The experience with ABRAC/OAB indicates that the success of the NIH–RAC was not transferable for biosafety assessments of field testing protocols for recombinant organisms.
- It would seem that the designation of lead regulatory agencies is necessary to avoid multiple reviews, which can become inefficient and unworkable.
- Research projects focused on biosafety questions can fill information gaps and thus help to resolve otherwise contentious public health and environmental protection issues.

What remains to be determined is the level of consumer acceptance of the products of biotechnology (56). Is there a linkage between the type of regulatory approval that is used to provide biosafety assurance and the level of consumer product acceptance? This question needs to be resolved.

BIBLIOGRAPHY

1. U.S. Department of Agriculture, Animal and Plant Health Inspection Service: *www.aphis.usda.gov/biotech/OECD/ustegs.htm#usdalaw*
2. S.N. Cohen et al., *Proc. Nat. Acad. Sci.* **70**(11), 3240–3244 (1973).
3. E.F. Smith and C.O. Townsend, *Science* **25**, 671–673 (1907).
4. M.D. Chilton et al., *Cell* **11**, 263–271 (1977).
5. P. Berg et al., *Science* **185**, 303 (1974).
6. M.N. Oxman, in A. Hellman, M.N. Oxman, and R. Pollack, eds., *Proc. Biohazards in Biological Research*, (Asilomar I), Cold Spring Harbor Laboratory, Pacific Grove, CA, 1973.
7. P. Berg et al., *Science* **188**, 991 (1975).
8. National Institutes of Health, *Fed. Reg.*, No. 48, p. 16459, April 15 (1983).
9. National Institutes of Health, 39 Fed. Reg., 39306 (November 6, 1974).
10. National Institutes of Health, 41 Fed. Reg., 27902–27943 (July 7, 1976).
11. National Institutes of Health, 41 Fed. Reg., 38426–38483 (September 9, 1976).
12. D.S. Fredrickson, *Recomb. DNA Techn. Bull.* **2**(2), 87–90 (1979).
13. *Genetic Engineering of Plants: Agricultural Research Opportunities and Policy Concerns*, National Academy Press, Washington, DC, 1984.
14. S.A. Tolin, *Recomb. DNA Techn. Bull.* **4**(4), 156–159 (1981).
15. E. Milewski and S.A. Tolin, *Recomb. DNA Techn. Bull.* **7**(3), 114–124 (1984).
16. E. Milewski, *Recomb. DNA Techn. Bull.* **8**(3), 102–108 (1985).
17. E. Milewski, in J.R. Fowle, III, ed., *Application of Biotechnology: Environmental and Policy Issues*, Westview Press Boulder, CO, 1987, pp. 55–90.
18. U.S. Congress, House of Representatives, Committee on Science and Technology, Subcommittee on Investigations and Oversight, June 22, Hearing No. 98/36, *The Environmental Implications of Genetic Engineering*, U.S. Government Printing Office, Washington, DC, 1983.
19. Office of Science and Technology Policy and Executive Office of the President, 49 Fed. Reg., 50856–50907 (December 31, 1984).
20. Office of Science and Technology Policy and Executive Office of the President, 50 Fed. Reg., 47174–47195 (November 14, 1985).

21. H.I. Miller, in D. Brauer, ed., *Biotechnology*, VCH, Weinheim, 1995.
22. Office of Science and Technology Policy and Executive Office of the President, 51 Fed. Reg., 23302–23350 (June 26, 1986).
23. 57 Fed. Reg., 6753–6762 (1992).
24. U.S. Department of Agriculture, 50 Fed. Reg., 29367–29368 (July 19, 1985).
25. U.S. Department of Agriculture Animal and Plant Health Inspection Service, 52 Fed. Reg., 22892–22915 (1987).
26. *Guidelines recommended to USDA by the ABRAC*, USDA Doc. No. 91-04, March (1992).
27. U.S. General Accounting Office, Report to the House Committee on Science and Technology, GAO/RCED-86-59, GAO, Washington, DC, March, 1986.
28. *Minutes of the Agricultural Biotechnology Research Advisory Committee*, USDA Doc. No. 91-01, pp. 25–30 (1991).
29. J.F. Fulkerson, *BioSci.* **37**(3), 187–189 (1987).
30. U.S. Congress, House of Representatives, Committee on Agriculture and Subcommittee on Department Operations Research and Foreign Agriculture, *Review of Current and Proposed Agricultural Biotechnology Regulatory Authority and the Omnibus Biotechnology Act of 1990*, Hearing 101/75, U.S. Government Printing Office, Washington, DC, October 2, 1990.
31. G. Simon, *Science* **252**, 629–630 (1991).
32. J. Doyle, *Altered Harvest: Agriculture, Genetics, and the Fate of the World's Food Supply*, Viking Penguin, New York, 1985.
33. *Crossfire*, with Tom Braden and Bob Novak, John Fulkerson, USDA, versus Jeremy Rifkin, FET. CNN-TV, April 20, Washington, DC, 1987.
34. I. Rabino, *Sci. Technol. Hum. Values* **16**, 70–87 (1991).
35. Committee on Biotechnology, Division of Agriculture and NASULGC, National Association of State Universities and Land-Grant Colleges, Progress Rep. 5, November (1986).
36. Committee on Biotechnology, Division of Agriculture and NASULGC, National Association of State Universities and Land-Grant Colleges, Progress Rep. 6, November (1987).
37. J. Fessenden Mac Donald, ed., *Agricultural Biotechnology: A Conversation about Risk*, Rep. 5, National Agricultural Biotechnology Council, Ithaca, NY, 1993.
38. *Field Testing Genetically Modified Organisms: Framework for Decisions*, National Academy Press, Washington, DC, 1989.
39. *Introduction of Recombinant DNA-Engineered Organisms into the Environment: Key Issues*, National Academy Press, Washington, DC, 1987.
40. S.A. Tolin and A.K. Vidaver, *Annu. Rev. Phytopathol.* **27**, 551–581 (1989).
41. D. Kennedy, *MBL Sci.* (Winter), 5–9 (1988, 1989).
42. C. Arntzen, *Science* **257**, 1327 (1992).
43. J.M. Tiedje et al., *Ecol.* **70**(2), 297–315 (1989).
44. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology — Field-Testing Engineered Organisms: Genetic and Ecological Issues*, OTA-BA-350, U.S. Government Printing Office, Washington, DC, 1998.
45. National Research Council, *Risk Assessment in the Federal Government: Managing the Process*, National Academy Press, Washington, DC, 1983.
46. H.I. Miller et al., *Science* **250**, 490–491 (1990).
47. H.I. Miller, *Science* **252**, 1599–1600 (1991).
48. H.I. Miller, *Policy Controversy in Biotechnology: An Insider's View*, R.G. Landes Co., Austin, TX, 1997.
49. V.T. Covello and J.R. Fiksel, Rep. NSF/PRA 8502286, National Science Foundation, Washington, DC, 1985.
50. M. Ratner, *Bio/Technology* **8**, 196–198 (1990).
51. U.S. Department of Agriculture Animal and Plant Health Inspection Service, 58 Fed. Reg., 17044–17059 (1993).
52. Office of Agricultural Biotechnology and USDA, Mission Accomplished: Final Report of the Office of Agricultural Biotechnology, Washington, DC, 1996.
53. J. Kaiser, *Science* **270**, 1911 (1995).
54. Committee on Biotechnology, Division of Agriculture and NASULGC, National Association of State Universities and Land-Grant Colleges. Progress Rep. 15 (November 1996). Available at: *www.nbiap.vt.edu*
55. Lisa Zannoni, Personal Communication, 1998.
56. S. Kilman and H. Cooper, *Wall Street J.* May 11, A1, A10 (1999).

See other entries AGRICULTURAL BIOTECHNOLOGY; ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE.

AGRICULTURAL BIOTECHNOLOGY, POLICY, AND NUTRITION

MARION NESTLE
New York University
New York, New York

OUTLINE

INTRODUCTION

Despite the widely recognized potential of agricultural biotechnology to improve the quantity and nutritional quality of the world's food supply, the first commercial products elicited extraordinary levels of controversy over issues of human and environmental safety as well as social values. This article examines the reasons why so potentially useful an application of molecular techniques to improving the nutritional status of populations has proved so controversial in the United States, Europe, and elsewhere. Focusing on the situation in the United States, it reviews key policy issues related to nutrition — economics, marketing, risk, and regulation — that have affected acceptance of the first genetically engineered food products. It reviews evidence indicating that early public acceptance of genetically modified foods was product-specific; people were willing to accept products believed beneficial, safe, and consistent with personal values. Because the failure to label genetically modified foods undermined public trust in industry as well as in government, the chapter addresses implications of the present controversy for future product development, industry actions, and public policies.

FOOD BIOTECHNOLOGY: PROMISE OR REALITY

Food biotechnology — the use of recombinant deoxyribonucleic acid (rDNA) and cell fusion techniques ("genetic engineering") to confer selected characteristics upon food plants, animals, and microorganisms — is well understood as a means to increase agricultural productivity, especially in the developing world. The great *theoretical* promise of biotechnology is that it will help solve world food problems by creating a more abundant and more nutritious food supply (1). Despite this widely recognized and undisputed promise, food biotechnology has elicited extraordinary levels of controversy. In the United States and especially in Europe, the first commercial food products of genetic engineering have been greeted with suspicion by the public, vilified by the press, confonted with boycotts and legislative prohibitions, and threatened with trade barriers. Such reactions reflect widespread concerns about the safety of the products, as well as about their economic impact, environmental effects, ethical implications, and social value (2). The reactions also reflect public fears of the unknown dangers of genetic engineering, along with deep distrust of the biotechnology industry and its governmental regulators. Biotechnology industry leaders and their supporters, however, have tended to dismiss such concerns as antiscientific and irrational, to consider "biotechnophobia" as the single most serious threat to research and commercialization, and to identify as their most important challenge the need to convince the public that the products are safe as well as beneficial (3).

The divergent viewpoints derive directly from the conflict between the two fundamental goals of the food biotechnology industry: to benefit humanity by developing agricultural products that will improve the nutritional status of populations, and to benefit the industry itself through successful marketing of products. Although the new products might well be expected to meet both goals, such is not always the case. The lack of a viable market constitutes a major barrier to research on food problems of the developing world, and the industry's need for rapid returns on investment drives virtually all decisions related to research and development. Indeed, financial imperatives have caused industry leaders to view legitimate public questions about the use, safety, or social consequences of particular food products as threats to the entire biotechnology enterprise and to make defensive marketing decisions that have only undermined public trust.

Theoretical Potential

There seems little doubt that biotechnology holds great promise for addressing world food and nutrition problems, most notably the overall shortfall in food production now expected early in the twenty-first century. No theoretical barriers impede the use of the techniques of molecular and cellular genetics to improve the quantity and quality of the food supply, increase food and environmental safety, and reduce food costs. Table 1 lists the wide range of potentially beneficial applications of food biotechnology now under investigation or theoretically possible. Such applications

Table 1. Theoretical and Current Applications of Food Biotechnology

- Improve the nutrient content, flavor, texture, or freshness of fruit and vegetables
- Increase levels of vitamins, protein, and other nutrients in plant food crops
- Modify seed storage proteins to increase concentrations of limiting amino acids
- Reduce saturated fatty acids in plant seed oils
- Increase plant production of specialty chemicals such as sugars, waxes, phytooxidants, or pharmaceutically active chemicals
- Enable fruits and vegetables to remain fresh during processing, transport, and storage
- Decrease levels of caffeine or other undesirable substances in plant food crops
- Increase resistance of crops to damage by insect, fungal, or microbial pests
- Increase resistance of crops to "stress" by frost, heat, salt, or heavy metals
- Develop herbicide-resistant plants to improve weed control
- Enable crop plants to be grown under conditions of low input of fertilizers, pesticides, or water
- Enable major crop plants to fix atmospheric nitrogen
- Develop plant foods that contain "vaccine" antigens
- Increase the efficiency of growth and reproduction of food-producing animals
- Create disease-resistant animals
- Develop animal veterinary vaccines and diagnostic tests
- Enable cows to produce milk containing human milk proteins
- Alter mosquitoes so they prefer animal blood to human, or convey vaccines
- Create microorganisms, enzymes, and other biological products useful in food processing
- Develop microorganisms capable of converting environmental waste products — plastics, oil, pesticides, or PCBs — into usable animal feeds

could well increase world food production, especially given the conditions of climate, soil, and environmental degradation characteristic of many developing countries. The potential benefits constitute the principal basis for industry arguments that biotechnology is not only the most important scientific tool to affect food production in the history of the world, but the *only* solution to the expanding global needs for food (4). These promises, however, have not yet been fulfilled, nor are they likely to be achieved in the immediate future, not least because many of the listed applications pose technical problems of formidable complexity. The slow progress of biotechnology in addressing world food problems should not imply that such problems cannot be solved. Given sufficient time, commitment, and funding support, technical barriers can be overcome. Whether doing so will help feed the world, however, is a matter of considerable debate.

Economic Realities

Investment Demands. Rather than technical problems, the most important barriers to addressing world food problems derive from the industry's need to recover research costs and maximize returns on investment. The potential returns are enormous; worldwide sales of genetically-modified crops could reach $300 billion by 2010 (5). To date, however, stock market returns have not reflected such projections. Although the food biotechnology industry increased in sales, revenues, and numbers of companies and employees in the 1990s, net losses also increased. The Monsanto company has been a notable exception; its stock prices rose rapidly in the mid-1990s and its agricultural products continue to be highly profitable (4). The generally poor performance of other food biotechnology stocks has been attributed to uneven management, corporate shortsightedness, and product failures. More recently international resistance to genetically modified foods has affected sales and confidence in the industry. Low levels of government investment also have impeded industry growth, as most federal biotechnology funding has supported drug rather than agricultural research. Only recently has the U.S. government begun to support agricultural biotechnology research to any significant extent.

The immediate need for returns on investment requires the industry to focus development efforts on products that are technically feasible and economically productive rather than on those that might be more useful to the public or to developing countries (6). Thus product development efforts concentrate on traits that most benefit agricultural producers and processors: control of weeds, plant diseases, ripening, insects, or herbicide resistance (7). For example, the Monsanto company's research budget (which is more than twice that of all of the public sector tropical research institutes combined) is applied almost entirely to temperate-zone agricultural problems (8). The company's principal agricultural products are soybeans and other plants genetically modified to resist the company's flagship herbicide, "Roundup," and "Bt corn" containing a insect-inhibiting toxin naturally produced by the soil bacterium *Bacillus thuringiensis.* Monsanto began selling "Roundup Ready" soybeans in 1996; by 1998 they had been planted on one-third of U.S. soybean farmland and covered 25 million acres (10.1 million hectares). By 1999 more than 35 percent of U.S. corn and 45 percent of soybean acres were grown from genetically modified seeds, and total worldwide acres devoted to such crops were expected to triple within the next five years (9). Monsanto's research "pipeline" mainly emphasizes "Roundup Ready" crops designed for animal feed, although a high-carotenoid canola oil designed to prevent vitamin A deficiency is a rare nutritionally focused exception (4). Projects to improve the nutritional content of basic food sources are expensive to produce and unlikely to be affordable by the populations they would most benefit. Although Third World agricultural problems and their biotechnological solutions are well defined, and many sources of private and public funding are available to support such projects, the sources are fragmented and poorly coordinated and often favor the priorities of the donors rather than recipients (10). Despite recent advances in cassava biotechnology, for example, nearly all international budgets for such research have been reduced (11). Ultimately biotechnology may well improve the world's food supply but to date it has not done so to any appreciable extent.

Marketing Barriers. To ensure adequate returns on investment, the biotechnology industry must create and sell new products. Because the United States vastly overproduces food, new products must compete in a market that is already highly competitive. In 1997, for example, the U.S. food supply provided an average of 3,800 kcal (18.2 MJ) per day for every man, woman, and child in the country, an increase of 500 kcal (2.0 MJ) per day since 1970 (12); most adults require one-half to two-thirds that amount and children much less. Because the amount of energy that any one person can consume is finite, such overproduction implies that a choice of any one food product will preclude the choice of another, making the food-marketing system extremely competitive. Food marketers compete for consumer purchases through two principal means: advertising and new product development. Retail sales of food and beverages generate about $800 billion annually in the United States, and food marketers spend $11 billion on direct consumer advertising, and about twice that amount on retail promotion. They introduce about 15,000 food products into the marketplace every year (13). Nevertheless, the food-processing sector rarely grows by more than 1 percent a year, a rate considered stagnant by comparison to that of other industries. In so competitive an environment, biotechnology is viewed as a critically important process for developing new products that will increase economic returns.

SAFETY ISSUES

From their inception, gene cloning experiments elicited safety concerns, mainly focused on the potential hazards of releasing new organisms with unknown properties into the environment. At a conference in 1975, scientists suggested stringent guidelines for research studies employing

Table 2. Safety Issues Raised by Food Biotechnology

- Adverse changes in the composition, absorption, or metabolism of key nutrients
- Unanticipated health effects resulting from genetic changes
- Increases in levels of naturally occurring toxins or allergens
- Activation of dormant toxins or allergens
- Introduction of known or new toxins, allergens, or antinutrients
- Induction of resistance to useful antibiotics through use of antibiotic marker genes
- Adverse environmental effects on wildlife and ecosystems
- Adverse changes in the nutrient content of animal feed
- Increased levels of toxins in plant byproducts fed to animals

rDNA techniques. The following year, the National Institutes of Health required researchers to follow similar guidelines. In subsequent years, as understanding of the techniques improved, concerns about safety diminished and the guidelines were modified accordingly. From the standpoint of the biotechnology industry and its supporters, genetically engineered foods are no different from foods produced by conventional genetic crosses; If they induce any risks at all, these are small and greatly outweighed by benefits. Nevertheless, the common genetic techniques for modifying foods, especially those involving bacteria that cause plant diseases (e.g., crown gall), antibiotics as part of the selection process, and genes from one living species inserted into another, continue to elicit debate. Table 2 summarizes the principal safety issues raised by the use of food biotechnology. Although most such concerns remain theoretical, some that could affect human nutritional status and health have a limited basis in observation or experiment, as discussed below.

Unintended Consequences: Tryptophan Supplements

Critics of food biotechnology insist that without prior experience, the techniques raise safety concerns that are difficult to define, predict, or quantify. As an example, they point to the demonstrable hazards of genetically engineered nutritional supplements of the amino acid tryptophan. Tryptophan is a normal constitutent of all body proteins that is sometimes sold as self-medication for insomnia and other conditions. In 1989 health officials linked tryptophan supplements from a single manufacturer to *eosinophilia-myalgia syndrome*, an unusual syndrome of muscle pain, weakness, and increased blood levels of certain white blood cells (eosinophils). Eventually, more than 1500 cases of illness and nearly 40 deaths were attributed to the supplements. Because tryptophan is an essential component of body proteins, investigators believed that the amino acid itself could not have caused harm, but that toxic contaminants must have developed during the manufacturing process. This process involved genetically modifying a strain of bacteria to produce unusually high levels of tryptophan, and then concentrating, collecting, and purifying the amino acid. To date, the toxin remains incompletely characterized. Although the genetic techniques do not appear to be directly at fault, their use in modifying a strain of bacteria created a situation — albeit inadvertently — that favored the formation of toxic products (14). This example suggests that concerns about the unknown hazards of biotechnology cannot be dismissed out of hand.

Allergenicity

Because genes encode proteins, and proteins are allergenic, the introduction of allergenic proteins into previously nonallergenic foods could be another unintended consequence of plant biotechnology (15). In support of this idea, a biotechnology company transferred an allergenic protein from Brazil nuts to soybeans, and researchers confirmed that people who are allergic to Brazil nuts react similarly to soybeans containing the transgenic Brazil-nut protein (16). The company had developed the Brazil-nut soybeans as a means to increase the content of methionine, a sulfur-containing amino acid, in poultry feeds. Feathers contain high levels of sulfur-containing amino acids, and poultry feeds must be supplemented with methionine — at additional cost — to promote optimal growth. Because the Brazil-nut protein is especially rich in methionine, its gene was a logical choice as donor. Nuts, however, are often allergenic; the researchers happened to have collected serum samples from people known to be allergic to Brazil nuts. Thus they had in place all components necessary to test for allergies to Brazil-nut proteins, a situation that is rarely the case for other food allergens.

True allergies to food proteins can be documented in less than 2 percent of the adult population, but many more people might be expected to develop food sensitivities as proteins are increasingly added to commercially-prepared foods. Soy proteins, for example, already are very widely used in processed foods, and genetically modified soy ingredients already are widely prevalent in the food supply (17). Most biotechnology companies use microorganisms rather than food plants as gene donors, however, and their proteins do not appear to share sequence similarities with known food allergens. Few have as yet entered the food supply, but their allergenic potential is uncertain, unpredictable, and untestable (18).

As discussed below, allergenicity raises complex regulatory issues. Under a policy developed by the U.S. Food and Drug Administration (FDA) in 1992, company scientists were required to — and did — consult agency staff about the need for premarket testing. Because testing demonstrated transmission of the allergenic protein, the company would have had to label its soybeans as genetically modified. Because the company could not guarantee that people would not eat soybeans intended for animal feed, it wisely withdrew the transgenic soybeans from the market. Supporters of the FDA policy interpreted these events as a clear demonstration of the policy's effectiveness. Others, however, argued that the policy failed to protect the public against less well-studied transgenic allergens to which they might be sensitive and therefore favored industry. Critics were especially concerned about the lack of a requirement for labeling, as avoidance is often the only effective way to prevent allergic reactions. In 1993 the FDA requested public comment on whether and how to label food allergens in transgenic foods and later proposed rules to help resolve safety issues

related to allergenicity, but by mid-2000 had taken no action on the matter.

Antibiotic Resistance

Genes for antibiotic resistance are used as part of selection processes. They can be expressed in transgenic plants when under the control of genes taken from higher organisms (but not bacteria), raising the possibility that they could be transferred to bacteria in the human intestine. Most experts consider this possibility remote but not entirely impossible, and antibiotic resistance is a principal concern of critics of genetically modified foods, especially in Europe. To avoid this possibility, the FDA suggests that companies evaluate the risks of transferring resistance to the antibiotics they are using, avoid using antibiotics effective against human diseases, and especially avoid using antibiotics uniquely effective against certain conditions (e.g., vancomycin and staphylococcal infections) (19).

REGULATORY ISSUES

Current debates about the regulation of food biotechnology center on the conflict between issues of safety on the one hand and a broad range of ecological, ethical, and societal issues on the other (20). For the industry and its supporters, safety is the only issue of relevance; because science supports the safety of most genetically engineered products, unnecessarily restrictive regulations appear to create barriers to further research and economic growth. Critics, however, view regulations as needed to protect the public not only against known safety risks but also against those that cannot yet be anticipated. They view safety as only one component of a far broader range of concerns about the impact of biotechnology on individuals, society, and the environment — issues that might also demand regulatory intervention. For government officials, biotechnology regulation must find the proper balance between oversight and encouragement of industry efforts to develop and market new food products (21). Current U.S. regulatory policies affect three aspects of food biotechnology directly — food safety, environmental protection, and intellectual property rights — and affect international trade indirectly. Thus far, these policies have achieved a balance that neither satisfies industry nor consumer groups.

Food Safety

In 1986 the U.S. White House developed a "Coordinated Framework" for regulating biotechnology based on the premise that the techniques led to products no different from those developed through conventional genetic methods. Thus existing laws and agencies should be sufficient for regulatory purposes. At the time regulation of food biotechnology involved no less than 3 offices reporting directly to the president; 4 major federal agencies; 8 centers, services, offices, or programs within agencies; and 5 federal committees — all operating under the authority of 10 distinct Acts of Congress (1). As might be expected, critics identified obvious flaws in this regulatory framework, among them lack of coordination, duplication of effort, overlapping responsibility, and gaps in oversight. The principal laws affecting food safety preceded the use of genetic engineering, however, and did not necessarily apply to the new methods.

This uncertain regulatory status caused the food biotechnology industry to demand more precise guidance from FDA. In response, FDA developed a formal policy for the regulation of genetically-modified plant foods (22). The policy presumed that foods produced through recombinant techniques raised no new safety or nutritional issues and therefore could be regulated by FDA's existing policies for foods considered Generally Recognized as Safe (GRAS). Instead, safety evaluation would focus on changes in the "objective characteristics" of foods — changes in nutrient composition or new substances, toxins, or allergens. FDA would invoke requirements for premarket safety evaluation, premarket approval, or labeling only when those characteristics were sufficiently altered. The biotechnology industry welcomed this policy as a strong incentive for investors, but consumer groups judged it inadequate not least because the foods would not be labeled. As early as 1992 it became evident that consumer choice in the marketplace would influence acceptance of genetically modified foods (23), and a federal study recommended a review of the entire regulatory framework in order to establish a more equitable balance between promotion of industry and protection of the public (21). By late 1994 the FDA had approved the marketing of tomatoes genetically altered to reach optimal ripening after harvest, milk from cows treated with recombinant growth hormone, virus-resistant squash, insect-resistant potatoes, and herbicide-resistant cotton (used to make seed oil for animal feed), and soybeans, none of which addressed nutritional characteristics directly.

Environmental Impact

The "Coordinated Framework" affirmed that the U.S. Department of Agriculture (USDA) and the Environmental Protection Agency (EPA) were the primary agencies for regulating agricultural biotechnology. The EPA was to regulate recombinant plants developed to control insects and other pests, and it did so by requiring biotechnology companies to obtain permits prior to the manufacture or release of their agricultural products. The EPA policies were designed to address concerns that widespread agricultural use of new kinds of living species might present direct risks to human health — risks generally agreed to be minimal. Instead, environmentalists were concerned that transgenic crop plantings might pose *ecological* risks — displace existing plants and animals, create new plant pathogens, disrupt ecosystems, or reduce crop diversity. They predicted that widespread use of genetically modified crops such as those containing the gene for Bt toxin might undermine ongoing efforts to promote sustainable agricultural practices by selecting for Bt-resistance (24). They also argued that increased planting of herbicide-resistant crops would increase reliance on toxic chemicals to manage pests (25). In 1994 the EPA proposed to extend pesticide laws to transgenic "plant-pesticides" such as Bt and to require their registration, meaning that manufacturers would have to conduct tests

of their nutritional and ecological impact as well as label them (26). Despite the EPA's assurance that the regulations would resolve uncertainties and attract investors, the rules appeared to favor large, established companies but discourage small, innovative companies.

Environmentalists were concerned that the proposals did not place enough emphasis on crops designed to resist chemical herbicides. Their ecological concerns have been encouraged by subsequent observations and research. Since 1996 researchers have reported preliminary signs of Bt-resistance in cotton plants and among moths and tobacco budworms, transmission of herbicide resistance from oilseed rape (canola) to related weeds, and higher mortality rates among bees fed proteins isolated from genetically modified rapeseed. Most famously, monarch butterfly larvae consuming pollen from Bt corn were reported to grow more slowly and die more quickly than larvae not exposed to such pollen (27). Although these observations are preliminary and require confirmation, they suggest a rational basis for environmentalists' fears that weeds and insects can develop resistance to currently available control methods, transgenic toxins can kill "friendly" insects such as bees and monarch butterflies, and genetically modified pollen can cross-fertilize conventional and organic crop plants. By late 1999 the EPA had not yet issued final rules on its plant-pesticide proposals.

Intellectual Property Rights

The U.S. intellectual property laws grant rights to patent owners to exclude everyone else from making, using, or selling the protected product for at least 17 years. Patents were first granted for plant varieties developed through asexual propagation in 1930. In 1970 Congress extended these rights to new varieties of plants developed through traditional genetic methods of sexual propagation. In 1980 the Supreme Court granted patent rights to microorganisms developed through recombinant techniques, and the Patent Office issued the first patent for such an organism. Patent rights were further extended to transgenic plants in 1985 and to animals in 1988 (28). The patenting of transgenic microorganisms and plants provided a major incentive for the growth of the food biotechnology industry. Within just a few years, however, industry and government officials in the United States, Canada, and Europe began challenging patent awards. By 1995, however, the U.S. Patent Office had issued 112 patents for genetically engineered plants. Among these were exclusive patent rights to one company for all forms of bioengineered cotton and to another for all uses of "antisense" genes such as those used to create tomatoes with a long shelf life (discussed below). The breadth of such patents seemed excessive and various groups soon filed lawsuits. Patent issues are especially pressing for Monsanto, as its U.S. patent for the Roundup herbicide expires in 2000.

The patenting of animals has generated even greater debate, particularly from animal-rights organizations and other groups who believe that the genetic engineering of farm animals might adversely affect family farmers, be cruel to animals, and endanger other living species.

Table 3. Principal Arguments for and Against the Patenting of Transgenic Animals

Arguments in favor
• Patent laws regulate inventiveness, not commercial uses
• Patenting is an incentive to research and development
• Patenting enables the biotechnology industry to compete in international markets
• Patenting is preferable to trade secrets
• Patenting rewards innovation and entrepreneurship
Arguments opposed
• Metaphysical and theological considerations make patenting untenable
• Patenting involves inappropriate treatment of animals
• Patenting reflects inappropriate human control over animal life
• Patenting disturbs the sanctity and dignity of life
• Most other countries do not permit patenting of animals
• Patenting could cause adverse economic effects on developing countries
• Patenting promotes environmentally unsound policies
• Animal patents will increase costs to consumers and producers
• Animal patents will result in further concentration in agricultural production
• Patent holders will derive unfair benefits from royalties on succeeding generations of patented animals

The successful cloning of a sheep ("Dolly") in 1997 (29) only heightened concerns about the ethics of "tampering" with animal life. The principal arguments for and against the patenting of genetically engineered animals are summarized in Table 3. Perhaps in response to concerns about such issues, the Patent Office ceased issuing patents for transgenic animals in 1988. In 1993 it resumed processing of the nearly 200 animal patent applications that had accumulated during that "self-imposed moratorium," but fewer companies were attempting to patent farm animals by that time, largely because persistent technical problems and costs had encouraged them to shift to more profitable areas of research.

PUBLIC PERCEPTIONS

Because food is overproduced, the food industry is so competitive, and overseas sales of American products are so important to the economic viability of agricultural producers, biotechnology companies have long viewed consumer acceptance of genetically modified products as critical to the industry. Thus various agencies in the United States and many other countries have conducted surveys of public perceptions of food biotechnology. In the United States these surveys cover a 15-year period. Despite differences in methods, year, and subjects, they have produced remarkably consistent information over time, and they reveal an internally consistent logic of considerable predictive value. They also explain why the responses to genetically modified foods in the United States have differed so substantially from those in Europe. Although U.S. survey

respondents have only a limited understanding of science and technology, they hold high expectations that food biotechnology will produce benefits for them and for society as a whole. Surveys find respondents to be concerned about the potential and unknown dangers of genetically modified foods, but believe that the benefits outweigh risks. For example, 75 percent answer yes to the question, Do you feel that biotechnology will provide benefits for you or your family within the next five years? The surveys indicate clearly that respondents prefer some transgenic food products to others, most favoring products that appear beneficial to health or society, save money or time, are safe, or improve the environment. For example, 77 percent of U.S. respondents say they would be likely to buy genetically modified foods that protected against insect damage or required fewer pesticide applications (30).

Safety considerations, although often the focus of biotechnology debates, do not emerge in these surveys as the most important public concern. Instead, respondents appear most troubled by *ethical* issues related to food biotechnology. They are more willing to accept genetically modified foods that involve plants rather than animals, that do not harm animals, and that do not involve the transfer of animal genes into plants (31). These views derive from value systems that encompass issues that extend beyond food safety to include fundamental social, cultural, and religious beliefs. The surveys also reveal substantial public distrust of government credibility in safety matters, its ability to regulate food biotechnology appropriately, and the ability of the biotechnology industry to make decisions in the public interest. Perhaps for these reasons, surveys invariably find a large majority of respondents to want genetically modified foods to be labeled. Demands for labeling of genetically modified foods are especially prominent in Europe where nearly everyone (96 percent in Great Britain) favors such action (32).

If industry leaders view public opinion as irrational and as evidence for the need to educate consumers about the safety and benefits of biotechnology, they are missing the most strikingly useful conclusion to be drawn from these surveys. Prior to 1996, consumer attitudes toward food biotechnology in the United States and in Europe were largely product specific. People were willing to accept genetically modified products perceived as valuable to public health and welfare (33). From the surveys it should have been evident to industry that most consumers would accept genetically modified products if they were demonstrably beneficial to the public as well as to the industry's economic interests. The failure of genetically modified foods to address needs perceived by the public as important, and the industry's refusal to label the products, explain much of the subsequent resistance to the products, especially in Europe.

PREDICTIVE IMPLICATIONS: RECOMBINANT PRODUCTS

For the first recombinant products approved for sale in the United States, survey results suggested an analytical framework — based on the value of a product, its safety, and its ethical value — for predicting the degree of difficulty a product might experience with public acceptance. Table 4 outlines the questions that comprise this framework. The more positive the answers, the more likely consumers were to accept the product. To the extent that the answers were negative or equivocal, consumer resistance was likely to increase. This framework predicted the degree of acceptance of products released through 1995. Beginning with the release of recombinant soybeans and corn in 1996, however, policies regarding labeling, food safety, and international trade also influenced public acceptance.

Table 4. Analytical Framework for Predicting Public Acceptance of a Genetically Modified Food Product

1. Is the food safe for people and for the environment?
2. Is the food valuable? Will it:
• Increase nutrient content?
• Increase food availability?
• Decrease food cost?
• Improve food taste?
• Grow better under difficult conditions?
• Reduce use of herbicides and pesticides?
3. Is the food "ethical"? Does it avoid:
• Harm to animals?
• Insertion of animal genes into plants?
• Economic harm to small farms or businesses?
• Economic harm to populations in developing countries?

Pharmaceuticals: Insulin

By the early 1990s the FDA had approved at least 15 recombinant drugs for use in human subjects. Recombinant insulin, for example, received approval in 1982 and was of unquestionable utility (6). It solved problems of scarcity and quality; it could be produced in unlimited amounts, and its amino acid structure was identical to that of human insulin and therefore superior to insulin obtained from the pancreas of pigs or cows. It was safe and raised no ethical considerations. Recombinant insulin readily met all three criteria for consumer acceptance, and it is neither surprising nor inconsistent that it and other recombinant drugs were accepted without protest in the United States as well as in Europe.

Enzymes: Chymosin

Recombinant enzymes used in food manufacture also were readily accepted. Chymosin, an enzyme used to coagulate milk to make cheese, was traditionally extracted from the stomachs of calves and sold as part of a mixture called rennet. It was difficult to extract, varied in quality, and was scarce and expensive. Through genetic techniques, the gene for chymosin could be transferred to bacteria that produced the enzyme in large quantities. The recombinant enzyme was approved for food use in 1990 (6). This action elicited no noticeable complaints from biotechnology critics, perhaps because the manufacturer did not publicize the enzyme's recombinant origins but also because obtaining the nontransgenic enzyme required the slaughter of baby calves. Transgenic chymosin also met

the three criteria for consumer acceptance: It was more useful, ethical, and just as safe as the enzyme it replaced.

Hormones: Bovine Somatotropin (rbST)

The history of recombinant bovine somatotropin (rbST), the first product to be approved by the FDA under its 1992 food biotechnology policy, best illustrates how issues of safety, societal benefit, and ethics contribute to consumer resistance. The product, a growth hormone that increases milk production in cows by at least 10 to 20 percent, elicited considerable debate in the United States in the mid-1990s and, more recently, in Canada and Europe. Its name reflects the controversy: proponents generally use the scientific name, rbST, whereas critics call it Bovine Growth Hormone (rBGH). For purposes of consistency, this article uses rbST. The Monsanto company developed rbST in the mid-1980s and promoted it as a method to increase the efficiency of dairy farming. Although such efficiency would seem to be of great benefit to consumers, critics soon raised questions about the product's effects on human health, animal welfare, and the economic viability of small dairy farms (6). They were especially concerned that rbST-treated milk would not be labeled as such. When the FDA approved Monsanto's rbST as a new animal drug in 1993, it ruled that labeling would be misleading as treated and untreated milk were indistinguishable by available methods. The level of protest against rbST was extraordinary; Supermarket chains announced that they would not carry milk from rbST-treated cows, and several states enacted legislation banning the hormone. Some dairy companies, concerned about consumer reactions, began to label their products as "BGH-free," but industry groups challenged the legality of this practice. The FDA permitted that designation to be used if accompanied by a disclaimer: "No significant difference has been shown between milk derived from rbST-treated and non-rbST-treated cows" (34). Protest against rbST could easily have been anticipated, as rbST raised safety, value, and ethics issues addressed by the questions in Table 4.

Safety Issues. Bovine somatotropin stimulates milk production, and the natural hormone is always present in cow's milk in low concentrations. Milk from rbST-treated cows contains both the natural hormone and rbST; these are almost identical. The hormone itself is unlikely to be harmful to humans, even though its concentration is higher in milk from treated cows. Its protein structure differs from that of the human hormone and is not biologically active in people. Like all proteins, the cow hormone is largely broken down in the human intestinal tract. In 1990 Monsanto-sponsored scientists reported that rbST milk was safe for human consumption and that the FDA studies had answered all safety questions. That same year FDA scientists reviewed more than 130 studies of the effects of rbST on cows, rats, and humans and concluded that the hormone did not adversely affect human health. The publication of this last report in a prestigious scientific journal was judged "unprecedented," as it appeared that the FDA was favoring a drug it had not yet approved. However, other expert groups also concluded that milk from rbST-treated cows was essentially the same — and as safe — as milk from untreated cows (35).

Despite this evidence critics continued to raise safety concerns about two factors that might be present in rbST milk: antibiotics and insulinlike growth factor-I (IGF-I). The concern about antibiotics derives from observations that cows treated with rbST develop udder infections (mastitis) more frequently than untreated cows (36). Because the infections are treated with antibiotics that linger in milk and meat, it is theoretically possible that consumed antibiotics could contribute to human antibiotic resistance. Although federal regulations require testing for antibiotic residues in milk, the FDA tests for only a small fraction of animal drugs in common use — just 4 out of 82 in one study — leading to charges that the agency lacks a comprehensive strategy for monitoring such drugs.

The IGF-I issue derives from concerns that the increased concentration of this factor in milk from rbST-treated cows might stimulate premature growth of infants or cancers in adults. Although IGF-I appears to be denatured in infant formulas and seems unlikely to be absorbed in significant amounts by the human digestive tract, the factor is readily absorbed from milk, is biologically active in rats, and is associated in epidemiological studies with increased risk of prostate cancer in men (37) and breast cancer in premenopausal (but not postmenopausal) women (38). Although the clinical significance of these observations is unknown, they have encouraged dairy and consumer groups to demand further examination of the data and to file suit against the FDA; they also have been used as a basis for refusal to license rbST in Canada and Europe.

Value Issues. For many years milk production in the United States exceeded demand, resulting in large surpluses of dairy products. The use of rbST was expected to further increase milk production. Biotechnology companies contended that use of the hormone would reduce farm costs because equivalent amounts of milk could be produced by fewer cows. Although it might seem logical that creation of more surplus milk would lead to lower prices, dairy prices are tightly linked to federal support programs and unlikely to change very much. Given this situation, rbST offered no evident cost benefits to consumers.

Ethical Issues. Because rbST-treatment of cows increases milk production, concerns have been raised about effects of the drug on the health and reproductive ability of animals. The more milk cows produce, the more likely they are to develop mastitis. In addition rbST is delivered through injection and can cause localized reactions at the point of entry. Despite industry assertions that appropriate herd-management practices can minimize such problems, they were reported regularly. In addition, increasing the supply of milk might be expected to accelerate long-standing trends toward the elimination of small dairy farms, and most commentors believed that at least some dairy farmers would be forced out of business. For these reasons, answers to some of the ethical questions are negative or equivocal.

Consumer Reactions. Taken together, public scepticism about rbST related to nearly all of the areas of concern listed in Table 4, suggesting that this product was an unfortunate first choice for commercialization. U.S. farmers already overproduced milk, and rbST offered no evident benefit to consumers in availability, price, or quality. That the product affected milk also was unfortunate, as this food often is promoted as conveying an image of purity. The primary beneficiaries of rbST therefore appeared to be its manufacturers and the large dairy farmers who were best able to exploit its use. Despite considerable resistance from farmers, the company has said that rbST broke even in 1996 and has been profitable ever since, but its annual report does not permit independent verification of this assertion (4). Because the use of rbST to produce commercial milk is not labeled, public acceptance of the hormone in the U.S. remains uncertain. One indicator of public opinion is the spectacular growth in sales of organic ("BGH-free") milk since 1996 (39). In Europe, use of the hormone is prohibited at least until the end of 1999. Although several international committees have reaffirmed the safety of rbST, European Community members have not yet reached consensus on its approval (40). In Canada, applications to market rbST have been pending for more than 15 years, largely because of conflicting opinions about the safety of the products for both cows and people. These events also could have been predicted from the questions in Table 4.

Foods: The "Flavr Savr" Tomato

Americans expect tomatoes to be available on a year-round basis. In 1997 farmers produced nearly 16 pounds (7.3 kg) per capita of fresh tomatoes and another 73 pounds (33.1 kg) for processing (12), but supermarket tomatoes, bred for disease resistance, appearance, and durability, have long been the bane of consumers longing for "backyard" taste and freshness. Beginning in the mid-1980s, Calgene, a California-based biotechnology company, invested $25 million and 8 years of effort to develop a tomato with a reversed (and therefore blocked) gene for ripening that would allow it to be picked and marketed at a more mature stage of taste (6). Calgene expected this "Flavr Savr" tomato to capture at least 15 percent of the market for fresh tomatoes as soon as it became available, and the company planned to sell — and label — it as genetically engineered to taste better. As the first company to develop a genetically modified food, Calgene worked closely with FDA to determine the tomato's regulatory status. FDA insisted that review committees focus exclusively on the tomato's *safety* and judged concerns about ethical issues or labeling as irrelevant. The agency decided in 1994 that all safety and nutritional questions about the new tomato had been resolved and approved its marketing. Although some groups threatened boycotts and "dumpings," most analysts believed that consumers would accept the tomato if its improved taste seemed worth the premium price, initially expected to be twice that of conventional tomatoes. From the answers to the questions in Table 4, some consumer resistance should have be expected. Although the Flavr Savr was as safe and nutritious as market tomatoes, it raised issues related to impact on small growers, and its benefit to the public was restricted to taste. Its higher costs, however, identified the Flavr Savr as a luxury product targeted to an upscale market. To Calgene, the tomato was well worth the huge investment of time, money, and effort as it paved the way for subsequent approval of the company's seed oils and other genetically modified crops. Eventually the Flavr Savr proved impossible to grow and ship in adequate quantities, and its acceptance in the markeplace could not be tested (41).

Food Crops: Soybeans and Corn

Experts predict that within five years virtually *all* of U.S. agricultural exports — worth $50 billion annually — will be transgenic or combined with genetically modified bulk commodities. Any resistance to acceptance of transgenic soybeans and corn would pose a serious economic threat (42). In the United States these products encountered little public opposition, perhaps because they are mainly fed to animals and their environmental hazards seem geographically remote from most people. Alternatively, the lack of protest reflected ignorance of the extent to which genetically modified ingredients pervaded the food supply (17). Also, until quite recently, most press reports about genetically modified crops appeared exclusively in business pages. Europeans, however, could not help but be informed; at the peak of coverage early in 1999, the seven largest British daily newspapers ran nearly 2000 column inches (5,000 cm) of copy about genetically modified food, nearly all of it unfavorable (43).

The controversy in Europe began in 1996 with the first marketing of unlabeled recombinant soybeans and corn. Since then the European Union and various member countries have issued outright bans, prohibitive regulations, or labeling requirements. Food producers and retailers have refused to use genetically modified ingredients in their products and have withdrawn products containing them from sale (9). The intense resistance in Europe, so much more extreme than in the United States, is nevertheless related to a similar set of consumer issues. Many Europeans have long-standing traditions of animal welfare, vegetarianism, and other value systems that might affect attitudes toward food biotechnology, as well as memories of Nazi eugenic experiments during World War II. In addition concerns about the spread of antibiotic resistance, fears generated by the 1996 food safety crisis over "mad cow" disease, and, more recently, alarms raised by outbreaks caused by foodborne pathogens have reduced public trust in government as well as in industry. European consumers, also were reacting to the perceived arrogance of American officials and companies who seemed to be forcing U.S. exports "down their throats" (44). In this context the aggressive marketing of unlabeled genetically modified soybeans and corn by American biotechnology companies only intensified public resistance. An advertising campaign by the Monsanto company emphasizing environmental and nutritional benefits of biotechnology through slogans such as "While

we'd never claim to have solved world hunger at a stroke, biotechnology provides one means to feed the world more effectively," and "Food labelling. It has Monsanto's full backing," produced the opposite of the effect intended and only increased public suspicion of genetically modified foods (45). Also inciting protest was the company's investment in "terminator" technology, a method for ensuring that genetically modified crops could not produce viable seeds. This technology would protect the company's proprietary rights to the seeds but adversely affect Third World farmers who grow 15 to 20 percent of the world's food from saved seed (46). Calls for an international ban on terminator research also brought negative public attention to food biotechnology.

By the late 1990s European surveys revealed nearly unanimous public support for labeling of food products containing genetically modified ingredients. Labeling requires strict segregation of genetically modified from conventional corn or soybeans and the use of genetic tests to distinguish them. Because of cross-pollination, some mixing of genetically modified with conventional field crops is inevitable, and authorities have yet to agree on the lowest level of "contamination" permissible for products to be labeled "GM-free." Labeling also has trade implications. Any differences between the food regulations of one country and another must be "harmonized" by the international Commission (Codex Alimentarius) that sets food standards. For several years, the Commission has considered standards for mandatory labeling of genetically modified foods but has yet to reach consensus (47).

POLICY IMPLICATIONS

Although consumers in the United States have been slow to oppose genetically modified foods, organized opposition appears to be growing. When the USDA proposed that standards for foods designated as "organic" could include those that had been genetically modified, 275,000 people wrote letters of protest and the agency was forced to withdraw the suggestion. Although news articles critical of biotechnology had been published for years in environmentalist and other specialist magazines, mainstream publications have begun to feature the issue and to raise public awareness. Coalitions of consumer groups, religious groups, chefs, and scientists have filed lawsuits and organized petition campaigns to force the FDA to require labeling and safety testing for transgenic products. One consumer organization collected nearly 500,000 signatures on a petition calling for mandatory labeling. By the end of 1999, 68 percent of American adults polled by Gallup wanted genetically modified foods to be labeled even if it meant paying higher prices (48). Feeling the pressure from organic competitors, the Gerber's and Heinz baby food companies have announced that they will not allow genetically modified ingredients in their products, as has the Archer-Daniels-Midland Company. In response to European rejection of genetically modified corn, the American Corn Growers Association advised its members to consider planting only conventional seeds in spring 2000. Labeling of genetically modified products seems likely to occur in the United States as well as in Europe.

In promoting public acceptance of the first genetically modified products, industry leaders focused on safety as the sole basis for discussion and characterized other concerns as unscientific or irrational. The industry's dismissal of concerns other than safety and its opposition to labeling missed an important point: Initial views of food biotechnology were product specific and, as such, were consistent and predictable. If the marketplace was to be allowed to determine the success of genetically modified products — as it does with all others — the products would have to be labeled. If the products were valuable to consumers, the label should have encouraged purchases as well as trust in the industry (18). The failure to label could well have been the single factor most responsible for the hostile European reception to genetically modified soybeans and corn.

The controversy over food biotechnology derives directly from the conflict between the industry's need to be profitable and the desire of consumers for products that are economically and socially valuable, as well as safe. To frame the debate in terms of rational science versus an irrational public is to do a disservice to both. Biotechnology is not inherently dangerous, and it should be capable of doing much good. The public is not inherently irrational and should be capable of judging whether genetically modified products are worth buying. This analysis suggests that the public will continue to be unconvinced that genetically modified foods are necessary or safe as long as the principal beneficiaries of the technology are the companies themselves. The current debates about genetically modified foods offer the industry an opportunity to address consumers' concerns about the credibility, safety, and ethical implications of the products. To improve credibility, the industry must bring its rhetoric more in line with reality. If industry leaders continue to state that food biotechnology is necessary to solve world food problems, they should be investing substantial resources into research on those problems. Companies, for example, could institute tithing programs that apply 10 percent of income to Third World research and development projects that might never prove profitable. Such programs might help convince the public that the industry recognizes its own conflict of interest and distinguishes its societal from its investment goals. Of course, the most effective way the industry can achieve credibility is to be credible. If biotechnology companies want to convince the public that their products are beneficial, they should develop beneficial products — those that promote public goals and truly promote sustainable agriculture, prevent environmental degradation, and improve nutritional quality.

ACKNOWLEDGMENTS

This article is an extensive revision and update of "Food biotechnology: politics and policy implications," in K.F. Kiple and C.K. Ornelas-Kiple, eds., *The Cambridge World History of Food and Nutrition,* Cambridge University Press, Cambridge, UK, 2000: used with permission.

The work was supported by research challenge grants from New York University and its School of Education, and by expert research assistance from Monica Bhagwan.

BIBLIOGRAPHY

1. Office of Technology Assessment, *A New Technological Era for American Agriculture*, OTA-F-475, U.S. Congress, Washington, DC, 1992.
2. T.J. Hoban, in D. Maurer and J. Sobal, eds., *Eating Agendas: Food and Nutrition as Social Problems*, de Gruyter, New York, 1995, pp. 189–212.
3. G.E. Gaull and R.A. Goldberg, eds., *New Technologies and the Future of Food and Nutrition*, Wiley, New York, 1991.
4. Monsanto, *1998 Annual Report* Online: *http://www.monsanto.com/*
5. J. Enriquez and R. Goldberg, *Gene Research, the Mapping of Life and the Global Economy Case N9-599-016*, Harvard Business School, Boston, MA, 1998.
6. U.S. Congress, Office of Technology Assessment, *Biotechnology in a Global Economy*, OTA-BA-494, U.S. Government Printing Office, Washington, DC, 1991.
7. Z.S. Olempska-Beer, P.M. Kuznesof, M. DiNovi, and M.J. Smith, *Food Technol.* **47**, 64–72 (1993).
8. J. Sachs, *The Economist*, August 14, pp. 17–20 (1999).
9. *The Economist*, June 19, pp. 19–21 (1999).
10. E. Messer, *Adv. Technol. Assess. Syst.* **9**, 371–378 (1992).
11. A.M. Thro et al., *Nat. Biotechnol.* **16**, 428–430 (1998).
12. J.J. Putnam and J.E. Allshouse, *Food Consumption, Prices and Expenditures, 1970–97*, USDA Stat. Bull. No. 965, U.S. Department of Agriculture, Washington, DC, 1999.
13. A.E. Gallo, *The Food Marketing System in 1996*, USDA Agric. Inform. Bull. No. 743, U.S. Department of Agriculture, Washington, DC, 1998.
14. R.M. Philin et al., *Am. J. Epidemiol.* **138**, 154–159 (1993).
15. D.D. Metcalfe et al., eds., *Crit. Rev. Food Sci. Nutr.* **36**(Suppl.), s1–s186 (1996).
16. J.A. Nordlee et al., *N. Engl. J. Med.* **334**, 688–689 (1996).
17. *Consumer Rep.*, September, pp. 41–46 (1999).
18. M. Nestle, *Nutr. Today* **33**(1), 6–12 (1998).
19. Food and Drug Administration (FDA), *Guidance for Industry: Use of Antibiotic Resistance Marker Genes in Transgenic Plants: Draft Guidance*, FDA, Washington, DC, 1998. Online: *http://vm.cfsan.fda.gov/*
20. P.B. Thompson, *Food Biotechnology in Ethical Perspective*, Blackie Academic & Professional, London, 1997.
21. U.S. General Accounting Office, *Food Safety and Quality: Innovative Strategies may be Needed to Regulate New Food Technologies*, GAO/RCED-93-142. GAO, Washington, DC, 1993.
22. D.A. Kessler et al., *Science* **256**, 1747–1749, 1832 (1992).
23. M. Nestle, *Bio/Technology* **10**, 1056 (1992).
24. J. Rissler and M. Mellon, *Perils Amidst the Promise*, Union of Concerned Scientists, Washington, DC, 1993.
25. R. Goldburg, J. Rissler, H. Shand, and C. Hassebrook, *Biotechnology's Bitter Harvest*, Biotechnology Working Group, Washington, DC, 1990.
26. Environmental Protection Agency, *Fed. Regist.* **59**, 60496–60518 (1994).
27. J.E. Losey, L.S. Rayor, and M.E. Carter, *Nature (London)* **399**, 214 (1999).
28. Office of Technology Assessment, *New Developments in Biotechnology: Patenting Life — Special Report*, OTA-BA-370, U.S. Congress, Washington, DC, 1989.
29. I. Wilmut et al., *Nature (London)* **385**, 810–813 (1997).
30. International Food Information Council, *U.S. Consumer Attitudes Toward Food Biotechnology*, Wirthlin Group Quorum Survey, Washington, DC, 1999.
31. T.J. Hoban, *J. Food Distrib. Res.* **27**, 1–10 (1996).
32. *The Economist*, June 13–19, pp. 113–114 (1998).
33. G. Gaskell, M.W.S. Bauer, J. Durant, and N.C. Allum, *Science* **285**, 384–387 (1999).
34. Food and Drug Administration, *Fed. Regist.* **59**, 6279–6280 (1994).
35. Office of Medical Applications of Research, *Bovine Somatotropin: National Institutes of Health Technology Assessment Conference Statement*, U.S. Government Printing Office, Washington, DC, 1990.
36. U.S. General Accounting Office, *Recombinant Bovine Growth Hormone: FDA Approval Should Be Withheld Until the Mastitis Issue Is Resolved*, GAO/PEMD-92-26, GAO, Washington, DC, 1992.
37. J.M. Chan et al., *Science* **279**, 563–566 (1998).
38. S.E. Hankinson et al., *Lancet* **351**, 1393–1396 (1998).
39. S. Gilbert, *N.Y. Times*, January 19, p. F7 (1999).
40. Food and Agriculture Organization of the United Nations, *Press Release*, FAO/UN, Rome, 1999. Online: *http://www.fao.org/*
41. M. Graves, *Los Angeles Times*, August 18, p. C1 (1997).
42. P.J. Longman, *U.S. News & World Rep.*, July 26, pp. 38–41 (1999).
43. E. Dorey, *Nat. Biotechnol.* **17**, 631 (1999).
44. Y.M. Ibrahim, *N.Y. Times*, November 7, pp. D1, D24 (1996).
45. S. Kilman and H. Cooper, *Wall Street J.*, May 11, pp. A1, A10 (1999).
46. R.A. Steinbrecher and P.R. Mooney, *The Ecologist* **28**(5), 276–279 (1998).
47. E. Masood, *Nature* **398**, 641 (1999).
48. S. Kilman, *Wall Street J.*, October 7, pp. A1, A15 (1999).

See other entries AGRICULTURAL BIOTECHNOLOGY; ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE.

AGRICULTURAL BIOTECHNOLOGY, SOCIOECONOMIC ISSUES, AND THE FOURTH CRITERION

WILLIAM LACY
University of California
Davis, California

OUTLINE

INTRODUCTION

The world economy is currently undergoing major structural changes. A central factor in these changes has been the development and diffusion of fundamentally new kinds of technologies, in particular, computers and the new biotechnologies. Social and economic changes that result from these profoundly enhanced capacities in science and technology are visible in every sphere of human life from health, transportation and communication to agriculture and the food system. However, each change is associated not only with new benefits but also with new risks, latent complications, and long-term consequences that are often poorly understood. Some have argued that the new biotechnologies may be the most radical experiment humankind has ever carried out on the natural world, in many ways representing our fondest hopes and aspirations as well as our darkest fears and misgivings. The technology according to some actually touches the core of our self-definition (1).

The new tools are arguably the ultimate expression of human control both helping us to shape and define nature itself as well as our very sense of self and society. The changes brought by biotechnology could deeply affect our individual and collective consciousness, the future of our civilization, and the biosphere itself. Until recently far more public attention has been focused on the other great technology revolution of this century, computers and telecommunications. However, after nearly 40 years of parallel development, the information in life sciences is slowly beginning to fuse into a single technological and economic force. Computers are increasingly being used to decipher, manage, and organize the vast genetic information as a new resource of the emerging biotechnology economy. This new field called "bioinformatics" is being used to download the genetic information of millions of years of evolution and thereby creating a powerful new genre of biological databanks. This new genetic information database may be used by researchers to remake the natural world. As Jeremy Rifkin has noted "These changes represent a turning point for civilization. We are in the throes of one of the great transformations in world history" (1, p. 4).

In the area of agricultural biotechnology, the president of a large U.S. public university noted, "Our society has moved into the era of high technology ... as we move into the new millennium, we will see more technological changes than we have experienced over the entire history of our nation. It promises to be one of the most exciting and challenging times in the history in mankind." He further observed, "Biotechnology, genetic manipulation and engineering research will have tremendous impact on the crops and animals we grow for food, affecting agriculture in ways never before dreamed possible" (2, p. 3). According to Rifkin (1), global agriculture could find itself in the midst of a great transition in world history, with an increasing volume of food and fiber being grown in enclosed tissue culture vats. The shift to indoor agriculture could bring significantly low prices, more abundant supplies of food, and massive displacement of millions of farmers in both the developing and developed world.

Despite the enormous optimism in the scientific community, national and state governments, and the private commercial sector, the applications of biotechnology have been fraught with concern and controversy within the both the scientific community and the broader public. Much of the initial public concern has centered on human and animal health and environmental safety issues. Issues of the environmental impact of the creation, mass production and wholesale release of thousands of genetically altered life forms into the environment and the potential irreversible change and wholesale reseeding of the earth's biosphere has been raised. Many groups of scientists and environmentalists and the local citizen groups have raised safety issues regarding the unexpected but possible consequences of introducing new life forms such as the production of a toxic secondary metabolite or protein toxin or the undesired self-perpetuation and spread of the organism. In agriculture, some have suggested that living natural inputs may be even more dangerous to society than the artificial products they replace. In addition there is concern that genetically engineered crops bearing resistance to nature herbicides may become weeds. Some fear that these herbicide-resistant crops may even cross with weedy relatives and spread resistance into sectors of the weed flora. In the area of human safety there has also been concern raised about the safety and human health issues surrounding genetically modified foods and pharmaceuticals. In the case of human health, these include potential allergenic or toxic effects resulting from genetic changes that are not completely understood.

THE FOURTH CRITERION

The debate, however, cannot be reduced to a simple risk controversy that focuses on health and environmental safety issues. While the three standard criteria often utilized to evaluate and approve new products and processes have been (*1*) human safety, (*2*) animal and environmental safety, and (*3*) efficacy, increasingly over the last couple of decades, a fourth criterion or fourth hurdle for product approval and regulation has been proposed. This refers to the social and economic effects of the product or a technology. With many citizens in both the United States and Europe the issue of food and agriculture modified from modern molecular genetics and biotechnology elicit deeper concerns about the relationship to the natural environment where there are strong dimensions of social and political risk. Increasingly it is recognized that the issues of agricultural biotechnology are not purely technical but also concern the balance between the different worldviews and values that enter the scene of each national regulation and product approval process. Assessing the risk associated with agricultural biotechnology therefore becomes a complex problem that is being heatedly debated. Many efforts have been directed

against the use of the fourth criterion on the basis that it inhibits trade and is in violation of a number of global trade agreements. Despite the explicit prohibition of the use of the fourth criterion to inhibit trade, the social criteria are being implicitly or explicitly included in a number of policy debates and decision-making processes. As a result of these developments and experiences with previous technologies, an increasingly accepted position among technology assessment professionals is that (*1*) all technologies have multiple effects, (*2*) many of these effects are potentially harmful and require conscious decisions, and (*3*) these critical decisions entail social, economic and moral as well as scientific analysis (3,4).

In the early 1990s this broader fourth criterion was employed in the European Common Market's ban on growth hormones in food products. At that time the Advocate General of the Court of Justice in the European Communities released an opinion on the legality of the hormone stating that "it was appropriate and justifiable to prohibit the administration of the five substances for fattening purposes, even in the absence of scientific evidence showing that they were harmful. A total prohibition was the only solution which could bring an end to the distortions of competition and barriers to intra-Community trade in meat, eliminate risks to public health, even if they were purely hypothetical ones, and avoid a further reduction in consumption" (5, p. 1).

Similarly, in the Austrian biotechnology regulations, the social and economic criteria are clearly present. Their regulations state that products containing or consisting of genetically engineered organisms must not cause any social unsustainability, no unbalanced burden on society or any social group that is unacceptable for economic, social, or moral reasons (6). Arguments regarding health or environmental risk are frequently countered with fears of socioeconomic hazards. In practice, the two perspectives are exceedingly difficult to separate. The absence of risk to environment, life, and limb appears to be the necessary, but inadequate, condition for acceptance among increasingly larger percentages of the Austrian population. As a consequence Austria passed paragraph 63 of the Genetic Engineering Act of 1994 which required that the genetically engineered products be assessed for the possible risk of social unsustainability before they could be licensed. These provisions seem to contradict the European Union (EU) Commission's ruling because they include aspects of assessment according to socioeconomic criteria that had been clearly rejected as the fourth hurdle to licensing. According to the EU Commission, socioeconomic criteria cannot be assessed by scientifically clear criteria and therefore would likely provoke legal issues and lead to a drain of capital.

Despite the EU's policies and prior to the passage of the Act, the Austrian parliament voted unanimously to set up a parliamentary inquiry commission to investigate technology assessment issues based on the example of genetic engineering. In its report in 1992, the commission demanded that social sustainability be taken into consideration in addition to the ethical requirements and environmental impact. One member of the commission indicated that the concept of social sustainability should entail safeguarding the balance of interests and the maintenance of consensual value orientations. Moreover, the inquiry commission recommended information for the public, participation by the public, mandatory disclosure (annual report on all genetic engineering activities), and measures to promote the public discussion of genetic engineering. Paragraph 63 of the Austrian Genetic Engineering Act specifically stated that genetically engineered products must not lead to social unsustainability and would not be approved for use "if it may be assumed on a technical basis that such products would lead to an unbalanced burden on society or on social groups, and if this burden no longer appears acceptable to the population for economic, social or moral reasons" (6, p. 303).

To implement the socioeconomic criteria for product approval, several models were explored. One focused on creating an interdisciplinary expert panel to design future scenarios and test them scientifically for their compatibility with the Constitution and societal values. The decision on introducing technology was to reside in diverse politically authorized groups. Criticisms of this scheme included the fact that experts would not be able to obtain consensus because of the pluralistic values of these groups. A second model involved the idea that anything that is accepted by society is socially sustainable. However, this meant turning the agendum into research on societal acceptance and focusing on the manipulation of public opinion and the creation of acceptance. The third approach focused on the participatory process where social sustainability would remain a transparent and preliminary working term only put into practice if a situation demanded it. The requirement that certain agricultural biotechnology products must not cause social unsustainability made them a mandate for public discussion and negotiation. The state's function was to provide a framework for the public dialogue to occur and to guarantee that framework from a legal point of view. However, it is unlikely that anyone was considering participatory procedures when social sustainability was conceptualized. It was something viewed instead as an expertise type of administrative procedure. Nevertheless, in time meaningful participatory procedures were developed for technology assessments, and they are now in place in Austria.

The compromise in Europe has been to introduce the precautionary principle to address uncertainty without specific reference to the fourth criterion or socioeconomic concerns. In the recent European Union's directive on genetic engineering and international laws, the precautionary principle was upheld. While it comes in many interpretations, the principle is that whenever there is a serious threat of irreversible damage, any lack of scientific evidence must lead to postponement or avoidance of the biotechnology. Built into this conclusion is time for deliberation. The use of precautionary measures allows for selection among possible risks and consideration of the severity of those risks. The decision regarding risks depend on who is affected in which ways and who derives which benefits, and this ultimately takes on socioeconomic dimensions.

The fourth criterion has also entered the discussion and negotations for an international biosafety protocol. In 1998 an ad hoc working group on biosafety met

and heard concerns from many members of developing countries about the impact of transgenic crops on their farming communities, and they voiced their support for adding socioeconomic issues to the protocol (7). The ability to implement certain new biotechnology products varies greatly across countries and regions (e.g., the enforcement of obligatory resistance management strategies for Bt crops will be difficult in a country like India, and there are legitimate concerns that where such strategies fail, the commercial usefulness of Bt transgenic crops will be limited to just a few years). Those advocating for socioeconomic factors in the biosafety protocol argue that it is important for individual countries to be free to examine, case by case, the social and economic impacts of an imported biotechnology. The decision to subjugate national interest to that of free trade is particularly problematic in the case of a developing country's agriculture where impact of biotechnology on farming systems will differ greatly from that in industrialized countries. Moreover the regulatory interests of countries that will primarily export genetically modified crops and animals will differ from those that do not have a domestic biotechnology industry. The immediate economic benefits of a burgeoning agribiotechnology market unencumbered by regulatory controls pose potential barriers to free trade, and these benefits will accrue to corporations based in countries that are in a position to export the genetically modified products. Therefore individual countries must be allowed to consider the socioeconomic impact of importing genetically modified plants and animals on a case by case basis in order to ameliorate the effects of disparities (7).

As noted earlier, some countries of the European Union, in particular, Austria, Denmark, Sweden, and Finland, have expressed the need to interpret the European directive on the deliberate release of living modified organisms in order to permit them to consider the socioeconomic impacts. The Norwegian Gene Technology Act developed by a non-EU country goes even further. The Act states that in "Deciding whether or not to grant the application, significant emphasis shall also be placed on whether the deliberate release represents a benefit to the community and a contribution to sustainable development" (7, p. 698). The Norwegian interpretation takes the discussion one step further: If there is any question that a negative impact may arise, this doubt must comply with environmental and the social factors that take priority and override the particular application of the biotechnology.

In another recent example of the potential utilization of broader socioeconomic criteria, Switzerland was faced with a Gene Protection Initiative that demanded the government to outlaw the release of genetically altered organisms into the environment as well as the patenting of transgenic animals and plants, of their components, and of the relevant processes. This initiative, defeated in a national vote, demanded that experiments of all genetically modified organisms require proof of the lack of alternatives and a statement of ethical responsibility. Schatz (8) noted that despite Switzerland being the corporate headquarters of a number of multinational biologically based corporations, the public has begun to view companies around the world as heartless giants and choose the high tech products of those giants as targets of their frustration. The Swiss basically have begun to doubt that their elected representatives can reign in the international conglomerates.

In contrast, the U.S. Executive Branch concluded a review of literature on the social consequences of rBST with this revealing sentence: "At no time in the past has the U.S. Federal Government prevented a technology from being adopted on the basis of socioeconomic consequences" (9). Despite this statement the recent report from the U.S. National Research Council (4) entitled *Understanding Risk: Informing Decisions in a Democratic Society*, has moved the U.S. debate on technology impact from a narrow scientific discussion of risk assessment to the broader issue of risk characterization. The committee was asked to review "the appropriateness of including in risk characterization such considerations as economic factors, equity issues, risk mitigation and tradeoffs, and technical control feasibility as well as environmental-equity issues and other issues of social context" (4, p. x). They were also charged to examine ways for improving public participation and building trust. As a consequence the study addressed such issues as the social, behavioral, economic, and ethnical aspects of risk that were viewed as relevant to the content or process of risk characterization (4).

This brief discussion illustrates that the social and economic concerns about biotechnology have become part of the policy and regulatory process. To date, these concerns have only selectively become a fourth criterion or fourth hurdle. The extent to which these concerns are made explicit may vary, but they are inherent in the debates in both the developed and developing nations. Therefore it is important for scientists, regulators, policy makers, and citizens to understand and evaluate not only elements of human safety, animal safety, environmental risks, and efficacy but also the range of socioeconomic impacts and concerns.

The potential social and economic impacts of agricultural biotechnology on the food and fiber system and society are just emerging. Consequently the proposed implications of biotechnology for the system represent only possible scenarios. The socioeconomic effects may include impacts on: (*1*) farmers, rural communities, and the food system; (*2*) the organization and structure of agribusiness and industry; (*3*) consumers; (*4*) science and technology transfer; and (*5*) the global economy and developing nations. The social impacts and consequences of any technology are likely to be dispersed in both time and space and occur through a wide variety of mechanisms. The social impacts of technology are controversial, in part, because the mechanisms that link technological innovations to their eventual consequences are generally opaque to both developers of the innovation, scholars of technology, policy makers, and citizens.

FARMERS, RURAL COMMUNITIES, AND THE FOOD SYSTEM

The impact of agricultural change on rural communities is largely proportional to the level of local dependence

on agriculture. Today nationwide, fewer than 40 congressional districts have more than 20 percent of their population living on farms. The overwhelming majority of farms that once existed in the United States no longer exist, and production is highly concentrated among the remaining farms characterized by productivity-enhancing technology. In 1978 there were 2.3 million farms in the United States while there were less than 2.0 million by 1999. Only 6 percent of U.S. farms, involving primarily the super-large farms, receive the majority of farm receipts (10). Over the last century, agricultural technologies have emerged that use ever greater levels of capital to enable fewer people to produce the nation's food. As a result income and opportunities have shifted from farms to the companies that produce and sell goods to farmers. As farmers focused on producing undifferentiated raw commodities, food system profit and opportunities have shifted to the companies that sell the farm inputs, and process, package, and market food. Consequently from 1910 to 1990 the share of the agricultural economy received by farmers dropped from 21 to 5 percent (11). Agricultural biotechnology will likely continue this trend with the profits accruing to the industries developing the biotechnology products. Finally, and importantly, a trend that appears in all sectors of American agriculture is a widening spread between what farmers receive for their production and what consumers pay at the supermarket.

The industrialization of agriculture in the United States has almost one-third of the total value of production of U.S. farms generated under contractual arrangements and mostly under market contracts. Large agricultural integrators tend to avoid capital investment in the means of production and pass the risk and costs on to their contract growers or to society at large. Under these conditions farmers contract to sell their products to a specific processor or contractor, but the farmer owns the product and the risks until the product is sold and makes all the managerial and production decisions. Production contracts are also increasing with the contractor owning the livestock or crop and paying the producer a flat fee plus additional payments for performance-based incentives. Under these conditions the producer or farmer becomes very similar to industrial laborers. The poultry industry is perhaps the most industrialized subsector of agriculture with 89 percent of poultry farms using contracts and about 86 percent of the total value of poultry production grown under contract. Competition in the hog, cattle, and lamb industries has been declining even before the recent rise in livestock contracting with the proportion of the market controlled by the four largest steer and heifer slaughter firms increasing from 36 percent in 1980 to 72 percent in 1990 and 82 percent in 1994 (10). The vast majority of small farms, however, are now buffered from the effects of technological change, since the farm is no longer the primary source of income for their owners. Consequently biotechnology will probably have less impact on the total number of farms in the United States and developed countries than previous mechanical and chemical technologies adopted by farmers during the last 50 to 75 years. Moreover it is likely biotechnology will not greatly exaggerate the decline in the number of farms, although it will certainly maintain present trends, which indicate that farming will continue to be one of the fastest declining occupations.

To better understand the potential impact of technologies such as agricultural biotechnologies, it may be important to briefly review the concept of the technological treadmill. Numerous scholars have argued that new production technology allows farmers to reduce the costs of production with early adopters of the technology reaping substantial profits. They produce more than their neighbors can with a comparable investment of time, labor, and capital. As more and more people adopt the new technology, however, total production rises and prices begin to fall. Those operating with the old technologies find themselves operating at a loss, and they often go out of business. On the other hand, those who adopt the new technology, find that higher profits disappear and they are producing more food to retain the same income level. However, the treadmill is more than a technology transfer. It also accounts for how societies consisting of many independent, owner-operated farms become societies that consist of a small number of land and capital-owning investors with masses of workers relegated to wage labor. The social transition described by the technological treadmill process is thus a change in social structure. The fundamental change that occurs is a shift from owner-operators, each with control over their work activity and relatively equal opportunity to succeed, to a society of owners and managers of capital who control the work life of laborers and who determine future directions of society through their investment policies and practices (12).

Genetic engineering is likely not the most important technology implicated in this transition, particularly in industrialized countries where this transition occurred long before the development of biotechnology. This technology may be far more important in affecting the social structure of agricultural economies in the developing world. Small farmers in those nations constitute a significant if not majority portion of the population, and they will become displaced or marginalized as urban populations begin to rely increasingly on industrialized agriculture from Europe, North America, Japan, and Australia. Even the successful few and large producers will have to share a larger portion of their farm profits with the companies that produce the biotechnology and become more dependent on those companies in a manner not unlike those of wage laborers dependent on their employers.

Other changes and impacts on family farms and rural communities include the shift in returns on production from labor to capital. Capitalists use technology to gain a larger share of the value of their product at the expense of labor. A new technology lowers costs and eventually dominates the industry. Those who work in the industry then are forced to accept wages offered by the owners of the technology. Thompson and others argue that this shift violates a farmer's right, for it reduces the farmer's autonomy and control in disposing of his primary assets, land and labor. Kloppenburg and his colleagues (13) have proposed a further impact that concerns the loss of a moral economy associated with traditional agriculture. They

argue that the institution of alienable property rights in land introduced commercial practices into food production that have inexorably undercut the moral economy. They propose that rural and urban people can invigorate a moral economy for contemporary agriculture that will reverse the commodifying influence of technology and its attendant impact on the human condition. The underlying assumption here is that if capitalism systematically consigns labor to a situation of wage servitude, it cannot be considered morally legitimate.

The extent of biotechnology's influence on the trend toward fewer and larger farms depends, in part, on how adoption effects the cost structure of farms. If biotechnology development significantly alters costs, returns, competitive positions, and the special location of production, and if certain trade and farm policies are implemented, the potential impact of biotechnology could be relatively important. It has been argued that these new technologies, like those of previous generations, will be adopted by well-financed, innovative farmers who are presumed capable to run the competitively large farms. However, others have argued that biotechnology innovations will provide widespread benefits to the full range of farmers because new technologies will be used in traditional ways. Regardless of which farmers are likely to benefit, however, biotechnology will probably increase the value added off farm at the expense of value added on farm.

Other significant changes in the farming community may result if the information and products of this technology bypass the Cooperative Extension system and the agricultural cooperatives. Previous products and information of biological research have been disseminated through the Cooperative Extension system. However, the development of new seed and chemical packages through biotechnology has emerged from private research. Public sector scientists may have limited knowledge with which to support extension education programs, with a consequence that extension, and potentially agricultural cooperatives, may gradually be reduced to playing a secondary role in farm change. Moreover many agriculturally based rural communities will continue the ongoing process of shrinkage and consolidation, as producers, and local supply and marketing firms continue to decline in numbers. Biotechnology may also accelerate the trend noted above regarding the integration of contract farming, already common in the United Status, where commodities such as poultry and most processed vegetables are produced on contract. These arrangements will further reduce the autonomy of farmers and will certainly reduce their contact with and need for extension education, agricultural cooperatives, and local farm suppliers. The new biotechnologies may also restructure the relationship between farmers and researchers. Until very recently farmers were seen as a primary clientele of public sector research. However, the entry of molecular biology into agricultural research has increasingly been accompanied by the insertion of the agribusiness sector between farmers and researchers. As a result it is quite possible that the interests of agribusiness sector will dominate the agenda setting in the public research arena (2).

Michael Gertler (14), participating in a National Agricultural Biotechnology Council annual meeting, focused on several reasons why he thought agricultural biotechnologies may become, or should be, social issues in rural agricultural communities. He noted that with the advent of the agricultural biotechnology products, some of which are proposed as environmentally friendly such as herbicide-resistant crops, the farmer is likely to incur increased costs and risks without assurance of gains. The use of expensive genetically engineered seeds do not guarantee a commensurate increase in yields. Furthermore supply companies and firms licensing particular generically engineered organisms are adept at charging what markets bear with the economic benefits arising from these technologies likely to be accrued by those holding the patents.

A second concern Gertler observed is the industrialization and accelerated structural change in the farm sector noted earlier. Biotechnology is being introduced in the context of increasing the industrialization of farming. This technology is an important development in responding to environmental, agronomic, and veterinary problems encountered when industrializing livestock and crop production. Gertler argues that it may permit further industrial development without addressing fundamental contradictions and efficiencies of this system. Although biotechnology may appear to be scale neutral, the level of investment required, the increased risk, and need for higher levels of management mean that larger and more capitalized farmers will likely benefit disproportionately. Moreover biotechnologies may create deeper divisions between farmers subscribing to different models or systems of production, between farmers and nonfarm rural populations, and between farmers and the nonrural public. For example, organic farmers may feel even more marginalized as the traditional industrialized agricultural food system embraces these new technologies and public section research is directed more extensively to production systems that overlook their needs. The organic growers, however, may experience increased demand from consumers distrustful of these new genetically engineered food products. Another potential impact is on the self-esteem of farmers as they are transformed from practitioners into the objects of agricultural practice. The proliferation of new genetically engineered products and processes may inhibit the ability of farmers to make educated choices with respect to crops and inputs appropriate to their regions and cropping systems. Gertler concludes by observing that farmers may eventually become more like consumers, less able to distinguish quality because of product proliferation, lack of information, and misinformation.

Perhaps the broadest analysis of potential social and economic impacts of modern technology and the new agricultural biotechnologies on rural communities focuses on the sweeping challenges to democratic rights that it poses. Langdon Winner argues that technical changes have the social effect that are equivalent to legal or constitutional changes. Here citizens would not tolerate such sweeping changes coming through government without due process, but scientists and business leaders

are able to bring about wrenching social change through the process of the introduction of new technologies that is totally isolated from public influence or participation. For Winner and others this amounts to a total usurpation of the most fundamental democratic rights and has been the basis for proposing the need for a fourth criterion to regulate technology (12).

Others such as Wendell Berry (15) have argued that industrialization undermines the moral meaning of work which is considered both the formation and expression of personal identity. The hard work that is necessary for traditional farming has the effect of providing the farmer a well-developed sense of self, an identity that attaches naturally to a set of interests arising from work. In contrast, Berry argues that the factory pattern of life encourages people to identify with leisure activities and to acquire interests that are not related to their identity or self-expression. This ecological perspective of work is embedded in his vision of community. Farmers depend not only on each other but on tradespeople, merchants and other members of the rural town. These constitute particular nonuniversal dependencies that establish strong moral bonds to specific community members. In such a community the farmer is linked to others in the community by their work activities that form their personalities and identities. As a consequence, Berry argues, community becomes meaningful as an ethical concept (15). The extent to which biotechnology continues a process of industrializing farming, its impact, in Berry's terms, goes far beyond the farm itself.

AGRIBUSINESS AND INDUSTRY

Many business and government leaders view biotechnology as a force to not only restructure farming and rural communities but also to catalyze a major change in the structure of worldwide agribusiness. They note that the application of molecular biology permits the various segments of the world's largest industrial sector, agribusiness, to form logical linkages to other economic sectors as was never before practical. This $1.3 trillion agribusiness sector, not counting feed and fiber, consists of four basic elements: input suppliers, growers, processors, and consumers. The system has experienced mechanical and chemical eras that contributed to increased productivity and efficiency and will likely continue to make significant contributions in the future. However, according to a number of business leaders, the new biological and biotechnological era will further increase both efficiency and productivity along with the ability to change the quality of food and feed. It will lead to consolidation and new forms of vertical and horizontal integration of the food industry (2,16).

The formation of new biotechnology companies increased dramatically, starting with the founding of Genentech in 1976, with more than 250 small venture capital biotechnology firms founded in the United States over the next decade. Proliferation of these risk-tasking companies helped raise billions of dollars from private investors and gave the United States a comparative lead in the early stages of biotechnology commercialization. By the late 1980s the number of these firms had grown to over 600. Despite consolidation which began in the industry with mergers, bankruptcies, and major multinational corporation investments, the number of companies have continued to grow. By 1998 the Genetic Engineering News, *Guides to Biotechnology Companies*, listed over 3500 companies worldwide and approximately 1500 in the United States with a substantial number involved in agriculture (17). However, multinational corporations now clearly dominate biodevelopment. These corporations presently control nearly a third of the fledgling bioindustry, a figure predicted to rise to 50 percent in the near future.

During the 1980s and continuing throughout the 1990s, these multinational corporations began diversifying into every field or specialty that uses living organisms as a means of production. The new biotechnologies appear to further reduce the distinctions among the traditional industrial sectors, rendering corporate boundaries virtually unlimited. These large multinational corporations specializing in chemicals, food, and pharmaceuticals have taken the leadership in agricultural biotechnology research and development (e.g., American Cyanamid, Agr Evo, Dow Chemical, Dupont, Eli Lilly, Merck, and Monsanto). At the forefront is Novartis, formed in 1996 by the merger of Sandoz and Ciba-Geigy. Novartis is now the world's largest agrochemical company, the second largest seed firm, the third largest pharmaceutical firm and the fourth largest veterinarian medicine company (16). At the same time companies like Monsanto moved rapidly to expand and consolidate their market share of several key crops with their genetically engineered seeds. For example, in 1997, 15 percent of the U.S. soybean crop was grown from genetically engineered seeds. By 1998 this had grown to 44 percent of the soybean crop and 36 percent of the nations corn crop with Monsanto's Round-up Ready (herbicide resistant) seeds controlling a majority of the market. In 1998 U.S. farmers planted more than 50 million acres of genetically modified crops about six times the acreage planted with such crops just two years earlier (18). Worldwide it was estimated that in 1999 GM crops were grown on over 100 million acres (19). Indeed, two-thirds of the genetically engineered crops available in 1999 are designed specifically to increase the sale of herbicides and pesticides produced by the companies selling the genetically engineered seeds (20). Lappe and Bailey conclude their analyses of the U.S. and international agricultural biotechnology developments by noting that never before in the history of the world has such a rapid and large-scale revolution occurred in a nation's food supply. According to bioindustry analysts, by the year 2025 some 70 percent of the industrial economy and 40 percent of the entire global economy will have, at its base, some form of biotechnology.

Michael Pollan, a writer for *The New York Times*, recently noted that with the advent of biotechnology, agriculture is entering the information age with a small number of multinational corporations positioned to become another Microsoft, supplying the proprietary operating systems to run the new generation of plants and animals (21). Most analysts predict biotechnology will continue and accelerate the trend toward increasing

concentration of power in a small number of large multinational corporations. Consequently development and commercial control of agricultural biotechnology will be in the hands of corporations that transcend geography boundaries and hold limited national allegiance. Within this context people question how we can ensure that democratic participation will occur in the decision-making processes surrounding the development and commercialization of biotechnology. This is difficult within the national boundaries and generally prohibited internationally given current government structures.

CONSUMERS

For consumers, the new biotechnologies could mean dramatic improvements in the productivity and efficiency of food production and processing, and the expansion and extension of food and nonfood uses of raw agricultural commodities. Consumers could benefit in the form of reduced prices, increased food safety, and more nutritional foods. Products in the pipeline, for example, could produce plants that lack allergenic proteins or have a healthier oil composition and may also provide benefits for developing countries such as the pro-vitamin A and iron-enriched rice (22). The new technologies also have a potential to change the very nature of food itself and to expand a range of possible food products. It is now possible to consider the production of new fabricated foods in which basic foods are broken down into their component parts (e.g., starch, fat, and sugar) and recombined into wholly new types of food. However, to date, the new food products and processes have been met with mixed reactions. On one level are the concerns about food safety from unexpected allergenic or toxic foods resulting from the insertion of a foreign gene into crops, food, or animals. At issue is whether a foreign gene can activate the expression of a latent toxic gene. On a more subtle level, the new biotechnologies may make it far more difficult in the future for the consumer to determine the composition of the food and to maintain a balanced diet.

Another impact of biotechnology has been the stimulation of new moral and ethical debates among consumers and the general public regarding the limits of science. Public concern about a range of scientific developments, including biotechnology, are resulting in a decline in public confidence in science and increasing public perception of the likelihood of environmental risks from genetically altered bacteria, plants, and animals. Krimsky and Wrubel argue that on the basis of their analysis the level of premarket public scrutiny of some of the first products of agricultural biotechnology has been unprecedented. Citizens are demanding earlier entry points and broader participation in technological decisions (23). The development of biotechnology is stimulating a wider range of public concerns about science that extend beyond human health, environmental risks, food safety, and animal health issues and includes such concerns as negative socioeconomic consequences and the morality of tampering with nature and life itself. Many environmental and consumer groups view transgenic food as a symbol of the assault on traditional sources of food. At issue here is the dignity of the food supply even more than its safety.

In 1999 reacting to the escalating public concern, several European supermarket chains banned GM (genetically modified) products from their house brands. Moreover in Britain, Unilever and Nestle announced that they would phase out genetically modified ingredients in their products. Meanwhile the European Union decided that products in which more than 1 percent of one of the ingredients was transgenic should be labeled and that the introduction of new GM crops would be suspended for several years. Even in the U.S. the Federal Drug Administration chose to hold public hearings around the country on whether it should adjust its role in regulating GM crops. These hearings were often confronted by consumer protests against "foods created by altering genes" (frequently characterized as "Frankenfoods"). Finally, leading food manufacturers in the U.S., Gerber and Heinz announced that they would permit no GM foods in their products. Many analysts have concluded that the next few years will be crucial for the future of GM crops and that in the end consumers, rather than the farmers that the industry had long considered its primary customers, will decide the fate of GM foods (22).

While, a small number of groups oppose any form of genetically modified food, the general focus of public policy consumer concerns has turned towards the question of labeling. Krimsky and Wrubel (24) note that labeling transgenic food would enable consumers to express social values in their food preferences, which is consistent with the trend towards "green consumerism." In 1999 labeling genetically modified food was high on the agenda of the Codex Alimentarius Commission with members of the European Union who were conscious of public pressure over GM foods; they argued that any food containing detectable GM ingredients should be labeled (24). From recent analyses of public and consumer concerns about biotechnology in Europe, people seem prepared to accept some risks as long as there is a perception of usefulness and no moral concern. But crucially, moral doubts act as a veto irrespective of people's views on use and risk. In one study the finding that risk is less significant than moral acceptability in shaping public perceptions of biotechnology held true for each European country and across all six specific applications of biotechnology (genetic testing, medicines, crop plants, food production, research animals, and xenotransplants). In the same survey, 74 percent of the respondents consider that genetically modified foods should be labeled and 60 percent believe that there should be public consultation about new developments in biotechnology (25).

In a column in *Nature*, the editor stated that the genetically modified foods debate needs a recipe for restoring trust. While there is no simple institutional formula for achieving this, some of the principles include (*1*) acceptance for the need to ensure the regulation of genetically modified foods based on the soundest possible science, (*2*) acknowledgment of the current limits to scientific certainty, (*3*) the need to find ways of facilitating public access to credible scientific information and of communicating it in a responsible form, (*4*) the need

for honest brokers, and (*5*) taking into account in food regulations broad public concerns. The editorial concluded if labeling all foods produced by genetically modified techniques turns out to be a necessary step in regaining trust on both sides, it should be a small price to pay (24).

SCIENCE AND TECHNOLOGY TRANSFER

Perhaps the most dramatic, immediate impact of the new biotechnologies is on science itself. During the last 20 years, the convergence of a number of new scientific techniques and biotechnologies, legal policy, and commercial developments has had a major impact on the way in which knowledge is generated and commercialized and on the evolution of agriculture and our food system. While some argue that biotechnology is a continuation of the application of biological techniques to improve plants, animals, and microorganisms, many biologists contend that biotechnology has revolutionalized the field. The knowledge and tools generated by molecular biology and biotechnology have stimulated a great deal of enthusiasm and redirected large sums of money in an effort to pursue knowledge in this area. At the federal level, financial support for biotechnology has grown steadily since the mid-1980s and reached several billion dollars annually in the 1990s. While 80 percent of the federal nonmilitary research budget has been devoted to the National Institutes of Health (NIH) program, support for agricultural biotechnology has been relatively meager, constituting less than 3 percent of federal expenditures for biotechnology.

The techniques and tools of biotechnology are facilitating basic research efforts to understand the intricate, complex, functioning of living organisms at the molecular and cellular level. This reductionist approach, often called logical positivism, continues and extends the basic methods and approaches of modern science. Modern biology attempts to reduce nature to small, definable pieces, subject to human manipulation, and separated from broader questions of value. From this perspective, scientists control, measure, reduce, and divide nature in order to generate knowledge. Biotechnology, particularly in agriculture, may truncate both the time and space required to develop new plant, animal, and food products. However, one concern is that this approach, while providing important but only partial knowledge, is rapidly becoming the dominant epistemology, often to the exclusion of important alternative ways of knowing. As a consequence lack of adequate support has occurred for critical complementary research to molecular biology and genomics agendas, and this includes whole-plant and animal-level research (e.g., traditional plant breeding), systems-level research programs (e.g., agroecology, farming systems), social assessments, and indigenous knowledge (2).

Another development stimulated by agricultural biotechnology with implications for the generation of knowledge is the increased concentration of research funds, scientific talent, and intellectual property at a small number of public and private institutions. In the public sector, every U.S. state could afford and has had conventional soils, breeding, and pathology programs. Every state cannot afford and will not be able to have a comprehensive agricultural biotechnology program. By the late 1980s and early 1990s, for example, eight states accounted for over half of the State Experiment Station expenditures and nearly half of all science years for agricultural biotechnology research (26). It is unclear how the absence of diverse and heterogeneous institutions and groups of scientists will affect the generation and dissemination of knowledge.

The new agricultural biotechnologies are also contributing to a changing collaborative relationship between the universities and industries. While partnerships between universities and industries have existed for several decades, the new types of university and industry relationships in biotechnology are generally more varied, wider in scope, more aggressive and experimental, and more publicly visible than the relationships of the past. The legal/contractual bases for these relationships depend on the goals and institutional characteristics of the partners, and consequently involve diverse approaches including: large grants and contracts between companies and universities in exchange for patent rights and exclusive licenses to discoveries; programs and centers organized with industrial funds at major universities, that give participating private firms privileged access to university resources and a role in shaping research agendas; professors, particularly in the biomedical sciences serving in extensive consulting capacities on scientific advisory boards or in managerial positions in the firms; faculty receiving research funds from private corporations in which they hold significant equity; and public universities establishing for profit corporations to develop and market innovations arising from research (27).

A notable example of these new types of collaborative arrangements between universities and industry is the five-year $25 million "strategic" research alliance announced in late 1998 between the University of California, Berkeley's College of Natural Resources and a unit at the Swiss biotechnology giant, Novartis. While large multimillion dollar industry grants to universities are not unheard of, this agreement applies not to a single researcher or team focusing on a specific topic but rather to the entire department of plant and microbiology. Under the agreement the Novartis until will provide funds and access to proprietary technology to Berkeley faculty members and graduate students, and in return it will receive first rights to negotiate licenses up to one-third of the inventions that result. Novartis is also considering the development of a facility on or near the Berkeley campus for 20 to 30 of its own scientists who would be available to work with university researchers and to share equipment and space (28).

The university and the private sector have very different goals for research and ways of pursuing those goals. When collaborating, the consequences of these two distinct and complementary research communities can be both positive and negative. In the United States, for example, university–industry collaboration may bring useful products to market more rapidly and promote U.S. technological leadership in a changing world economy. Second, in light of funding stagnation within U.S. Department of Agriculture (USDA) and in many

states, such collaborations are a means of raising new funds for university research and support for graduate education. Third, these joint efforts may expand the scientific network, increasing communication between some university and industry scientists and provide some university scientists access to cutting-edge research tools, proprietary materials, and vast databases owned by the particular company (29).

However, a number of concerns have been voiced regarding the impact of these new relationships. First, long-term research, previously a major emphasis of the public sector, may decline. The private sector has short-term proprietary goals, and as a consequence funding for research is also generally short term, spanning one or two years. In contrast, nearly all the federal NIH extramurally funded programs and USDA Hatch-based funded projects are for three years or longer. Moreover dependence on private sector funds will generally change not only the time frame but also the stability of funding. It seems unlikely that these university–industry relationships will provide stable long-term funding, nor will they significantly address the capital needs of the universities. For example, in a study of executives of 210 agricultural, chemical, and pharmaceutical corporations, 59 percent reported supporting university research totaling $340 million for more than 1500 projects. However, most said their support lasts two years or less and involves research contracts for less than $100,000 (30).

Universities are also concerned about ensuring that research projects are generally originated by faculty members and not adopted as a result of outside pressure, either implicit or explicit. If a sufficiently large and influential number of academic scientists and engineers become involved with industry, a whole range of research agendas, traditionally the purview of the university community, might be de-emphasized. Furthermore the scientific community could become desensitized to the environmental or social impacts of proprietary research. Some research that lacks commercial application could be neglected entirely.

As noted earlier, with increased focus on knowledge and technology as intellectual property, particularly in the biological arena, there has been an enormous increase in patents, licensing, and material transfer agreements. Many analysts suggest that these new practices and processes may impede or limit the pace and direction of scientific efforts, restrict scientific communication, or undermine an academic scientist's ability to carry out research. The potential restriction of communication is particularly true of university scientists with private sector grants, who often must delay public discussion of work, or its results, pending review by the sponsoring company. Even some scientists with public funding feel inhibited about discussing their work, for fear that some private company with the money, equipment, and time will utilize their ideas and perform the experimental work before they can (29). Companies sponsoring university biomedical research often ask scientists to go beyond the standard secrecy requirements needed to obtain a patent for products related to their research. While NIH calls for a delay of only one or two months while an application is filed, 58 percent of the companies in a recent survey ask researchers to keep data secret for more than six months (31). The net effect of these various developments appears to be a reduction of the free flow of information. Many have argued that open communication and the freedom of thought is the best path towards realizing the social benefit from science (12).

A final impact involves potential conflicts of interest and/or scientific misconduct. In interviews, public and private sector scientists alike stress the potentially detrimental effects of restrictive agreements between the universities and corporations. These effects include favoritism, unwarranted financial advantages through privileged use of information or technology derived from the publicly funded research, and shelving of research of interest to the public but not to the corporation (2). In a recent article entitled "University–industry research must get closer scrutiny," Mildred Cho observed that "one major reason for concern is that if faculty members are profiting financially from their research either through royalties from, or as investors in, companies that market products based on their discoveries, the outcome or direction of their work may be affected. They might, for example, be tempted (consciously or unconsciously) to design studies that are more likely than not to have an outcome favorable to the product" (31, p. B4). A recent study (32), for example, designed to measure how drug company money might influence scientists, points directly toward a need for disclosure of industry relations and funding sources. This study found that 96 percent of the researchers who wrote favorable articles about a controversial class of drugs for treating hypertension and angina also had financial ties to the marketers of those drugs. In contrast, among those who published articles critical of the drugs, only 37 percent had financial ties. Conflicts were disclosed in only 2 of the 70 papers. As one researcher at George Washington Medical Center noted, researchers like to think that they are not influenced by their financial ties, "but the pressures may be too subtle for them to realize" (33, p. A41). These divided loyalties and conflicts of interest betray the public trust. According to Krimsky, the most significant social consequence of change within scientific institutions is "the disappearance of a critical mass of elite, independent and commercially unaffected scientists to whom we turn for vision and guidance when we are confronted by technological choices" (3, p. 79). Thompson (12) has further noted that as little as the public might care about the institutional effects of biotechnology within science, they may well be among the most far-reaching.

DEVELOPING COUNTRIES AND THE GLOBAL ECONOMY

The new technologies offer the hope of increasing crop yields where population growth is outstripping the food supply. Microbiology in conjunction with plant propagation and breeding is already creating more drought and resistant varieties of cassava, oil palms, and groundnuts. Yet despite biotechnologies' great promise for feeding the world's rapidly growing population, particularly in developing countries, science and policy

makers admit it will not be easy to ensure that this technology has the desired positive effects. Much that has been said about the social consequences for family farming and rural communities in industrialized countries applies more dramatically to the resource poor farmers in developing countries. Several analysts have predicted that biotechnology will have an unfavorable impact on the rural poor in Africa, Asia, and Latin America while benefiting relatively better-off farmers in those regions. As farms become larger and fewer, more people both in absolute numbers and as a percentage of the population in developing countries are being affected. Those who are affected are much worse off to begin with and are more vulnerable to displacement.

There is legitimate concern arising that the developed nations will use the new technology to undercut traditional Third World exports, such as vanilla, sugar, cocoa butter, and other important cash crops. Genetic engineering processes are being used to transform the production of certain agricultural commodities into industrial processes. In principle, any commodity that is consumed in an undifferentiated or highly processed form, could be produced using new biotechnological processes, and product substitutions could be easily introduced. Similarly, although with greater difficulty, tissue culture techniques could be used to produce edible plant parts in vitro. Several companies are now capable of final production of a natural vanilla product in the laboratory. A genetic modification of oilseed plants to convert cheap oils (e.g., palm or soybean oil) into high-quality cocoa butter is well advanced. Biotechnology is also being used to produce substitutes for sugar as an industrial sweetener. Even moderate success in realizing these product substitutions would have profound effects around the world, most immediate and important would be the restructuring of global markets (34–36).

Another concern is that biotechnology will increase the disparities between the developed and developing nations. With the shift in applied research and associated product development from the public to the private sector, the benefits from the new biotechnologies may become less widely available. Furthermore the products developed are unlikely to be ones that are important to the poor developing countries, particularly in the tropics. Only a small amount of the estimated $2.5 billion dollars of research spending on agricultural biotechnology around the globe is carried out in the developing world. According to Robert Herdt, from the Rockefeller Foundation, between $50 and $75 million per year is spent on agricultural biotechnology in the developing world, with about half of that conducted by the Consultative Group for International Agricultural Research (CGIAR) centers. However, the financial support for that system has weakened in significant part because of declining U.S. support. While the International Service for Applied Agricultural Agrobiotechnology data on field trials of genetically engineered crops reveal over 3700 trials of genetically engineered crops through the end of 1995, none of the field trials to date has been directed specifically at increasing output. Instead, most of the work has been done on transforming crops to be herbicide resistant (40 percent) and insect resistant (22 percent). Research on some of the traits most needed in the developing world such as the ability to tolerate low soil fertility, the ability to tolerate soil salinity or alkalinity, and techniques for producing biological pesticides has gone unstudied. This could further widen the gap between the agricultural production methods in the North and the less developed practices in the South (37).

A further concern revolves around the controversy over the property rights in genetic resources. Biotechnology is playing a key role in conserving genetic diversity worldwide at the same time as it is accelerating the privatization of these genetic resources. Biotechnology is providing new incentives to patent commercially valuable genetic resources as well as providing the means to both enhance those resources and protect them from patent infringement. It has been argued that native genetic resources (i.e., germ plasm or seeds) are owned by indigenous farmers, by their governments, or collectively, by the whole society and considered the common heritage of humankind. Intellectual property rights may deprive farmers in developing countries of something they currently have. Farmers are losing the right to plant seed freely from land races or other publicly available varieties (12). This is probably unlikely since most legal codes protect any existing uses of the raw materials from which new seed varieties or plants are derived. For example, the convention on Biological Diversity, which came into force in December 1993, provides an internationally legally binding instrument that explicitly recognizes national sovereign rights over the genetic resources existing within a country's territory (38). However, although indigenous farmers may have a legal right to use plants in traditional ways, they lack the resources and knowledge to protect those rights. Moreover many countries, especially developing countries rich in genetic diversity, have called for measures to protect their interests and to insure that they share in the benefits derived from the use of their resources by others (39). Issues of ownership, access to genetic resources, and intellectual property protection of genetically engineered products are currently receiving considerable attention in various international forums. These complex and highly politically charged debates will likely continue into the future.

Finally, many developing countries have no basic and limited applied research capacity, marginal capabilities to adapt biotechnological advances to local conditions, and few resources to attract transnational corporations to conduct their own research. In conclusion, agricultural biotechnology may shift the geographic location of agricultural production from one Third World country to another or from the Third World to the First World. The literature on social consequences of agricultural biotechnology for developing countries includes very little in the way of detailed ex ante studies. However, for many Third World countries that are dependent on one or two agricultural commodities for their continued viability, this production and market restructuring, and increased productivity gaps, could result in a collapse in existing markets. Significant numbers of farmers and farm workers could find themselves with no products to

sell. This could increase the already high Third World debt and exacerbate the deficit imbalance of payments in Third World countries. If this were to occur, political instability, already a problem in the developing world, would doubtless increase.

STRATEGIES FOR INCORPORATING THE FOURTH CRITERION

This article has discussed a number of both positive and negative social and economic impacts biotechnology may have on (*1*) farmers, rural communities, and the food system; (*2*) the structure and organization of agribusiness and industry; (*3*) consumers; (*4*) science and technology transfer; and (*5*) developing countries and the global economy. Genetic engineering and biotechnology are areas in which nontechnical decision-making inputs, on initially technical issues, are found to be increasingly important. It was noted that reduction of risks to health and environment are usually emphasized as the task of any regulation but that goals based on social and ethical principles also appear within the scope of genetic engineering laws and regulations. Indeed, it is argued that the health and environmental risks, as well as the socioeconomic hazards, are difficult to separate. Moreover the issues of risk to environment, life and limb appear to be a necessary but inadequate conditions for public acceptance. Although the socioeconomic criteria, often referred to as the fourth hurdle or fourth criterion, have been clearly rejected in licensing, it is equally clear that socioeconomic criteria are being taken into account in numerous contexts (e.g., the Austrian Genetic Engineering Act; the Norwegian Gene Technology Act; the UNEP Technical Guidelines for Safety and Biotechnology; release applications for genetically engineered organisms in Denmark, Sweden, and Finland; the Swiss Gene Protection Initiative).

As a consequence a number of alternatives have emerged for incorporating the socioeconomic issues into a broader public discussion and eventually into decision making. One process or procedural approach described by Thompson (12) is discourse ethics where certain morally relevant constraints on discourse must be met for it to be reasonably successful. First, discourse must be open to all competent speakers whose interests will be affected. Second, people must be free to construe the issue and their own interests in whatever terms they deem appropriate. Third, participants must be free of rigid inflexibility that precludes them from adopting a hypothetical stance toward their own and others' interests and values. Fourth, the process must be free of external coercion and, fifth, statements must be focused exclusively on establishing the best reasons for accepting a prescription or conclusion (12). Thompson admits that since these conditions are rarely met, actual public debates over biotechnology are unlikely to reach an ethically defensible consensus. However, the debates will be greatly enhanced if informed by efforts to approximate these ideal discourse conditions — if not in a public forum, then at least under some controlled circumstances in which these issues can be seriously pursued.

Seifert and Torgensen (6) also examined various mechanisms whereby socioeconomic criteria can be included in the decision-making process for new technologies such as genetic engineering. Their discussion of the Austrian efforts focused on incorporating social sustainability in consideration of regulations for genetic engineering. They suggested three forms to enhance broader participation in the debates. The first is discursive meditation whose main emphasis is on clarifying controversial issues. The best example of this process in the United States is the annual public forum conducted by the National Agricultural Biotechnology Council that brings together diverse participants and stakeholders in the arena of agricultural biotechnology to discuss and clarify concerns surrounding agricultural biotechnology. Consensus conferences are a second form that serves the task of a development of a political will. Institutional innovations to provide for independent mediation among competing positions on issues like biotechnology is a third form. For example, the Dutch consumer protection organization Consumer and Biotechnology is an example of such an institutional innovation. The function of this institution is to mediate among nongovernmental organizations, the biotechnology industry, the authorities, agricultural associations, and consumer organizations.

All these processes and procedures, however, have two fundamental problems. First, the results are generally not binding and require the approval of the decision makers for their implementation. To partially overcome this limitation in Denmark, for example, the institution of consensus conferences was established in association with the parliament so that the results of the procedures and process could be considered in parliamentary decision making as quickly as possible. The second problem concerns the vulnerability of key parties to manipulate the results. In all the processes there is a risk of being coopted or overwhelmed by powerful interests and ultimately serving only as justification for the implementation of these new technologies. Therefore the goals and motives of the organizers, the credibility of spokespersons who both raise and discuss the range of issues surrounding biotechnology, the funding sources for these procedures and processes, the process by which content, scope, and audience for these processes are made, and how this information and the outcomes of the conferences will be pursued must be kept as transparent as possible.

Barling and his colleagues (40) in examining social aspects of food and agricultural biotechnology in Europe have observed that public concern about these new technologies are primarily focused on issues of trust, choice and care for a sustainable society of natural balance. They have recommended improving consumer choice and promoting greater public involvement in decision making. Complementing labeling, they propose transparency right through the food chain. This would entail the use of a comprehensive system of segregation and certification of genetically modified crops and their products from nongenetically modified crops at each stage of the food chain, and for this to be reflected in the final labeling information. This degree of transparency, they argue, would allow consumers to make more fully informed choice

of foodstuffs in line with their more deeply felt values on such issues and would provide for a more democratic and participatory basis for transparency. Finally, in attempting to predict the effects of the use of genetically modified organisms in agriculture and food production, they propose integrating the precautionary principle more actively into risk management. As noted earlier, the precautionary principle is applied in circumstances of scientific uncertainty reflecting the need to take action in the face of potentially serious harm in the absence of scientific proof. As a consequence the precautionary principle is not simply a matter of science but is socially and politically informed. They argue that the incorporation of wider social concerns, as articulated by different social actors, should be included in risk analysis to produce a more socially embedded and accepted process of risk analysis of the applications of modern biotechnology. This would build greater trust in and acceptance of the regulatory process and provide a more socially responsible, plural, and accountable form of decision making. Implementation, of course, remains quite complex and highly problematic.

While efforts to enhance the dialogue and to ensure wider participation in the debates is laudable, Youngberg (41) at a National Agricultural Biotechnology Council annual meeting challenged the participants to move beyond preoccupation with the dialogue process and to begin to explore new and innovative ways to involve the broader society such as the sustainable agriculture community in the biotechnology decision-making process itself. He noted that there is a critical difference between broad participation in the dialogue about biotechnology, and the actual involvement in planning and decision-making phases of agricultural biotechnology research development and the introduction of these technologies into the marketplace. He continued by suggesting that the time has come for the biotechnology industry to begin exploring the principles and processes of participatory decision making and to initiate a serious assessment of ways to implement concrete decision-making opportunities involving all elements of the agricultural biotechnology constituency, including farmers, public interest group representative, and other citizens. He concluded by noting that the dialogue offers only sporadic, short-term opportunities for interaction while ongoing relatively intimate interactions characteristic of participatory decision making would create authentic opportunities to directly influence the biotechnology agenda. This action would likely create greater trust, result in more comprehensive and enlightened planning that includes meaningful consideration of the socioeconomic issues, potentially save money, time, and resources, and make possible endorsement not mere acceptance of the new agricultural biotechnologies.

CONCLUSION

Although introduced as the fourth criterion, it may be more appropriate in evaluating public research agendas as well as the regulation and approval of new agricultural biotechnology products and processes to consider the broader socioeconomic effects as the first criterion. As most scientists and policy analysts acknowledge, biotechnologies are the tools and means to achieve particular socioeconomic goals. These biotechnologies are options, albeit compelling options, among many and involve choices. As such they should be framed and evaluated in terms of local, regional, national, and international social goals and values. In the final analysis the key question is not how we learn to accept and live with the new biotechnologies but rather under what conditions, price, costs, gains, and benefits on a personal, community, national, and global scale. In a democracy the public has an obligation and a right to be informed, to participate, and to shape the developments of technology in terms of the broader socioeconomic values of their respective society. The effective public representation of the public interest also ensures that society avoids the potential abuses of power by those with vested interests (16). In the case of agricultural biotechnology as we have seen, the public is increasingly exercising their obligations and rights.

Thompson concludes his thoughtful analysis of the ethics of food biotechnology by looking at the relationship among science, trust, and democracy. He notes that to the extent "that democracy is understood as a form of government distinctive for its receptivity to participation and resting upon consent of the governed, the events that turn ordinary people into enemies of science can be seen to compromise government, rather than science" (12, p. 237). How we balance scientific criteria embedded in traditional risk analysis and the social criteria embedded in the fourth criterion may well determine the extent to which the future of biotechnology is able to live up to its promise as the first important science of the twenty-first century.

BIBLIOGRAPHY

1. J. Rifkin, *The Biotech Century: Harnessing the Gene and Remaking the World*, Tarcher/Putnam, New York, 1998.
2. L. Busch, W.B. Lacy, J. Burkhardt, and L.R. Lacy, *Plants, Power and Profit: Socioeconomic and Ethical Consequences of the New Biotechnologies*, Basil Blackwell, Oxford, UK, 1991.
3. S. Krimsky, *Biotechnics and Society: The Rise of Industrial Genetics*, Praeger, New York, 1991.
4. P.C. Stern and H.B. Fineberg, eds., *Understanding Risks: Informing Decisions in a Democratic Society*, National Academy Press, Washington, DC, 1996.
5. G. Weber, Extension Biotechnology Coordinator's Memo, USDA-ES, Washington, DC, March 24, 1990.
6. F. Seifert and H. Torgersen, *Public Understand. Sci.* **6**, 301–327 (1997).
7. T. Crompton and T. Wakeford, *Nature Biotechnol.* **16**, 697–698 (1998).
8. G. Schatz, *Science* **281**, 810–811 (1998).
9. U.S. Executive Office of the President, *Use of Bovine Somatotropin BST in the United States: Its Potential Effects*, Washington, DC, 1994, pp. 35–36.
10. U.S. Department of Agriculture, *A Time to Act: A Report of the USDA National Commission on Small Farms*, USDA, MP-1545, Washington, DC, 1998.
11. S. Smith, *Choices* **7**(1), 1–10 (1992).
12. P. Thompson, *Food Biotechnology in Ethical Perspective*, Blackie Academic and Professional, London, 1997.

13. J.J. Kloppenburg, J. Henrickson, and S. Stevenson, *Agric. Hum. Values* **13**(3), 33–42 (1996).
14. M.E. Gertler, in R.W.F. Hardy, J.B. Segelken, and M. Voronmaa, eds., *Resource Management in Challenged Environments*, National Agricultural Biotechnology Council, Ithaca, NY, 1998, pp. 137–145.
15. W. Berry, in W. Vitek and W. Jackson, eds., *Rooted in the Land: Essays on Community and Place*, Yale University Press, New Haven, CT, 1996, pp. 76–84.
16. R. Hindmarsh, G. Lawrence, and J. Norton, in R.G. Hindmarsh, G. Lawrence and J. Norton, eds., *Altered Genes: Reconstructing Nature, the Debate*, Allen & Unwin, St. Leonards, Australia, 1998, pp. 3–23.
17. Genetic Engineering News, *Guides to Biotechnology Companies*, GEN Publishing, Larchmont, NY, 1998.
18. D. Strickland, *Genetic Eng. News* **19**(21), 1, 12, 62, 72 (1999).
19. D. Ferber, *Science* **286**(5448), 1662–1666 (1999).
20. M. Lappe and B. Bailey, *Against the Grain: Biotechnology and the Corporate Takeover of Your Food*, Common Courage Press, Monroe, ME, 1998.
21. M. Pollan, *N.Y. Times Mag.* October 25, 1998, pp. 44–51, 62, 63, 82, 92, 93.
22. M. Enerink, *Science* **286**(5448), 1666–1668 (1999).
23. S. Krimsky and R. Wrubel, *Agricultural Biotechnology in the Environment: Science, Policy, and Social Issues*. University of Illinois Press, Urbana, IL, 1996.
24. *Nature* **398**(6729), 639 (1999).
25. Biotechnology and the European Public Concerted Action Group, *Nature* **387**, 845–847 (1997).
26. W.B. Lacy and L. Busch, in J.F. MacDonald, ed., *Agricultural Biotechnology at the Crossroads: Biological, Social and Institutional Concerns*, National Agricultural Biotechnology Council, Ithaca, NY, 1991, pp. 153–168.
27. W.B. Lacy, L.R. Lacy, and L. Busch, in M.C. Hallberg, ed., *Bovine Somatotropin and Emerging Issues*, Westview Press, Boulder, CO, 1992, pp. 3–32.
28. G. Blumenstyk, *Chron. Higher Educ.*, December 11, 1998, p. A56.
29. W.B. Lacy, in S. Wolfe, ed., *Knowledge Generation and Transfer: Implications for the 21st Century*, forthcoming.
30. D. Blumenthal, E.G. Campbell, N. Causino, and K.S. Locus, *N. Eng. J. Med.* **335**, 1734–1739 (1996).
31. M.K. Cho, *Chron. Higher Educ.*, August 1, 1997, pp. B4–B5.
32. H.T. Stelfox, G. Chua, K. O'Rourke, and A.S. Detsky, *N. Eng. J. Med.* **338**(2), 101–106 (1998).
33. G. Blumenstyle 1998b. *Chron. Higher Educ.*, May 22, 1998, pp. A40–A41.
34. N.P. Barbosa, *Int. J. Sociol. Agric. Food* **4**, 161–168 (1996).
35. R.M.A.A. Galhardi, *Futures* **27**(6), 641–656 (1995).
36. H.J. Shand, in J.F. MacDonald, ed., *Agricultural Biotechnology and the Public Good*, National Agricultural Biotechnology Council, Ithaca, NY, 1994, pp. 73–86.
37. R. Herdt, in R.W.F. Hardy, J.B. Segelken, and M. Doionmaa, eds., *Resource Management and Challenged Environment*, National Agricultural Biotechnology Council, Ithaca, NY, 1998, pp. 33–40.
38. G. Hawtin, in R.W.F. Hardy, J.B. Segelken, and M. Doionmaa, eds., *Resource Management in Challenged Environments*, National Agricultural Biotechnology Council, Ithaca, NY, 1998, pp. 113–121.
39. W.B. Lacy, *Crop Sci.* **35**(2), 335–345 (1995).
40. D. Barling et al., *Environ. Toxicol. Pharmacol.* **7**, 89–93 (1999).
41. G.I. Youngberg, in J.F. MacDonald, ed., *Agricultural Biotechnology and the Public Good*, National Agricultural Biotechnology Council, Ithaca, NY, 1994, pp. 147–150.

See other entries AGRICULTURAL BIOTECHNOLOGY; ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE.

ANIMAL BIOTECHNOLOGY, LAW, FDA REGULATION OF GENETICALLY MODIFIED ANIMALS FOR HUMAN FOOD USE

STUART M. PAPE
PAUL D. RUBIN
Patton Boggs, LLP
Washington, District of Columbia

OUTLINE

BACKGROUND

Since 1981, when scientists at Ohio University were credited with producing the first genetically modified animals by transferring genes from other animals into mice, there have been significant scientific advancements in the field of animal biotechnology. Despite these advancements, genetically modified animals intended for human food use have not yet been commercially distributed in the United States. The use of animal biotechnology for human food use, however, is expected to become prevalent within the next decade.

To date the primary focus of biotechnology policy has been directed toward agricultural biotechnology

products — which have been marketed in the United States for a number of years. Agricultural biotechnology products have been highly controversial and have been subject to significant criticism from a variety of interest groups, scientists, and members of the public (particularly in Europe). Many observers expect the controversy and criticism directed toward genetic modification of animals to be even more intense.

Although not actually involving genetic manipulation of animals, the public's reaction toward the use of recombinant bovine somatotropin (rBST), also known as recombinant bovine growth hormone (rBGH), in cows is instructive (1). In 1993, after a comprehensive review of safety and efficacy, the Food and Drug Administration (FDA) approved a new animal drug application (NADA) for a product, called Posilac, that contained rBST. The genetically engineered hormone, which was found by FDA to be identical to natural pituitary-derived bovine growth hormone, was approved for injection into cows to increase milk production. (Although the cows received the rBST injections, they were not genetically modified (2).)

Subsequently, based on the alleged impact on cow health and milk, numerous challenges were made to FDA's finding that food products from cows treated with rBST are safe for human consumption (3–6). Challengers argued that the possible adverse health effects of Posilac were not addressed, in part because long-term toxicology studies to ascertain human health safety were not required by FDA or conducted by the NADA applicant, Monsanto.

In response to these challenges, the FDA conducted a comprehensive audit of the human food safety sections of the NADA supporting the drug approval (2). The audit reviewed all of the studies relied upon to determine the human food safety of rBST. FDA concluded that an examination had not been performed on antibody data during the course of the original review of the Monsanto application (2). FDA subsequently reconsidered all of the studies and concluded that there were no new scientific concerns regarding the safety of milk derived from cows treated with the drug (2). The determination that long-term studies were not necessary for assessing the safety of rBST was based on studies that demonstrated that rBST is biologically inactive in humans even if injected, and that rBST and pituitary-derived bovine growth hormone are biologically indistinguishable.

The public's reaction toward the use of rBST in cows provides evidence of the extensive controversy that may result when changes are made to animals that affect the human food supply. Not surprisingly, the controversy is expected to be even more intense if genetic modifications are involved.

Compared with agricultural biotechnology, however, the visceral reaction opposed to the use of biotechnology techniques may be tempered in the United States by the fundamental difference between the current regulation by FDA of agricultural biotechnology and genetically modified animals intended for human food use. Whereas most genetically modified agricultural products are currently only subject to a voluntary notification system by FDA, genetically modified animals intended for human food use are subject to extensive agency regulation. FDA is currently reevaluating its policy toward agricultural biotechnology and is considering adopting a mandatory review policy. The mandatory and arduous nature of animal biotechnology regulation may satisfy the "safety" concerns raised by critics of biotechnology as applied to agricultural products, and could conceivably mitigate opposition to genetically modified animals. Other issues, however, including religious and/or ethical-based concerns, could subject animal biotechnology to the same public backlash that has recently plagued agricultural biotechnology.

BENEFITS OF ANIMAL BIOTECHNOLOGY FOR HUMAN FOOD USE

Biotechnology Overview

The Genetic Code. Living organisms contain cells that contain DNA (deoxyribonucleic acid) in their chromosomes. DNA contains the genetic code, or the genome, for an organism. The genetic code is derived from a four-letter alphabet, A,C,G,T (Adenine, Cytosine, Guanine, and Thymine), and based on the sequence and number of genes, the hereditary traits and characteristics of the living organism is determined.

The chemical and physical composition of DNA does not vary from organism to organism. It is only the sequence and number of letters in the genome that create differences between different living organisms — the physical and chemical composition of the actual DNA, and letters themselves, are constant from organism to organism. In practice, this means that DNA from any organism is capable of functioning even if it is transferred into a different organism.

Definition of Biotechnology. Biotechnology has been defined as the manipulation of the DNA molecules of living organisms. By means of selective breeding of animals and crops, humans have practiced basic biotechnology for centuries. The transfer of DNA occurs naturally through sexual reproduction and has been utilized in plant and animal breeding programs for centuries in order to alter the traits in living organisms.

Modern biotechnology techniques, however, are technologically superior to traditional breeding in that (*1*) precise genetic manipulation can alter specific genes while leaving others unchanged and (*2*) genes not native to an organism may be added from a distinguishable organism (i.e., a distinguishable species).

Intragenetic combinations involve the transfer of genes that are native to a species (e.g., adding an additional growth hormone gene that is ordinarily found in the species). Intergenetic combinations, which involve the addition of a nonnative gene from one species to another (e.g., adding a nonnative disease resistance gene from one species to another), allow the transfer of desirable traits found in nature from one organism to another. In other words, rather than relying solely upon conventional breeding, which can be structured to exploit traits that exist naturally within a breed, intergenetic combinations permit the breeder to add genes from outside the breed. Furthermore unlike conventional breeding, specific

genes can be added or modified within a breed without potentially altering other genes as well. In other words, biotechnology permits the transfer of DNA from one species to another; DNA may be exchanged between plants, animals, bacteria, or viruses in order to alter the genetic information contained in the genome.

Examples of Animal Biotechnology

Fish. Fish comprise a significant portion of the diet and are a major source of protein for people throughout the world. At present, the majority of animal biotechnology research is being conducted on fish. Seafood has been the focal point of animal biotechnology research due to the simpler biological make-up of fish compared with farm animals. Although no transgenic fish have yet been approved for food use in the United States, investigations of transgenic fish are being conducted throughout the world at the present time.

Research on seafood biotechnology is currently focused on a variety of genetic changes, including (*1*) improving the growth rate of fish, (*2*) increasing fish size, (*3*) improving the food conversion capabilities of fish, (*4*) improving the nutritional profile of fish, (*5*) altering the color, flavor, or texture of fish, (*6*) improving disease resistance of fish, (*7*) improving temperature resistance of fish, and (*8*) using fish as "biopharm animals" to create drugs or other chemicals for human use.

Meat and Poultry. In comparison with fish biotechnology, genetic modification of farm animals is still in its infancy. Genetic modification of farm animals is generally far more complex than for plants or fish due to the genetic complexity of the organisms, and difficulties with embryo transfer. Biotechnological developments ultimately may be capable of increasing the muscle mass of cattle (i.e., cattle with less fat), increase growth, improve digestive capabilities, increase disease resistance, and improve the nutritional profile of meat and poultry products (including eggs).

Meat and poultry are currently produced by numerous breeds and varieties based on years of domestication and selection. Genetic modifications, however, are capable of adding traits and characteristics that are currently impossible to develop under conventional breeding techniques. For example, it may be possible to genetically modify poultry to improve digestive function — such that the poultry would be capable of digesting lower-quality animal feeds (which are less expensive and readily available even in developing nations). Animals may also be genetically altered to produce a human or veterinary drug or biologic, food additive, or other product that can be harvested from the milk, blood, or other tissues of the animal (i.e., a "biopharm animal").

Animals may also be genetically altered to improve disease resistance to specified pathogens. Such genetic alterations could complement animal vaccination programs, and decrease the human health risk associated with ingestion of meat and poultry. In addition genetic modification could reduce or eliminate the need to use antibiotics, and thus may help to address the potential problem of antibiotic resistance.

Finally, poultry could also be genetically modified to improve the nutritional profile of whole eggs for human ingestion. For example, in the near future it may be possible to create eggs with high levels of protein and lower levels of cholesterol.

REGULATORY CATEGORIES OF ANIMAL BIOTECHNOLOGY

For regulatory purposes FDA has characterized uses of animal biotechnology into three major categories, each subject to special regulatory issues: (*1*) biopharm animals, (*2*) somatic cell therapy, and (*3*) transgenic gene modification.

Biopharm Animals

Biopharm animals are animals that have been genetically modified to produce a human or veterinary drug or biologic, food additive, or other product that can be harvested from the milk, blood, or other tissues of the animal (7). The genetic modification is designed to harness the metabolic capabilities of the animal to produce a product in lieu of using chemical synthesis or other traditional production methods (7). Although biopharm animals that are salvaged may also end up in the human food supply, the vast majority of biopharm animals are not in themselves intended for human food use. Rather, only the products harvested from the milk, blood, or other tissues of the animal are ordinarily intended for human use.

In general, products derived via biopharm animals will be regulated by the regulatory agency with experience in regulating that type of derived product, regardless of the breed of the biopharm animal or the method used to genetically modify the animal (8).

For example, if a genetically modified animal produces a human biologic (which is expected to account for the majority of products derived from biopharm animals), FDA's Center for Biologics Evaluation and Research (CBER) would conduct the safety and efficacy review (although if the biopharm animal were also used for human food use, FDA's Center for Veterinary Medicine (CVM) would be expected to consult with CBER regarding food salvage and safety issues).

A guidance document issued by CBER in 1995, entitled *Points to Consider in the Manufacture and Testing of Therapeutic Products for Human Use Derived from Transgenic Animals*, provides an overview of the FDA regulatory considerations associated with biopharm animals. The guidance document outlines FDA's concerns with regard to the use of transgenic animals to produce FDA-regulated drugs and biological products for human use.

Among the issues CBER will evaluate are (*1*) the generation and characterization of the transgene construct, (*2*) creation and characterization of the transgenic founder animal, including genetic stability and expression, (*3*) establishment of a reliable and continuous source of transgenic animals, (*4*) generation and selection of the production herds, (*5*) maintenance of the transgenic animals, including monitoring, feeding, and use of by-products, and (*6*) purification and characterization of the transgenic product.

Somatic Cell Therapy

Somatic cell therapy involves treating somatic cells, cells of the body that compose the tissues and organs, with new DNA to change the function of the recipient somatic cell. Somatic cell therapy may be accomplished via individual animal injections that modify specified cells in the body in order to express a protein, hormone, or enzyme. Individual steers, for example, could be modified to produce more muscle mass without having to modify the breeding herd (thereby avoiding calving difficulties that might be caused by additional muscle mass in a brood cow) (9). Somatic cell therapy does not ordinarily change the heritable traits of the animal.

Somatic cell therapy is expected to be ordinarily regulated by the FDA and subject to NADAs, unless the purpose is to prepare animals to be used in the production of regulated biopharm animal products (8). In addition, if the genetic modification more appropriately falls within the regulatory category of a food additive or color additive, FDA is expected to review the product under its food-related regulatory requirements rather than its drug-related requirements.

Transgenic Breed Modification

Transgenic breed modification involves germ line modifications made to affect growth characteristics or quality of food products derived from the target animal. The modifications are made to eggs or sperm and are heritable. FDA has indicated that animals derived from traditional breeding and selection, including artificial insemination and in vitro fertilization (IVF), would be excluded from the definition of transgenic breed modification animals (8).

As with somatic cell therapy, transgenic breed modifications are expected to be ordinarily regulated by the FDA and subject to NADAs. As with somatic cell therapy, if the modifications more appropriately fall within the food additive or color additive requirements, FDA is expected to review the applicable products under those regulatory requirements rather than its drug regulatory requirements.

U.S. REGULATORY AGENCIES: OVERVIEW

The FDA is the primary U.S. regulatory agency responsible for regulating genetically modified animals intended for human food use. Although other regulatory agencies, such as the U.S. Department of Agriculture (USDA) and Environmental Protection Agency (EPA) may share jurisdiction over animal biotechnology in certain circumstances, FDA will be the focal point in the United States for establishing regulatory policy over animal biotechnology.

FDA is composed of five separate centers, and the center responsible for regulating a specific product derived via animal biotechnology will vary depending on the product type rather than the process used to create the genetic modification. It is anticipated that CVM will be responsible for regulating the majority of products derived via animal biotechnology. CVM is responsible for regulating "animal drugs," which are defined as including most products intended to improve (*1*) animal growth, (*2*) animal feed efficiency and digestive capabilities, (*3*) animal carcass characteristics, and (*4*) animal disease resistance (i.e., an antibiotic effect).

Other centers within the FDA, however, may have primary responsibility for regulating certain types of animal biotechnology products. It is anticipated that FDA's Center for Food Safety and Applied Nutrition (CFSAN), for example, would have primary responsibility for regulating animal biotechnology intended to improve the nutritional profile of food used for human use. Genetically modified fish, such as that intended to increase the level of omega-3 fatty acids present in the fish (in order to improve the nutritional profile of the fish when ingested by humans), would likely be regulated as a "food additive" by CFSAN. It is anticipated, however, that CVM may also play a significant role in reviewing the product due to CVM's expertise in evaluating animal health issues. CFSAN would also be expected to have primary responsibility over animal biotechnology used to improve the color profile of animals intended for human use. It is anticipated that salmon, genetically modified to contain increased pink muscle tissue, would primarily be regulated by CFSAN as a "color additive."

FDA's Center for Drug Evaluation and Research (CDER), which is primarily responsible for reviewing the safety and efficacy of drugs intended for human use, would have primary jurisdiction over drugs produced for human use from biopharm animals (i.e., animals genetically modified to produce drugs for human use that are harvested from animal milk, blood, or other tissue). Similarly CBER, which is primarily responsible for reviewing the safety and efficacy of biologics (e.g., vaccines and many products derived via biotechnology) intended for human use, would have primary jurisdiction over biologics produced for human use from biopharm animals.

USDA's Animal and Plant Health Inspection Service (APHIS) is responsible for regulating animal "biologics." It is therefore anticipated that a genetic modification of an animal intended to produce a vaccine-type response to a disease in the animal (an immune-response) would be primarily regulated by APHIS. If applicable, APHIS would also conduct a food safety review for animal vaccines.

For products of animal biotechnology regulated by CDER and CBER, however, if the genetically modified animals are also intended for human food use, CVM would consult with the appropriate FDA center regarding the safety of the genetically modified animal for human food use.

Finally, numerous federal and state regulatory agencies, including the U.S. Fish and Wildlife Service and EPA, may have partial regulatory responsibility to review the environmental effects of animal biotechnology (e.g., introduction of genetically modified fish into the environment). For animal biotechnology applications that are managed by CVM, it would be expected that the assessment of potential environmental effects of such products would be coordinated by CVM under the National Environmental Policy Act (NEPA). Table 1 identifies the regulatory agencies expected to assert primary jurisdiction over various types of animal biotechnology products.

Table 1. U.S. Regulatory Agencies: Expected Primary Jurisdiction Over Biotechnology Products

Type of Product	Primary Agency Jurisdiction
Biopharm animal	
Produces a human drug	FDA, CDER
Produces a human biologic (e.g., vaccine)	FDA, CBER
Produces a food additive for use in human food	FDA, CFSAN
Produces a color additive for use in human food	FDA, CFSAN
Produces an animal drug	FDA, CVM
Produces an animal biologic (e.g., vaccine)	USDA, APHIS
Somatic cell therapy or transgenic gene modification	
Increases animal muscle mass	FDA, CVM (animal drug)
Increases animal growth	FDA, CVM (animal drug)
Reduces the amount of fat present in the animal	FDA, CVM (animal drug)
Improves digestive capabilities of the animal	FDA, CVM (animal drug)
Improves animal disease resistance via a vaccine antibody/antigen response	USDA, APHIS (animal biologic)
Improves the nutritional profile of an animal for improved nutrition upon ingestion by humans[a]	FDA, CFSAN (food additive)
Modifies the color of the animal for improved appearance for human food use[b]	FDA, CFSAN (color additive)

[a]One example would be the genetic modification of fish to increase omega-3 fatty acid content.
[b]One example would be the genetic modification of fish to increase the amount of pink muscle, which would be more aesthetically pleasing when intended for human food use. This would be a complex issue, however, as the "color additive" review would not assess animal health. Accordingly, it would not be surprising if CVM consulted with CFSAN regarding animal health issues. In addition, FDA could attempt to regulate such fish via the NADA process.

FDA REGULATION OF BIOTECHNOLOGY

Background: FDA Regulation of Biotechnology in General

Congress has not enacted any new statutory provisions specifically governing products derived via biotechnology processes. In 1986 the FDA issued a General Statement of Policy on Biotechnology (10) that indicates that because FDA regulates products rather than processes, products of biotechnology may be regulated under existing statutory authority. The Policy established the following general principles that should be followed in determining the safety of food produced by biotechnology (11):

1. The cloned DNA, as well as the vector DNA, should be properly identified.
2. The details of construction of the production organism should be available.
3. There should be information documenting that the inserted DNA is well characterized (i.e., the exact nucleotide sequence of the insert and any flanking nucleotides should be characterized) and is free from sequences that code for harmful products.
4. Food produced should be purified, characterized, and standardized.

The Policy also included a variety of considerations that should be evaluated for determining the safety of food produced by microbial isolation that has been genetically manipulated (11):

1. The microbial isolate used for production should be identified taxonomically, and if the strain of the isolate has been genetically manipulated, it should be determined whether each strain contributing genetic information to the production strain is identified.
2. The cultural purity and the genetic stability of the isolate should be maintained.
3. Fermentation should be performed with a pure culture and monitored for purity.
4. It should be determined whether the microbial isolate used for production also produces antibiotics or toxins.
5. It should be determined whether the isolates are pathogenic.
6. It should be determined whether viable cells of the production strain are present in the final product.

In addition a 1992 federal government oversight document confirms that federal government regulation should focus on the characteristics and risks of the biotechnology product — not the process by which it is created (12). Where oversight is warranted, the extent and type of oversight measure(s) must be commensurate with the type of risk being addressed, must maximize the net benefits of oversight by choosing the oversight measure that achieves the greatest risk reduction benefit at the least cost, and must consider the effect that additional oversight could have on existing safety incentives (12).

FDA Regulation of Genetically Modified Animals Intended for Food Use

Overview. The FDA has determined that existing FDA regulatory requirements are capable of ensuring the safety and efficacy of genetically modified animals intended for food use. FDA officials have noted that regulatory

determinations will focus on the resulting product of the biotechnology method, rather than the process (9), and that, accordingly, most genetically modified animals intended for human food use will be regulated as new animal drugs (9).

In 1994 FDA published a status report summarizing proposed regulatory approaches and issues for FDA regulation of animal biotechnology (8). The status report identifies several considerations underlying FDA's oversight of animal biotechnology products. First, FDA will seek consistent regulation of similar products. Second, the use of existing statutory authority and administrative processes is expected to save resources, offer a measure of consistency, and minimize disruption by taking advantage of many existing regulations, FDA's scientific surveillance staff, and existing FDA guidelines. Third, a mechanism should be established to inform the public which animal biotechnology products are regulated by FDA. There is no centralized program currently regulating investigations and field trials of transgenic animals, though there is a program for transgenic plants. Fourth, clear lines are needed to define those types of animal biotechnology where governmental oversight is required and those where it is not. Fifth, scientific flexibility will be required by FDA, particularly with regard to safety assessments.

FDA Regulation of Animal Drugs. The vast majority of genetically modified animals intended for human food use are expected to be regulated by FDA as animal drugs. As noted, FDA focuses on the effect of a product in determining regulatory jurisdiction, rather than the process used to produce the product. For example, growth hormones may be delivered to animals via injection, somatic cell therapy, or transgenic breed modification. Regardless of which method is used, the animals receive additional growth hormone. Accordingly, each of these methods would be regulated by FDA under its animal drug regulatory requirements (9).

Drugs are defined under the FFDCA as including (*1*) articles intended for use in the diagnosis, cure, mitigation, treatment or prevention of disease in man or other animals (13), and (*2*) articles (other than food) intended to affect the structure or function of the body of man or other animals (e.g., production drugs and hormones, anesthetics, contraception drugs). Animal drugs do not include vaccines designed to prevent animal disease (which are regulated as veterinary biologics by APHIS), or food or color additives (e.g., genetic modifications that change the color of fish, or improve the nutrition of animal meat for human food consumption). Finally, genetic modifications developed via traditional breeding techniques are not regulated by FDA as animal drugs.

It is interesting that animal clones are currently not expected to be regulated as animal drugs. If a clone is identical to a traditional animal (i.e., it is not a clone of a transgenic animal), FDA is not likely to assert jurisdiction since there would be no distinction between the cloned animal and the traditional animal. FDA's Center for Veterinary Medicine regulates products (not processes), and cloning is a process.

If, however, cloning is used to "produce" transgenic animals that produce an animal drug, CVM would regulate the production of the animal drug and require NADA approval. It is unclear, however, if FDA would regulate the safety of food derived from cloned animals that are not genetically modified. Theoretically the cloned animals would be indistinguishable from their noncloned parents — and therefore should not present a food safety concern (14).

For all animal drugs, FDA is responsible for evaluating (*1*) the safety and efficacy of the drug on the target animal, (*2*) labeling and promotional claims for the animal drug product, (*3*) environmental safety, (*4*) manufacturing and quality controls (ensuring that the product may be consistently manufactured to comply with established specifications under good manufacturing practices), and (*5*) the safety profile of the animal drug when provided to animals that are ultimately ingested by humans (e.g., toxicity and potential for adverse health effects).

Ordinarily, with regard to human food safety concerns based upon animal drug residues, FDA is expected to focus primarily upon the effect of potential chronic low level exposure to the drug residues. Drug residues are ordinarily not expected to produce acute toxicity in humans. For traditional animal drugs, FDA ordinarily relies upon toxicological studies to determine the "no observed effect level" (NOEL) — which is defined as the highest dose at which the drug produces no adverse effect. Based upon the NOEL, FDA utilizes a safety factor to establish an "acceptable daily intake" (ADI) — which is defined as the highest amount of drug residue that should be safely allowed in the edible tissues of the target animal.

As part of the NADA review process, FDA also stringently reviews all drug residues remaining in human food. For instance, an FDA regulation provides the following with regard to analytical methods used to identify and evaluate drug residues (15):

> Applications shall include a description of practicable methods for determining the quantity, if any, of the new animal drug in or on food, and any substance formed in or on food because of its use, and the proposed tolerance or withdrawal period or other use restrictions to ensure that the proposed use of this drug will be safe. When data or other adequate information establish that it is not reasonable to expect the new animal drug to become a component of food at concentrations considered unsafe, a regulatory method is not required.

FDA has also provided the following examples of the types of studies that may be used to evaluate the existence and safety profile of drug residues potentially found in food-producing animals (15):

1. Complete experimental protocols should be employed for determining the drug residue levels in the edible products (including residues present in muscle, liver, kidney, fat, and possibly skin, milk, and eggs), and the length of time required for residues to be eliminated from such products following the drug's use (studies should be conducted under appropriate conditions of dosage, time, and route of administration to show the levels of the drug and any metabolites in test animals during and upon cessation of drug treatment).

2. If an animal drug is provided via animal feed or water, appropriate consumption records of the medicated feed or water and appropriate performance data for the treated animal should be evaluated.
3. If an animal drug is to be used in more than one species, drug residue studies or appropriate metabolic studies should be conducted for each species that is food-producing.
4. If residues of the animal drug are suspected or known to be present in litter from treated animals, it may be necessary to obtain data with respect to such residues becoming components of other agricultural commodities because of use of litter from treated animals.

For genetically modified animals, FDA will evaluate how different the transgenic animal is from the traditional animal. FDA will ordinarily conduct a case-by-case assessment based on molecular biology research, toxicological studies, and perceived stability of the gene pool. As noted, however, the statutory human food safety requirements for genetically modified animals are the same as those for other animal drugs. FDA is therefore expected to require the food products produced from genetically modified animals to be as safe as those from nontransgenic animals.

However, the standard battery of toxicology studies used to establish the safety of "traditional" animal drugs are not expected to be appropriate for assessing the safety of a transgene in a genetically modified animal. Unlike traditional drugs, the genetically modified genes will not be eliminated from the animal, and therefore the concept of a "withdrawal period" would not apply. Accordingly the safety profile of the genetically modified gene and any expression products must be evaluated — and if safety issues arise due to expression products, FDA may be required to establish an appropriate tolerance for a level of acceptable use.

In evaluating the safety profile of genetically modified animals, FDA may also take into consideration the fact that mammals are often important indicators of their own safety, since adverse consequences of introduced genetic material will generally be reflected in the growth, development, and reproductive abilities of the mammals. Accordingly, if genetically modified mammals are healthy, FDA should be expected to take this fact into consideration when conducting its scientific analysis.

Fish, however, are known to produce toxins that are harmful to humans — but for which the fish have developed a natural resistance. Accordingly, unlike mammals, healthy fish may impose safety issues when intended for human food use — and the FDA is therefore unlikely to rely on the health of genetically modified fish as demonstrative of safety for human health.

REGULATION OF BIOTECHNOLOGY BY AGENCIES OTHER THAN THE FDA

USDA

Animal and Plant Health Inspection Service. APHIS regulates animal biologics, plants, plant pests, and nonhuman animal pests. The Animal Virus, Serum, and Toxin Act of 1913, which provides for the regulation of veterinary biological products (16), defines a veterinary biological product as including:

> all viruses, serums, toxins (excluding substances that are selectively toxic to microorganisms, *e.g.*, antibiotics), or analogous products at any stage of production, shipment, distribution, or sale, which are intended for use in the treatment of animals and which act primarily through the direct stimulation, supplementation, enhancement, or modulation of the immune system or immune response. The term "biological products" includes but is not limited to vaccines, bacterins, allergens, antibodies, antitoxins, toxoids, immunostimulants, certain cytokines, antigenic or immunizing components of live organisms, and diagnostic components, that are of natural or synthetic origin or that are derived from synthesizing or altering various substances or components of substances such as microorganisms, genes or genetic sequences, carbohydrates, proteins, antigens, allergens, or antibodies (17).

USDA issues licenses for biological products and establishments if the applicant meets standards designed to ensure the safety, purity, potency, and efficacy of the product (18). Animal biologics derived from biotechnology are expected to be regulated in the same manner as products that are prepared via conventional techniques (19). APHIS has also published a guidance document describing its method for conducting a risk analysis for veterinary biologics (20).

Food Safety Inspection Service. The Food Safety Inspection Service (FSIS) regulates products prepared from domestic livestock and poultry pursuant to the Federal Meat Inspection Act (FMIA) (21) and the Poultry Products Inspection Act (PPIA) (22). The FSIS is required to inspect meat and poultry products intended for human food to ensure they are wholesome, not adulterated, and properly marked, labeled, and packaged. Although both FDA and USDA share adulteration and misbranding jurisdiction over meat and poultry products, FSIS has primary jurisdiction over general compliance issues (22).

FSIS also regulates the slaughter for food use of livestock and poultry involved in biotechnology experiments under its regulations for livestock and poultry involved in research (23). FSIS regulates the slaughter of genetically modified animals (including experimental animals) somewhat differently than conventional animals. Specifically, FSIS has noted: "For nontransgenic livestock or poultry derived from transgenic experiments, the data should be submitted to FSIS, and would have to show that the animals to be slaughtered for food use do not have the experimental transgene, and consequently are equivalent to the parental line and, thus, are not adulterated as a result of the experiment" (24). If an animal with a transgenic modification is to be slaughtered, review and approval in accordance with the regulations would be required (25).

In general, for genetically modified meat and poultry, FSIS is expected to consult with FDA regarding safety and labeling issues. The FDA is expected to have primary responsibility for evaluating the safety and efficacy of the genetically modified animals intended for human food use.

FSIS, however, would still be responsible for conducting the inspections of genetically modified meat and poultry.

Environmental Issues and EPA Involvement

EPA regulates all pesticides under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and controls the use of genetically engineered microorganisms under the Toxic Substances Control Act (TSCA). EPA is not principally involved in regulating genetically modified animals intended for human food use, but rather it is expected to consult with FDA and USDA in evaluating environmental issues.

Significant environmental concerns may arise when genetically modified species are released into the environment. A genetically modified gene may spread more widely than anticipated, and environmental and ecological changes may occur as a result of competition between the transgenic variety of the species and the natural variety. Accordingly environmental issues and biocontainment strategies are expected to be evaluated in addition to the more traditional safety and efficacy review.

CONCLUSION

The majority of genetically modified animals intended for human food use are expected to be regulated by CVM as animal drugs. CVM does not currently intend to issue a standard set of guidelines on how the food safety determination for transgenic animals should be conducted. Accordingly CVM advises companies seeking approval for genetically modified animals to consult with CVM in order to develop appropriate protocols for evaluating human food safety issues.

The legal and regulatory climate appears hospitable toward the development of genetically modified animals for human food use. The federal food and drug laws provide the FDA with a degree of regulatory flexibility, and within the next decade, as FDA conducts its safety evaluations for these products, modifications to the existing regulatory regime may be implemented without substantial difficulties. The genetic modification of animals for human food use is still in its infancy, however, and the public and political reaction to such products is largely unknown at the present time.

BIBLIOGRAPHY

1. Executive Branch Study, *Use of Bovine Somatotropin in the United States: Its Potential Effects*, U.S. Government Printing Office, Washington, DC, January, 1994.
2. Food and Drug Administration, *1999 Report on the Food and Drug Administration's Review of the Safety of Recombinant Bovine Somatotropin*, FDA, Washington, DC, 1999.
3. *Ben and Jerry's Homemade, Inc. v. Lumpkin*, 1996 U.S. Dist. LEXIS 12469 (N.D. Ill. Aug. 27, 1996).
4. *International Dairy Foods Ass'n v. Amestoy*, 92 F.3d 67 (2nd Cir. 1996).
5. *Stauber v. Shalala*, 895 F. Supp. 1178 (W.D. Wis. 1995).
6. *Barnes v. Shalala*, 865 F. Supp. 550 (W.D. Wis. 1994).
7. G.A. Mitchell et al., Regulatory Issues in Agricultural Biotechnology, Conference on Urban/Rural Environmental, Food and Agricultural Issues, November 14, 1997, Citing L.M. Houdebine, in *Transgenic Animals Generation and Use*.
8. J.C. Matheson, *Possible Approaches and Issues for FDA Regulation of Animal Biotechnology: A Status Report*, Presentation at 1994 Workshop at Davis, CA, April 6, 1994.
9. J.C. Matheson, *FDA Veterinarian* (May–June) (1999).
10. 51 Fed. Reg. 23309 (1986).
11. 51 Fed. Reg. 23313 (1986).
12. 57 Fed. Reg. 6753 (1992).
13. FFDCA, §201(g)(1)(B).
14. M.A. Miller and J.C. Matheson, III, *FDA Veterinarian* (March–April) (1998).
15. 21 CFR 514.1(b)(7) (1999).
16. Pub. L. No. 62-430, 31 Stat. 828 (1913) (codified at 21 USC §151).
17. 62 Fed. Reg. 31326, 31328 (1997) (9 CFR §101.2).
18. 9 CFR Part. 102 (1999).
19. 51 Fed. Reg. 23336, 23340 (1986).
20. Animal and Plant Health Inspection Service, Guidance Document: *Risk Analysis for Veterinary Biologics*, APHIS, Washington, DC, 1994.
21. Federal Meat Inspection Act, 21 U.S.C. §§601–695.
22. Poultry Products Inspection Act, 21 U.S.C. §§451–470.
23. 9 CFR §309.17 (1999); 9 CFR §381.75 (1999).
24. 56 Fed. Reg. at 67054, 67055 (1991).
25. 9 CFR §§309.17 and 381.75 (1999).

See other entries AGRICULTURAL BIOTECHNOLOGY; FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE.

ANIMAL MEDICAL BIOTECHNOLOGY, LEGAL, LAWS AND REGULATIONS GOVERNING ANIMALS AS SOURCES OF HUMAN ORGANS

ANDREW TREW
John Carroll University
Cleveland, Ohio

OUTLINE

INTRODUCTION

How far do we need to regulate xenotransplantation, the use of animal organs, tissue, and cells for transplantation into human subjects? Xenotransplantation offers life-saving potential and brings hope of an end to waiting lists for transplants, and current ethical dilemmas surrounding who should receive transplants, at a time when demand for organs outstrips supply. However, xenotransplantation also involves fears about crossing the barriers between the species which are the result of gradual evolutionary changes over millions of years. This creates a real risk that human patients may be harmed by infections, transferring with the animal organs, cells, or tissue. Retroviral infection of transplant patients might take years before emerging as new diseases and could meanwhile spread to wider populations, creating public health implications. Regulatory provisions, nationally and internationally, address many of these fears.

In the United States, regulatory policy is based mainly on 1996 Federal Drug Administration (FDA) Guidelines. These provide for patient protection, through informed consent and patient autonomy provisions. There are stringent biocontainment and long-term post-operative monitoring provisions that involve major restrictions on civil liberties. The Guidelines also supplement existing Animal Welfare Laws, and provide detailed risk management strategies regarding the use of donor animals.

Patent laws are also central to regulatory frameworks for xenotransplantation. Genetically engineered (transgenic) animals are likely to be used as donors. For example, transgenic pigs have been designed to include human genetic material in order to reduce rejection problems associated with the use of animal organs (e.g., transgenic pig hearts). Transgenic techniques (in effect, transgenic animals) have been patented, by corporations who own patents to the associated immunosuppressive drugs. The advantages of using transgenics include a possible reduction in acute rejection. The risk of viral infection crossing from animal to human remains. Moreover, the use of drugs to suppress the rejection of organs, also suppresses the immune system, and may open the way for viruses, contained in the donor animal's DNA, to transfer more easily into human patients. There are also questions about the long-term viability of xenografts within human beings, including the ability of animal organs to perform effectively within another species.

There are also concerns about rearranging the DNA and genome of animals, and using patent laws to claim ownership over animals and thus control their availability or use in xenotransplantation. There are deep-seated legal concerns in Europe that conceptually the patent regime does not, or ought not to, extend to life forms. A movement toward the treatment of life as a commodity is central to this concern. Corporate "creators" of transgenic animals are likely to impose controls on the centers chosen for xenotransplantation clinical trials, which will promote corporate and not public interests. Transplant centers wishing to advance research in xenotransplantation, and provide greater equality of access to the new techniques, may be tempted to proceed using "natural" animals and develop new immunosuppressive drugs. This is permitted within the terms of the FDA Draft Guidelines. However, the risk of zoonosis (infectious agents crossing from animal to human) might be equally high whether transgenics or natural animals were used as donors in xenotransplantation procedures.

The lack of federal legislation, or an advisory agency, may possibly lead to a two-tier development of clinical trials, one using transgenics and the other using natural animals. Given the unquantifiable nature of any viral infection of the patient, and/or the public at large, there is growing national concern about the appropriate level of regulatory intervention. Xenotransplantation may continue to exist in a largely unregulated setting similar to existing provisions concerning assisted reproduction and human-to-human transplants. Policing of the FDA Guidelines, through the Institutional Review Board (IRB), at a local level, is clearly inadequate. Long-term monitoring of patients to detect viral infections may raise insurmountable restrictions on civil liberties. On the other hand, if there were to be viral infection of the public by xenotransplant recipients, it may be too late to regulate.

In the international context, the World Health Organization (WHO), the Council of Europe, the European Union, and the Organization for Economic Cooperation and Development (OECD), have all addressed the risks and benefits of xenotransplantation, emphasizing the need for international cooperation on appropriate public health policies and regulation. International concerns center on issues of justice, equality of access, and the availability of xenotransplantation in Third World countries. International conventions may be needed to avoid the possibility of forum shopping, as has occurred in the context of human organs for transplants. Forum shopping involves patients searching out a market where organs from wild, poorly screened animals are procured for use by transplant centers, in lucrative safe havens provided by countries where no xenotransplant regulations exist. These centers would offer tempting solutions to desperate patients. Any resulting viral infections would not respect political boundaries.

Internationally there may also be an inadequate scientific base on which to decide whether xenotransplantation ought to go ahead and what level of regulatory supervision would be appropriate to minimize risks.

Recent reports that infections may be able to cross over from animals into human cells add fuel to the view of some experts that a moratorium is appropriate. A moratorium on human cloning was proposed by the National Bioethics Advisory Commission, and accepted by President Clinton. However, it would appear xenotransplantation is not currently on the Commission's agenda. There is therefore increasing need for public debate and informed dialogue between scientists and lawyers. Discussion might center on regulatory polices that balance the benefits of research into new medical treatments, against the risks of emergent retroviral infection in patients, that could develop into public health concerns. Ought we to wait, or to go ahead now to regulate or permit a largely free market in xenotransplantation? Do the perceived risks outweigh the benefits of this frontier surgery? Or should we be looking to alternative sources for transplant material, such as embryonic stem cells or cloned material? Perhaps we should change the laws on organ and tissue donation, or develop more artificial spare parts for human patients? There is clearly a need for further public debate.

REGULATORY CONTEXT

Encouraging Increased Organ Donation

There is a severe shortage of organs from human donors for transplantation. The 1998 United Network for Organ Sharing (UNOS) statistics indicate a huge waiting list of patients, many of whom will die before a suitable organ is available. For UNOS statistics see: www.unos.org. As of December 1999 there were 69,550 patients on the national waiting lists. During 1999 only 21,941 transplants were carried out, spread over 272 medical institutions. The United States government has proposed measures to allocate organs to those most in need, which UNOS has opposed. Other countries experience similar shortages (e.g., statistics from the OECD in Ref. 1). Current legislation, the National Organ Transplant Act 1984 (2), relies on donations that are voluntary. Donor status depends on statements in driver's licences, in living wills, or donor cards, and in practice, the consent of relatives or proxies of the deceased. In the case of renewable organs or tissue, donations depend on individuals' altruistic conduct. Shortages may partly be the result of cultural or religious prohibitions. There are also deep-seated taboos. These are associated with respect for the dead or revulsion at removing organs and interfering with the integrity of a recently loved human being. Attempts to reduce organ shortages center on (*1*) developing artificial organs such as the electric heart, which may form a "bridge" until a human organ is available, (*2*) creating new criteria for legal brain death and/or creating a legal presumption of donation (as in the laws of Belgium and France), and (*3*) creation of a market for the sale of organs and tissue (3). There are two more recent proposals. First, organs might be developed from cloned human stem cells. This might enable genetically compatible transplantation, subject to removing any inherent genetic defects present in the donor's body. Second, organs may be taken from nonhuman animals to transplant into human subjects. This is known as xenotransplantation, as opposed to allotransplantation where the donor is human.

Regulatory Policy

What role might the law play in the development of xenotransplantation? Transplantation from animals to humans represents a turning point in medicine. However, it has not received the same publicity as animal and human cloning. The prospect of replicating humans has raised ethical and legal debate worldwide, leading to a ban in many countries (4). Replicating parts of humans, by creating transgenic animals with human genes, for example, donor pigs, has caused less outcry. Public awareness of the issues is lower, and experts have called for more open public debate (5) and education. Discussion focuses on appropriate ethical and regulatory policies to deal with public health risks potentially associated with xenotransplantation. Ought xenotransplants to be banned in the same way as many jurisdictions have outlawed human cloning research? The response in the United States has been a "no" in the form of FDA Draft Guidelines (6). These envisage controlled use of the donor animal, which is either "natural" or transgenic. However, internationally, WHO (7) and other international bodies are concerned about emerging infectious diseases that "cross over" from animal to human species. The central regulatory issue for public health is how far to control the possible spread of such retroviral infections, both to patients, and by patients, who have received transplanted organs or tissue from animal sources. The question is similar to the HIV/AIDS experience. How far can the potential spreading of a retroviral infection justify restrictions on patients' civil liberties?

There are existing warning signs that infections do indeed transfer from animals to humans. Recent cases off CJD (Creutzfeldt Jakob Disease) have occurred in Britain (8), which appear to originate in a form of crossover infection from cattle with BSE (Bovine Spongeoform Encephalopathy). Recent evidence suggests a strong connection between HIV and a form of simian immune deficiency (SIV) (9). Outbreaks of Hanta virus and Dengue fever are also warning precedents about the reality of such risks. The problem is that if xenotransplants are banned by law, research that may be vital will not proceed. Regulatory concerns therefore involve effective ways to strike a balance between the risks and benefits of xenotransplantation and conflicting interests.

In general, the United States is reluctant to legislate in the area of biotechnology and genetics, preferring to allow the marketplace, self-regulation, and existing patent regulation to govern. Extensive use of patents has led to the primacy of the interests of biotechnology corporations. For example, the U.S. government prefers a free market approach; it has refused to adopt the 1999 Protocol on Biosafety, which regulates trade in genetically engineered plants and animals, despite its adoption by over 120 countries (10).

Xenotransplantation is heavily dependant on the manufacture of genetically engineered animals (the animal of choice being the pig) and on associated immunosuppressive drugs, which control rejection of animal organs by humans. Both the animals and drugs are subject to patent protection. A central issue of justice is how far ought corporations to own the newly engineered species of transgenic animals, which in turn provides ownership rights to DNA and the building blocks of life (see the discussion below). A small number of patent owners in effect control access to xenotransplantation under patent licences granted to medical centers who wish to pioneer these new treatments.

Regulatory policy also needs to address how far existing concepts of informed consent and patient autonomy are adequate to protect the patients who consent to xenotransplant protocols. Given current organ shortages, patients may be tempted to opt for xenotransplants as a life-saving possibility, without full appreciation of the risks and outcomes. It might be suggested that a national body be established to guide Institutional Review Boards (IRBs) in design of protocols.

DEFINING XENOTRANSPLANTATION

Xenotransplantation is an emerging field. The scope of definition may change. Currently xenotransplantation covers a range of animal types as sources in the "xeno" part of the definition, and a range of procedures within the "transplantation" part of the definition. (See FDA Guidelines 1996, Para. 1.2.) The key issue is the use of 'live' tissue organs or cells. Other animal material used in treatments is regulated separately (see the discussion below).

"Xeno"

The "xeno," or nonhuman source of transplant material, covers three animal types. First, there are animals that arise in nature. Pigs, baboons, and even sheep have been used in pioneering transplants. The animals may be taken from the wild, bred in captivity, or reared under laboratory conditions. Second, there are transgenic animals. Such animals are genetically engineered to include genes from another species. These include transgenic pigs that have been genetically modified to include a small number of human genes to overcome specific aspects of the rejection problem when animal organs are transplanted into human patients (11). Third, there are cloned animals (12). Animals may be cloned from laboratory raised natural animals, for example, baboons or pigs. Such clones might be created following the successful use of a natural animal in transplant procedures, in order to ensure consistency of donor quality. Clones may also be created from transgenic animal donors identified as optimum specimens. There are other possibilities in the future. Animal stem cells may be used to "grow" an organ rather than the whole animal. Such techniques (13), if applied to human genetic material, may largely dispense with the need for xenotransplants in the future. Organs may also be cloned from human stem cells in the future. The organs could provide a compatible match for human patients, if genetically engineered to remove inherent defects that initially caused the patient's need for transplantation.

"Transplantation"

Transplantation refers to any procedure that involves the use of live cells, tissue, and organs, from a nonhuman animal source ("xeno" as above), whether transgenic or nontransgenic, transplanted or implanted into a human, or used for ex vivo perfusion. Use of a bioartificial liver support system might be included, where live sterile pig liver cells (hepatocytes) form part of liver dialysis treatment. Equally, animal neural cells are used in treatments for Parkinson's and Huntington's diseases, and baboon bone marrow is used as a treatment for AIDS.

Associated Definitions

The use of natural and transgenic animals as sources for treatments may include nonliving animal products. Such products have been used for some time without unforseen risks arising through transfer from animal to human. These nonliving materials are regulated separately. Porcine heart valves are classified as medical devices. Porcine insulin, on the other hand, is classified as a drug. Bovine serum albumin is classified as a biologic.

FEDERAL REGULATION IN THE UNITED STATES

Animal Regulatory Framework

There is no federal statue specifically regulating the use of animals as sources of organ donation. The National Organ Transplant Act (1984) only applies to human donors. However, animals used as donors will be subject to the Animal Welfare Act 1966, as amended in 1985 (14). Regulatory authority under the Act vests in the Secretary of the U.S. Department of Agriculture (USDA). The law regulates handling, sale, transportation, and humane treatment of a wide range of animals intended for research or experimentation, including a list of animals determined by the Secretary of Agriculture. The list covers warm-blooded animals, such as nonhuman primates, dogs, and rabbits. Legal provisions extend to health certification of animals prior to sale or transport, treatment of animals, humane care, and documentation requirements, together with appropriate use of anesthetics and other drugs.

The Food Security Act 1985, Subtitle F-Animal Welfare (15), provides for institutional supervision of protocols, standards of housing, and humane care of animals to be undertaken by the IACUC (Institutional Animal Care and Use Committee). The Act also expands the provision of humane and ethical care for, and use of, laboratory animals. The National Institutes of Health (NIH) Office for Protection from Research Risks (OPRR) provides an IACUC Guidebook, detailing procedures for the review of protocols which involve animal use (16). The Guidebook covers issues such as minimization of pain, euthanasia and research methodology, including the numbers of animals required, animal health and husbandry, facilities record keeping, and the health and safety of workers.

The IACUC, in the context of xenotransplantation, will need to address the difference, from a legal perspective, between the use of transgenic animals and natural animals. Transgenics, being regulated by patent law (17),

are subject to contractual and licensing restrictions as to their use and disposal, especially restrictions on the use of gametes or on breeding or cloning. Biosafety provisions in laboratories are crucial to avoid legal actions based on licencing infringements or negligence. Management should design procedures governing the secure physical and biological containment of transgenic animals and material, especially in relation to escape of animals or pathogens into the environment. Welfare issues specific to transgenic animals include possible suffering caused by exposure to infectious animals and physical abnormalities caused by mistakes in the genetic manipulation that leads to the creation of the transgenics. Similar, more familiar restrictions are likely to apply to naturally arising animals from controlled herds. Physical conditions of animals generally are regulated, including standards governing heating, ventilation, space, and sanitation. The need for physical exercise and for psychological well being are recognized. Procedures involving pain and distress to animals are defined and are to be minimized. Detailed regulations on animal welfare (18) ensure compliance with the Animal Welfare Act and its various amendments. The regulations incorporate provisions on licencing and registration, research facilities, record keeping, containment policies, holding periods, minimizing the numbers of animals used in research protocols, destruction of animals, and inspection of premises and records. The regulations set standards that cover the health of warm-blooded animals, breeding, and the operation of animal facilities. A valuable '*Guide to the Care and Use of Laboratory Animals*' is published by the National Research Council (19). This details the housing, management and medical care of animals, including euthanasia of animals, together with institutional policies and laboratory hazard containment, health and safety provisions. The Guide also provides details of national accreditation standards for animal facilities.

FDA Regulatory Supervision

In addition to the Draft Guidelines (discussed below), the FDA is also involved through IND (Investigational New Drug) approvals of clinical trials. These regulations require sponsors of transgenic experiments to obtain written FDA approval. One example of such approval is the use of neural cells from pigs for treatment of degenerative neurological diseases such as Parkinson's and Huntington's diseases. This is an alternative to the use of human fetal cells in the treatment of these conditions. Other trials have involved the use of livers from pigs as a bridge to human organ transplantation and the use of living sterile pig liver cells (hepatocytes) in a dialysis machine. These trials are all subject to FDA regulations in 21 CFR part 312 (20). IND applications are handled by the FDA Center for Biologics Evaluation and Research (CBER). Its first IND applications relating to xenotransplantation date from 1995. CBER had expressed concern about the possibility of infectious retroviruses from porcine blood derived cells infecting human patients. This was confirmed in the literature in 1998 (20). CBER publishes *Vision*, a valuable guide to its role and current activities, on the Internet (21).

FDA Draft Guidelines on Xenotransplantation

The 1996 Draft Public Health Service Guidelines as amended in May 2000, provide a regulatory framework for xenotransplantation (22), addressing the central risks to patients, medical teams, and the public at large. A final version of the Guidelines had not been produced by February 2000. However, detailed discussion with experts did take place in January 1998 (23) about the advances in research relating to viral risks as well as regulatory, ethical, and public health concerns. In general, the Guidelines attempt to contain such risks, including possible viral infection (zoonosis), and long-term retroviral risks to patients and their contacts.

Issues raised by the Guidelines can be conveniently divided into four main headings: (*1*) the medical team, (*2*) institutional requirements, (*3*) animal donors, and (*4*) patient concerns.

The Medical Team. Transplanting animal organs, tissues or cells into human patients requires interdisciplinary expertise. The Guidelines envisage a team comprising infectious disease and zoonosis specialists (both physicians and veterinarians), an immunologist, the director of a microbiology/virology laboratory, and an infection control specialist. Since there are currently some 270 transplant centers in the United States, it is clear that such teams may not be available at smaller sites. For all health care workers, including veterinarians and laboratory staff, there is a need for a continuing health education program to alert workers to the risks associated with both the handling of animal tissues and organs, and to provide support for xenotransplantation patients (para. 4.3.3). Educational materials should be developed to include instruction on safety procedures, appropriate protective clothing, and detailed risk management strategies. The seriousness of potential long-term risks of infection to workers is recognized in detailed provisions about worker surveillance (collection of baseline sera), postexposure protocols for monitoring or treatment of infected personnel, and maintenance of long term records regarding nosocomial (hospital/laboratory) based exposures to health risks.

Institutional Requirements. Centers engaging in solid organ transplants are required to be members of the OPTN (Organ Procurement and Transplantation Network) and to comply with legislation regarding the housing and treatment of laboratory animals, and also with the Public Health Service Act (24). Protocols will need to be reviewed by the center's IRB, IACUC, and Biosafety Committee. The Guidelines require members of the various boards reviewing protocols to have high levels of expertise in assessing and evaluating potential risks of infection. The risks could affect animals and laboratory workers; patients, their families, and contacts; and the health care team and the public at large. However, multidisciplinary expertise is unlikely to be available except in larger institutional settings. There is clearly a need for institutions to share expertise in the design and approval of protocols, and avoid duplication of research. Self-education is also vital to ensure awareness of latest research and developments in xenotransplantation trials.

In addition there is a case for establishing a national xenotransplantation advisory agency to integrate matters concerning the institutional needs of transplant centers, oversight of regulatory issues, and an ongoing review of risks to patients, health care workers, and the public. This would also be valuable in coordinating the collection and comparison of health care records.

The Guidelines provide that records of the progress and outcomes of xenotransplantation should be kept indefinitely. An Institutional Xenotransplantation Record would contain full details of the procedures, animal sources, and all those concerned with the patient's care. The risks of hospital acquired (nosocomial) infection are to be minimized by following infection control policies. Where appropriate, isolation precautions should be followed. Hence a permanent Nosocomial Health Exposure Log is to be kept. This should track risks to employees of potential transmission of xenogeneic infections. A third permanent document, covering individual health care records, should follow the patient throughout the clinical stage and record the results of postoperative surveillance.

Although a national registry is envisaged, as well as a serum and tissue archive, there is currently no provision for an integrated federal advisory resource. Those who advocate a federal agency might consider a recent precedent in Great Britain where a regulatory agency, the United Kingdom Xenotransplantation Interim Regulatory Authority (UKXIRA) was established in 1998. A regulatory or advisory agency could service institutional needs in this context and include a central registry for statistics, avoidance of duplication in early trials, and sharing of both research data and details of xenogeneic infections. Equally, there is a need for a national training resource to provide the IRB, IACUC, and indeed hospital legal departments and management with updated information about scientific, legal, and ethical concerns. Without some form of regulatory or advisory agency, equality of access, to both the xenografts and to research findings, may be restricted by the vested interests of the patent owners, who control access to transgenic animals and associated immunosuppressive drugs. This may lead institutions who are not chosen as initial centers by corporate patent owners, to pursue the use of "natural" animals, such as pigs or baboons as organ or tissue donors, and attempt to overcome the problems associated with early xenotransplant experiments.

Animal Donors. Central to the Guidelines on Xenotransplantation are detailed provisions regulating the use of animals as sources of organs, tissue or cells. The animals should not be procured from the wild (para. 3.1.4) from abattoirs (para. 3.1.7), or be imported from overseas or be immediate offspring of such imports (para. 3.1.5). However, imported animals of a species not available in the United States, may provide donor material if their history is properly documented. For example, provision is made to ensure imported animals are free from transmissible diseases (e.g., BSE-Mad Cow Disease). The animals bred and reared in captivity should have documented lineage to ensure disease-free animal donors. Failure to follow the FDA Guidelines as a standard of care could result in negligence liability. The use of transgenic animals (e.g., pigs) may reduce some risks of rejection but still carry the potential for retroviral infection for transplant recipients (para. 5.5.). However, their use is not specifically addressed in the text of the Guidelines. Whatever type of animal is used as donor, the Guidelines provide detailed safety provisions governing animal housing, feeding, screening, and surveillance (para. 3.2).

Generally, any facility that houses animals used in xenotransplantation procedures must comply with the National Accreditation Standards (detailed in Ref. 19). The animals and humans entering the facility would be monitored and screened for diseases and infections. Quarantine for animals (para. 3.5) is to be provided for at least three weeks before animals are used in xenotransplantation procedures. Animals should be screened for infectious agents and viral agents, which research shows can infect human cells in vivo or in vitro. Routine serum samples should be obtained from the animals, and animal deaths investigated thoroughly. As in patient records, documentation about the health of the animal donors and the herds (if any) from which they are drawn from, be maintained indefinitely. Equally important is control of animal feed (para. 3.2.1.3). Recycled and rendered animal materials may be a risk factor associated with prion associated diseases. These include BSE in cattle and CJD (Mad Cow Disease) in humans. Other methods of risk containment cover screening for infectious agents (para. 3.3) and include preclinical studies to identify the nature of species specific endogenous retroviruses and their potential to transfer to human patients as infections. Such studies might extend to an assessment of the potential of samples from xenografts to infect human cells.

What is missing from the FDA Draft Guidelines is any attempt to prohibit the use of certain animals as donors. There is clearly an animal welfare issue here. More especially there is a threat of species-crossing infection, posed by the use of nonhuman primates such as baboons. The Guidelines fail to provide a detailed assessment of the risks of using natural animals in general. Clearly, the extensive treatment of animal issues in the Guidelines illustrates the degree of concern about the possible spread to patients of animal viral infections that might be contained in the xenograft. This raises a central issue of regulatory policy. If the FDA, having the benefit of national and international expertise in drawing up the Guidelines, provides such detailed risk management procedures, might it be that the risks are substantial enough for there to be a strong case for enforceable federal legislation, or even for a moratorium? There are clearly international regulatory problems, since patients may readily spread their infections travelling abroad. Concerns about animals ultimately reflect a fundamental concern for the impact of xenotransplantation on patients.

Patient Concerns. Xenotransplantation is an invasive and experimental procedure. The patient is entitled to standard legal protections governing informed consent to risks. These are ensured by patient autonomy under the Patient Self-determination Act 1990, IRB supervision of protocols, and through leading court decisions such as

Canterbury v. Spence (25). Informed consent also protects medical teams against actions for the tort of battery. Risks communicated to the patient include "inherent and potential hazards of the proposed treatment, the alternatives . . . and the results likely if the patient remains untreated" (25, p. 776). Both the risks and outcomes of the surgery should be explained to the patient in lay language and on a signed consent form (26). When reviewing protocols IRB members have a responsibility to include easily understood language to describe the surgery.

The Guidelines envisage that the issues specific to xenotransplantation would be fully explained to patients prior to their signing consent documents. Explanations would cover risks of both known and unknown infections being passed from the animal donor to the patient. Patients would be made aware that any infection may not occur until some time after the surgery and that it may be possible to infect both partners and third parties. Because of the possible outcomes, patients should be prepared to be monitored, possibly for life, to undertake regular tests, and to provide up to date information about changes of address. Further restrictions include not donating blood, gametes, tissue, or body parts for use by human beings. Moreover patients face the risks of passing infection to those involved in any future sexual relationships. The Guidelines advise patients to treat the risk of infecting third parties in the same way as transmission of HIV, suggesting that the use of barriers during intercourse may minimize risks to partners. All these provisions reflect attempts to monitor future unquantifiable risks of infection to patients, their families, or the general public. These restrictions clearly involve severe incursions into traditional civil liberties and raise questions of how far they could be enforced as a contractual obligation rather than a simple consent form. How could they actually be monitored? Although the patient will be reliant on drug or other therapies for a long period of time after the xenograft, compliance with "guidelines" of conduct will depend on individual responsibility. Otherwise, the early patients could well be condemned to a life of quarantined isolation.

How patients can give meaningful "informed" consent to such a range of risks and restrictions on personal freedom is uncertain. Indeed, absent federal legislation, how could a patient be restrained by an institution and his or her liberties be curtailed? Suppose that the patient is making a good recovery but must submit to the long-term surveillance required by the Guidelines. Over what time period should this extend? What if any retroviral infection takes years to emerge (as in the case of HIV)? The Guidelines suggest warning the patient that xenogeneic infectious agents may be transmitted by unprotected sexual activity. How could that be monitored? If we treat this as a case where risks and benefits need to be balanced, how could patients realistically be given advice about outcomes or future potential for viral infections when so little is known? In reality, "consent" is meaningless here. Moreover the institution that foresees possible public health risks might well be held accountable for initiating a chain of events involving risks to third parties. A xenotransplant patient post operatively is in effect a walking source of novel and unknown infection. Knowing the patient is the source of foreseeable risk of harm, how might this impact on negligence liability? On a product liability basis, we might argue that the "manufacture" and sale of a transgenic animal, with foreseeable potential to harbor retroviruses, could lead to third party liability. This may indeed justify federal intervention. At the minimum greater public discussion is needed (27).

PUBLIC HEALTH RISKS

Identifying public health risks has been central to the Guidelines from the outset. In the absence of a national regulatory agency for xenotransplantation, the heavy burden to monitor patients and health care workers indefinitely, and to maintain tissue and data archives, falls on individual transplant centers at the point where an infection threatens public health. Public health risks could derive from (*1*) patients, (*2*) animal donors, or (*3*) health care workers.

Risks from Patients

The consent section of xenotransplantation protocols require patients to be apprised of the risks of transmitting infections to family or close contacts "with whom the recipient participates in activities that could result in exchange of bodily fluids" (para. 2.5.3). Educating the patient about risks of infection is left to the transplant center. It is unclear how this might be done outside the framework of xenotransplant protocols and consent forms. Counseling may be appropriate, and clearly centers will need to consider their obligation to ensure informed consent is not subject to any pressure. Education of a patient's close contacts about the uncertainty regarding risks, and of the need to tell their doctor about any unexplained illness, is seen as the individual responsibility of the patient (para. 4.2). Suitable indemnity clauses may need to be inserted into the consent section of protocols. Conduct that might spread infections, such as needle sharing, unprotected intercourse, donation of blood should be prohibited. In the event of a clinical episode, which might lead to infection, it is envisaged that the state health department, the Center for Disease Control) (CDC) in Atlanta and the FDA will be notified (para. 4.1.1.6).

Every transplant center is obliged to maintain permanent archives of biological specimens relating to the patient, taken before the xenotransplant, immediately after the operation, and then at regular intervals for the rest of the patient's life. The intervals could be regulated by individual protocols (para. 4.1.1.2). Similar data of potential public health value would include the provision of permanent archiving of major organs and tissue samples from autopsies of patients who have died after xenotransplantation. Secure storage facilities would have to be provided within the transplant centers. All of these obligations raise the need for detailed in-house policies in collecting and maintaining long-term archival data on patients. So for there is no provision for early warning, or sharing of information, that might alert to potential public health concerns. The Guidelines do recognize the need for a national registry of data to track common features among xenograft recipients and to assess long-term

safety (para. 5.1). However, it is not clear how or when such an agency could be created.

Risks from Animals

When it comes to animals as a public health concern, the Guidelines provide for detailed systematic record keeping to link biological samples from donor animals to transplant recipients. This is "essential for public health investigation and containment of emergent xenogeneic infections." (para. 3.7). Once again, maintaining the integrity and security of long term cryogenic storage of samples is the responsibility of individual transplant centers. Transplant centers are expected to provide education, and suitable educational materials, for health care workers who face risks involved with animal donors or human recipients of xenografts.

Risks from Health Care Workers

Occupational health care programs should provide details of risks, ensure that standard precautions are followed, including protective clothing and disposal of waste, and monitor for possible infection among workers. Workers forming part of a xenotransplant team that handles donor animal material or provides care for patients would be subject to health checks including the archiving of baseline sera, obtained from workers, before exposure to xenografts or recipients. In the event of exposure to infection, or unexplained illness after exposure, there is provision for reporting and recording incidents (para. 4.3.3.).

Risk Assessment and Wider Concerns

The provisions relating to public health thus center on specific risk management provisions. However, the Guidelines contain no provision to ensure maintenance of safety provisions, no means to provide effective exchange of know how, no ways to standardize data collection, analysis, and exchange among individual centers. No effective provision exists to ensure uniformity of standards regarding protocol approval and self-education of the members of IRB or IACUC. Without objective consideration of risks, there is a danger that institutional review of protocols will initially be driven by commercial interests. In those centers where non-patented animals are used, there is no accessible national resource to provide expert advice to IRBs and research groups. Appropriate federal laws or regulatory agency could provide this.

Risk Reduction, "Natural" and Transgenic Animals

Overall, assessment or reduction of risks is difficult absent clinical trials with real patients. Recent research has indicated that the potential for infection of human cells by animal donor material is considerable, and this sounds a warning note. The use of transgenic organs genetically modified to include some five to ten human genes would help to reduce problems of rejection by human patients. However, retroviral transfer would remain a serious risk to patients. Already the pig genome is known to contain retroviruses that are endogenous, essentially integrated into the genetic makeup of the pig. If retroviral transfer does take place, the threat to public health from infected patients could be considerable. Against any evidence of foreseeable harm to both the patient and public, to proceed with clinical trials would seem at best negligent. An alternative attempt may be to transplant porcine material into monkeys in order to obtain the clues as to the long-term viability of placing a pig organ in a human.

The Guidelines do not ban the use of nonhuman primates as donors, although there are restrictions as to their procurement. For example, wild caught animals ought not to be used, nor captive free-range animals, in order to minimize some infectious risks. In the past a variety of animals had been used in the early xenotransplantation experiments. Since 1963 chimpanzees, baboons, sheep, and pigs have been used. In the so-called Baby Fae case, in 1984, surgeons at Loma Linda Hospital in California transplanted the liver of a seven-month-old baboon into a newborn human infant. The baby survived for 20 days and then died of kidney failure. More recently the University of Pittsburgh Medical Center transplanted baboon livers into two patients who were dying from Hepatitis B. One, aged 35, lived for 71 days in 1992. The other patient aged 62 died after 26 days. Baboon bone marrow was transplanted into an AIDS patient in 1995. These early cases do not indicate how far the risk of acquiring infections might be significantly reduced by breeding and rearing nonhuman primates in captivity and isolation. Moreover there are widespread ethical concerns about breeding baboons, monkeys, or other animals to farm their organs and tissue for human use.

Experts seem to agree that the pig offers a better choice of donor, whether transgenically engineered or not. However, although pig heart valves and insulin have been used in human patients, inserting an animal organ into a human patient raises novel issues. Although breeding pigs in isolation may reduce some infectious risks, the fear is that genetically inherited retroviruses in the makeup of the pig genome may transfer to the human patient. Even if transgenic pigs were used, it seems that "human proteins expressed on the surface of transgenic pig cells can act as receptors for viruses." Opinion is divided; see Refs. 5 and 21 (28). Producing a virus-free transgenic would require identifying viruses as yet unknown. Indeed, the impact on humans of viruses that are of no danger to pigs raises the familiar argument that a virus like that in Michael Chrichton's *Andromeda Strain* may mutate into a wider population. It is also conceivable that transplant recipients may infect other animals. Clearly, concerns about public health are the basis for extreme caution.

Individual patients may in desperation agree to become guinea pigs in early surgery in having accepted the likelihood of an uncertain future or death. But suppose that the surgery were reasonably successful using transgenic pigs and new immunosuppressive treatments. How would it be possible in a democracy to control the sexual behavior and personal freedom of xenotransplant recipients, as envisaged in the Guidelines, so that there were minimum risk to public health? It is surely unrealistic not to be deeply concerned about a scenario where the public became infected with new animal-to-human viral infections. There are some precedents

already. In Britain and other parts of Europe, outbreaks of CJD (Mad Cow Disease) in humans was transmitted from cattle infected with BSE. In Australia, the deaths of veterinarians examining an infected horse were attributed to cross-species infection. Worldwide, in the case of the HIV virus, the likelihood is that it transferred to humans from the sooty mangabey and from chimpanzees (9), originating in strains of SIV (Simian Immunodeficiency Virus). These examples offer cautionary tales for those who advocate going ahead with xenotransplantation. For some it may be a question of the "unnatural" way in which species barriers are crossed. Is nature showing us a red light here? Jean Rostand, the French biologist, wrote soon after the discovery of the DNA sequence in the 1950s that in Nature there may be some defined "no trespassing" signs (29), which we cross at unknown peril.

A XENOTRANSPLANTATION MORATORIUM?

Could xenotransplantation take us to the point where we should pause before crossing into new territory in biotechnology? The evidence of risks is strong. If there is a public health implication, the public has an interest in being informed about the risks and benefits of xenotransplantation. The transfer of organs from animals raises a possible but unquantifiable risk of xenotropic organisms. These may cause infections of an unknown kind in humans that were not harmful to donor animals. Endogenous retroviruses may transfer with the animal organ, with risks of new types of infection. Were the viruses to mutate in humans, it is conceivable that animal donor species could be infected through contact with humans. Research into the interplay of human and porcine viruses in the influenza epidemic of 1918 lends support to this idea. Moreover the patient will be more open to infection generally after immunosuppressive drug treatments, and consequently react in a different way to infections derived from animal donors than would a healthy person. A moratorium on xenotransplantation might be based, either on a version of the so-called Asilomar agreement, which involved scientists curtailing their own use of science through self-regulation, or on a federal ban on funding, as is proposed in the case of human cloning research. The issue whether a moratorium is required has so far not been addressed by the President's National Bioethics Advisory Commission and is unlikely to be the subject of federal legislation in the near future. Unless the issues of public and patient risks are raised to a higher profile within the media, public debate on the future of xenotransplantation regulatory policy is likely to remain largely uninformed.

DRAFTING CONSENT FORMS

It may be too late for a moratorium. There are now proposals being presented to IRBs in medical centers for clinical trials with backing from major pharmaceutical companies. If transplants are to proceed, it is vital that attorneys representing medical centers be aware of the public health risks and monitoring issues in the Guidelines and in the literature. Drafting consent forms within protocols ought not to be left to the medical/research team. Patients are "consenting" both to life-saving surgery and to becoming an additional risk of infection for others after surgery, which is unusual. Thus attorneys need to consider novel contractual and negligence liability clauses. They should be aware of the competing interests represented by the commercial sponsors anxious to promote patented commercial advantage, the patients who desperately seek potentially life-saving surgery but lack expert independent advice, and the ultimate "consumers" of infections, the general public. There may be a need to create a forum of legal and ethical experts, for group discussion and education of patients' families or partners about potential risks, and to explore issues of liability. Public indemnity insurance may also require revision. Positive undertakings would also be required. Perhaps patients would need to submit to a form of quarantine, the length of which seems difficult to stipulate. Expert scientific opinion ought to be invoked. However, the basis of medical care is consensual. Unless the risk to public health is certified by some state or federal agency, how could patients legally be detained against their will without the hospital facing false imprisonment litigation? Without a clear federal enforcement mechanism, the FDA Guidelines are really a paper tiger.

FDA REVISED GUIDELINES

After considering concerns from scientists and ethicists, the FDA issued Revised Draft Guidelines at the end of May 2000 (29a). These provide the FDA with closer regulatory oversight of clinical trials involving xenotransplantation. No clinical trials are allowed unless investigators have first submitted applications for prior FDA authorization. The FDA will propose further regulations and guidance on protocols for industry. There are further limitations on the use of nonhuman primates, coupled with FDA acceptance that there is insufficient information about risks from the use of animals as donors in xenotransplantation. Risk minimization procedures are strengthened regarding the screening and surveillance of animals, and those coming into contact with them, including patients.

Taking into account the needs to counsel patients, their families and close contacts about risks, the Guidelines place new obligations on sponsors for the design and monitoring of clinical trials and for counseling. This might seem inadequate, given the vested interests of sponsors. However, the FDA expects a coordinating national role will be played by the Secretary's Advisory Committee on Xenotransplantation, soon to be established. In part, the Committee will oversee protocol designs, and evaluate wider scientific, medical, and public health issues related to xenotransplantation. To meet wider public health monitoring concerns, the FDA will now require maintaining of both records and specimens from animal donors and patients for a period of fifty years.

The revised Guidelines go some way toward meeting the concerns raised since the 1996 Guidelines were published. However, there remain real concerns about the ways in which long term monitoring of patients might be achieved, and the preventative measures necessary to address any widespread infection of the public at large. Moreover, there

is a growing need to view the problems of risk management and surveillance as an international and not merely an U.S. policy concern.

INTERNATIONAL REGULATORY ACTION

The regulation of xenotransplantation is of growing international concern. Equality of access to pioneering surgery tends to grow slowly over a period of time. The pace of medical advances may be deterred by overregulation in a given country. There are cultural differences with regard to health care coverage and approaches to the use of animals in surgery. Further, there is a general concern that if patent owners control the medical market for the use of transgenic animals, some countries may be kept out of the early exchange of research data and clinical expertise. Health, unlike the environment, has not been the subject of major international regulation. Nonetheless, both international organizations and individual countries are engaged in promoting regulatory initiatives, in some cases involving a moratorium on clinical trials pending further expert and public debate. These regulatory initiatives indicate a shared international perception about the risks from animal to humans of viral infections that could threaten public health as transplant recipients travel from country to country.

World Health Organization

WHO is concerned about the emergence of new communicable diseases. Although lacking regulatory authority, WHO reports provide a catalyst for international research and cooperation, leading in turn to international treaties and conventions. In 1998 WHO published a Consultation on Xenotransplantation (7,30), providing a framework for international policy. International agreement is urged "to ensure that xenotransplantation is developed in conformity with accepted ethical and legal standards, based on the need to respect human dignity and individual rights together with community interests" (para. 8.1.14.). The WHO report is neutral on whether to advocate xenotransplantation. Recommendations (para. 8) provide a check list of issues which countries thinking of undertaking xenotransplantation need to address. These cover maximizing individual and public health by managing risks of zoonoses; procurement of healthy animals; risk assessment, counseling, and monitoring of transplant recipients and their contacts; developing archives of human recipient and animal donor tissue; and ensuring animal welfare. Informed public debate should include sensitivity to national, cultural, and religious norms. Equality of access should be encouraged through information exchange, nationally and internationally. A multidisciplinary xenotransplantation review body should be established on regulatory policy, patient welfare, archives, and global exchange of information.

Council of Europe

The Council of Europe, which represents some 40 states (31), has adopted a Convention on Human Rights and Biomedicine (32). This provides the basis (under Articles 19 and 20) for the current development of a common regulatory policy on organ transplants (32). A Second Protocol on Xenotransplantation may be added to the Convention following the proposal in 1999 to establish a moratorium on xenotransplantation. In 1997 the Committee of Ministers passed a Recommendation on Xenotransplantation, aimed at regulations to minimize "the risk of transmission of known or unknown diseases and infections to either the human or animal population" (33). Regulations would cover issues such as research and clinical trials, sources and welfare of donor animals, and long-term review of animal donors and transplant recipients. Subsequently, in 1998, the Parliamentary Assembly of the Council of Europe received two Reports on xenotransplantation prepared by member representatives (34). Both endorsed a "legally binding moratorium on clinical xenotransplantation," based on risks of a major pandemic affecting public health. The moratorium is linked to proposals to stimulate research into risks of infections to humans and animals, and full examination of appropriate ethical and legal policies.

Most recently, in 1999, the Parliamentary Assembly adopted a Recommendation on xenotransplantation (35), formally urging the Committee of Ministers to work for rapid introduction of a legally binding moratorium, to consider development of a Second Protocol to the Convention on Human Rights and Medicine, and to plan an international strategy with WHO to balance ethical, legal, medical, and public health aspects of xenotransplantation. In addition the ministers are urged to lobby for a worldwide legally binding agreement for a moratorium. Despite the optimistic tone of all theses documents, they are unlikely to be supported by the United States, given the reluctance to legislate on biotechnology issues. However, declarations by some 40 countries might add weight to the argument that it is time for the President's National Bioethics Advisory Commission to address xenotransplantation regulatory policy in detail.

The European Union

Within the 15-member states of the European Union (EU), there has been a long and heated discussion about appropriate legal policies to balance the availability of patent protection to biotechnology companies against the philosophy, prevalent in Europe, that patenting life forms raises serious ethical and moral objections. Safeguarding the dignity and integrity of the person is seen as a central legal issue. In 1998 the EU adopted a Legal Protection of Biotechnological Inventions directive for implementation by the member states no later than July 30, 2000 (36). In general, the new law relates to existing uncertainties about the scope of Article 53(b) of the European Patent Convention, as applied to genetically modified animals. Article 53 (b) excludes from patentability "animal varieties or essentially biological processes for the production of . . . animals." Specifically, this complex directive includes legislation of relevance to xenotransplantation. The law identifies the difference between inventions and discoveries. Attempts to patent a mere DNA sequence or partial sequence will not be permitted, as being discoveries about natural phenomena. On the other hand,

even inventions may be unpatentable if they fall within Article 6 "where their commercial exploitation would be contrary to public order or morality." The list of inventions includes "uses of embryos for industrial or commercial purposes," which might prevent patents on embryonic stem cell usage to create organs or tissue as an alternative to xenografts." Furthermore, "processes for modifying the genetic identity of animals" are unpatentable, if "likely to cause them suffering without any substantial medical benefit to man or animal," as are "animals resulting from such processes." Would the transgenic pigs, engineered to contain some human genes, fall within this Article? The "substantial medical benefit" of xenotransplantation, as we have seen, is subject to risks of substantial harm to patients from the use of genetically altered animal donors.

Methods of National Regulation

Some countries, such as India, are adopting a similar model of regulation to that proposed by the Council of Europe. India has declared a strict ban on organ transplantation including xenotransplantation clinical trials. A doctor who attempted to transplant a pig's liver into a patient with severe heart problems was prosecuted in 1997 (37). At the same time India's Council of Medical Research is drawing up guidelines. Another model for regulation is the establishment of a national regulatory authority. This has been adopted in countries such as Britain, where there is a xenotransplantation regulatory authority, UKXIRA. The Canadian government has initiated proposals for a National Advisory Board on Xenotransplantation based on recommendations of a National Forum on Xenotransplantation exploring clinical ethical and regulatory issues, which was convened by Health Canada in November 1997. Canada's proposals are likely to involve a standard-based regulatory system linked to the Food and Drugs regulations, and the National Animal Care Committee. In France, a similar expert advisory committee oversees safety issues, and legislation is under review. In Sweden, a Committee on Xenotransplantation was established during 1998; public opinion was tested through a questionnaire and public hearings held on xenotransplantation. In Australia, the government has taken initial steps to establish a regulatory body. The Australian Health Ethics Committee has been developing guidelines to assist local Institutional Ethics Committees on scientific and ethical issues. The context for the xenotransplant debates in Australia was a proposal to commence diabetes treatments using pancreatic islet cells from pigs. Another model is to include xenotransplantation in constitutional reforms covering organ transplants. Switzerland is considering a constitutional amendment to cover organ, tissue, and cell transplants from both human and animal donors. The Swiss Science Council is also researching the ethics and risks of xenotransplantation.

Establishing National Regulatory Bodies

The main regulatory shortcoming of the current FDA Guidelines in the United States remains the lack of enforceability. The role of the law might be seen as standard-setting here. Enforcement of the Occupational Health and Safety laws provides a model that might be valuable in xenotransplantation policy. The regulatory framework in Britain offers an alternative models to the current FDA Guidelines.

British research pioneered the techniques for creating transgenic pigs as an alternative to "natural" donor animals such as baboons. UKXIRA offers a national focus for monitoring research and proposed clinical trials, a framework for regulatory issues, and a means of assessing the scientific evidence about risks and current techniques (38). Established in 1998, UKXIRA promotes links with other scientific and regulatory bodies across the globe. Ideally such an agency would encourage exchange and minimize duplication of data between government departments, and provide a cross-departmental focus. Regulation is to be provided through legally binding Codes of Practice and Health Service Circulars. As for animal donors, existing legislation, the Animal (Scientific Procedures) Act 1986, provides for an Inspectorate to oversee licensed use of animals in scientific or medical procedures and experiments. A Code of Practice governing the welfare of donor animals used in xenotransplantation is being developed by the Inspectorate, and UKXIRA is initiating parallel provisions on biosecurity. Clinical trials are governed by a legally binding Health Service Circular (39) which contains detailed guidance provisions, in the form of Directions to the National Health Service. Health authorities cannot undertake xenotransplantation treatments without prior written approval by the Secretary of State for Health under Section 17, National Health Service Act 1977. The guidance sections address information about the medical team, sources of animals, their housing and welfare, infectious risk controls, patient consent and monitoring. The role of UKXIRA is to advise the Secretary of State for Health on "the acceptability of specific applications to proceed with xenotransplantation in humans." Thus the British model clearly ties together the general regulatory framework governing health provision and animal welfare, with binding specific regulations administered by a national expert regulatory agency. In this way coordination of information and coherent national standards for xenotransplantation clinical trials are ensured.

REGULATORY LIMITS TO THE USES OF ANIMALS AS DONORS

What is the future for animals as donors? The patenting of animals and other life forms continues to be an ongoing legal controversy central to xenotransplantation.

Comparative Legal Issues

Since 1980, when the Supreme Court decided the *Diamond v. Chakrabarty* case, the way has been open to engineer life forms and gain patent protection in the United States. The original patent law was amended in 1995 after the Trade in Related Aspects of Intellectual Property (TRIP) agreement at the Uraguay Round of GATT extended protection to 20 years from the date of application (40). The Chakrabarty case turned on the genetic manipulation of bacteria that might be used to break up oil slicks, which was thus a rearrangement of life forms. However, the Court viewed

this as a narrow case. It did not seem to foresee the huge impact that would result from extending patent protection, beyond the purposes envisaged by Thomas Jefferson's understanding of "inventions," and the concepts adopted since the original Act of 1793 which exclude from patentability "discoveries" about Nature. After all, the discovery of the structure of the DNA "double helix" was not patented by Crick and Watson in the 1950s. Extension of U.S. patent laws to protect the discovery, rather than invention, of what seem to be the rules of Nature governing the sequences, arrangement, and rearrangement of DNA has caused animated controversy in the European Patent Office. As a result there was reluctance to register the U.S. Oncomouse, as the first patent on a genetically engineered mouse used in human cancer research (41). Initially the argument prevailed, both in Europe and in the United States, that the techniques for sequencing and manipulating genetic material are inventions.

In 1998, however, legislation in the form of a directive (36), was passed by the European Union (EU), covering the legal protection of biotechnological inventions. This attempts to deliniate boundaries between discoveries and inventions. For example, a "mere DNA sequence without indication of a function" is not patentable because it is treated as a discovery. Although inventions would normally be patentable, as discussed earlier, the directive excludes certain classes of inventions if "their commercial exploitation offends against *ordre morale* and morality." The inventions are detailed in Article 6, and include "processes for modifying the genetic identity of animals." The United States continues to pursue policies permitting patents on a wider basis. The driving force may be considerations of trade. There are billions of dollars invested in biotechnology companies, based on potential revenues from genetic patents, including those on animal life and human stem cells.

Biodiversity and Animal Donors

Could the Chakrabarty case be challenged? The Supreme Court decision seems to contain a fundamental conceptual error. If genetically "created" bacteria can be patented, then any life forms, including human beings, might be altered and patented, producing possible adverse effects on biodiversity. Control over patenting of transgenic plants, such as wheat and tomatoes, or transgenic animals, such as the Oncomouse or the pig engineered to contain human genes, proved a logical extension of Chakrabarty. The ability to selectively breed and "own" hybrids of plants or animals is familiar to us. Explorers over centuries have rearranged the geographical location and enabled exposure to other species of countless varieties of flora and fauna. All this could be seen as interfering with "natural" biodiversity, and in 1992 the Biodiversity Treaty attempted both to encourage and regulate biodiversity. Nonetheless, the United States refused to sign the subsequent Biosafety Protocol on the trade in genetically altered plants and animals, endorsed by more than 120 countries, in February 1999 (10). The Protocol is graphic indication of the spread of genetically manufactured life forms. Biotechnology enables us to bypass evolution and genetically engineer plants and animals and potentially human beings. This in turn enables us to impose a new evolutionary shift on plants, animals, and eventually humans, and it sets in motion a chain of events over which we have little control. Controling the nature of Nature may thus lead to harmful ecological changes. Many would argue that genetically engineered animals, plants, or seeds can effectively improve the quality of agricultural output, promote resitance to disease, and reduce the risk of famine. Yet trade in and use of genetically modified organisms may result in major adaptation to the food chain, mingling wild and genetically engineered species. Within a period of years, biodiversity and evolutionary changes may produce instability as a result of our short-term perception of agricultural problems and their solutions.

The recent refusal of the United States to ratify the Biosafety Protocol shows the pressures that biotechnology industries can exert in this area of patents. Countries in which initial development of the new genetics occurred, notably Britain, have supported a more global perspective through a regulatory interventionist policy. By rejecting the Protocol, the United States will encounter barriers to trade and restricted access to markets overseas in which to license patented plants and animals.

Conflict of Laws Issues

Research and overseas investment by American biotechnology companies may be particularly affected by conflict between the U.S. Patent regime and the 1998 EU directive on biotechnology patents adopted by the 15 EU countries. On the other hand, xenotransplantation advances within the EU may be hindered if reciprocal recognition of U.S. patents on transgenics is denied within its countries. American companies would presumably be unable to protect overseas licencing restrictions within the EU jurisdiction. EU countries wishing to pursue xenotransplantation, using U.S. originated transgenic pigs, may not have access to these animals and may develop unpatented transgenics or possibly use "natural" animals for transplantation, coupled with developing new types of immunosuppressive drugs.

In the United States, patents continue to be registered that cover human genetic material, whether or not incorporated into animals. Ought we not to be concerned about human or animal life being owned and commodified through intellectual property rights? The ongoing Human Genome Project, which aims to map all human genes, was certainly not intended to lead to the patenting of gene sequences and human or animal genetic material. Rather, it was a public project, a multinational research exercise, to map the human genome, backed by many governments, including the United States. In the event, American-based corporate interests have successfully patented individual human gene sequences and a variety of transgenic animals and the techniques for their "manufacture," after overtaking the slower pace of discovery in state-funded research projects. Why is it not contrary to law to own and control the use of manufactured animal life? If the animal were a human, it would be unconstitutional to own a complete person, on grounds of slavery (42). Yet it is not unlawful through intellectual property rights to own or as

it were to "enslave" the genetic building blocks that make up both animals and humans.

BIBLIOGRAPHY

1. E.g. Statistics from the OECD in E. Ronchi, *OCDE/GD(96)* **167**, 5–9 (1996).
2. 42U.S.C.273 (1984); *Xenotransplantation: Science, Ethics & Public Policy*, Institute of Medicine, Washington, DC, 1996.
3. A. Trew, *Toledo Law Rev.* **29**, 271, 306–315.
   R.M. Arnold and S. Youngner, *Transplant. Proc.* **27**, 2913 (1995).
4. E.g. Council of Europe, *Convention on Human Rights and Biomedicine*, ETS No. 164, Council of Europe, 1997.
   Protocol (Banning Human Cloning), ETS No. 168 (1998), adopted by 23 European states; *Human Fertilization and Embryology Act*, UK, 1990.
5. R.A. Weiss, *Br. Med. J.* **317**, 931–934 (1998).
   D. Butler, *Nature (London)* **391**, 320–325 (1998).
6. FDA Draft Guidelines, *Fed. Regist.* **61**, 49920–49932 (1996).
7. World Health Organization (WHO), *Xenotransplantation: Guidance on Infectious Disease Prevention and Management*, WHO/ZOO/98.1, WHO, Geneva, 1998.
8. Department of Health Advisory Group of the Ethics of Transplantation, *Animal Tissues into Humans*, H.M. Stationery Office, London, 1997.
9. F. Gao, *Nature (London)* **397**, 436–441 (1999).
10. *N.Y. Times*, February 25, Sect. A, p. 1 (1999).
11. D. Dickson, *Nature (London)* **377**, 185 (1995).
    H. Vanderpool, *Lancet* **351**, 1347–1350 (1998).
12. R. Wilmut, *Nature (London)* **385**, 810–813 (1997).
13. U.S. Pat. 5,843,280 (1998) approved for primate embryonic stem cells. See advice by President's National Bioethics Advisory Commission November 20, 1998: *www.bioethics.gov*
14. PL89-544. The amended text is in 7USC 2131–2156.
15. PL99-198.
16. *NIH Publ.* **92-3415.**
17. U.S. Constitution Article 1, Section 8; Title 35 USC para. 101.
18. CFR Title 9, Chapter 1, Subchapter A-"Animal Welfare."
19. National Research Council, *Guide to the Care and Use of Laboratory Animals*, National Academy Press, Washington, DC, 1996.
20. 21CFR part 312.
21. The Website for CBER is at *www.fda.gov*
    C. Wilson, *J. Virol.* **72**(4), 3082–3087 (1998).
22. *Fed. Regist.* **61**(185), 49919–49932 (1996); Health & Human Services Fact Sheet 26th May 2000: *www.hhs.gov/news/ress/2000/pres/20000526.html.* Test of revised Guidelines: *www.fda.gov/cber/xap.xap.htm*
23. U.S. Department of Health and Human Services, *Developing US Public Health Care Policy in Xenotransplantation*, US DHHS, Bethesda, MD, 1998. For text, use "Search" at Web site: *www.fda.gov*
24. 42 USC 262,264.
25. *Canterbury v. Spence*, 464 F2d. 772 DCCircuit (1972).
26. 45 CFR 46; 21 CFR Part 50.
27. F. Bach, *Nat. Med.* **4**, 141–144 (1998).
28. W. Heinene, *Lancet* **352**, 695–699 (1988).
    United Kingdom Xenotransplantation Interim Regulatory Authority, *Workshops on Porcine Endogenous Retroviruses*, UK Xenotransplant, Interim Regul. Auth., 1999.
29. J. Rostand, *Can Man be Modified?* (J. Griffin, transl.), Basic Books, New York, 1959.
    Institute of Medicine Report, Washington, DC, 1996, Ch. 2 (note 6 above).

29a. FDA Revised Guidelines on Xenotransplantation Full Text (May 2000): *www.fda.gov/cber/gdlns/xeno0500.txt*

30. *Economic Aspects of Biotechnology Related to Human Health*, OECD Report, 1998, DSTI/STP/BIO(98)8/FINAL.
31. Council of Europe Website is at *www.stars.coe.fr*
32. *Explanatory Report to the Convention on Human Rights and Biomedicine*, Council of Europe, 1997, DIR/JUR(97)1.
33. Council of Europe Recommendation R(97)15.
34. Council of Europe, 1998, Doc.8166; Doc.8264.
35. Council of Europe, *Xenotransplantation*, Recommendation 1399, Council of Europe, 1999, (1).
36. European Union, 1998, Directive 98/44/EC.
37. G. Mudur, *Br. Med. J.* **318**, 79 (1999).
38. UKXIRA, *First Annual Report, May 1997, August 1998*, Department of Health, UK, 1998.
39. Department of Health, HCS1998/126; Web site: *www.open.gov.uk/doh/coinh.htm*
40. *Diamond v. Chakrabarty* et al., 65 L ed 2d 144 (1980); TRIPS amendment to USC Title 35, para. 154 (1995); Utility patents defined: U.S. Constitution Article 1, Section 8; USC Title 35, para. 101–103.
41. U.S. Pat. 4,736,866 (1988).
    A. Kimbrell, *The Human Body Shop*, Harper Collins, New York, 1995, pp. 188–202.
    U.S. Office of Technology Assessment Report, *Patenting Life*, U.S. Office of Technology, Washington, DC, 1990.
42. U.S. Constitution, Amendment XIII: Neither slavery nor involuntary servitude ... shall exist within the United States.

See other entries ANIMAL, MEDICAL BIOTECHNOLOGY, POLICY, WOULD TRANSGENIC ANIMALS SOLVE THE ORGAN SHORTAGE PROBLEM?; FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE; see also RESEARCH ON ANIMALS entries; TRANSGENIC ANIMALS: AN OVERVIEW.

ANIMAL MEDICAL BIOTECHNOLOGY, POLICY, WOULD TRANSGENIC ANIMALS SOLVE THE ORGAN SHORTAGE PROBLEM?

JEFFREY L. PLATT
SARAH C. HENSEL
Mayo Clinic
Rochester, Minnesota

OUTLINE

INTRODUCTION

The transplantation of organs or tissues between individuals of different species, such as the transplanting of a baboon heart into a human, is called xenotransplantation. Developing the technologies needed to safely carry out xenotransplants has been a long-standing goal, as it would allow the widespread application of organ transplantation and, potentially, other benefits, as discussed below. Successful application of xenotransplantation, however, has been impeded by several hurdles. Recent progress in identifying the molecular basis of the hurdles to xenotransplantation, however, has allowed the genetic engineering of source animals to address these hurdles, those being (*1*) the immune response of the recipient against the organ graft, (*2*) the functional limitations of the foreign tissue, and (*3*) the possibility of a xenograft introducing an infectious disease into the recipient. This article summarizes the rationale and hurdles to xenotransplantation, the potential application of genetic engineering to address these hurdles, and the ethical issues related to xenotransplantation. The reader is referred to other reviews for more detailed consideration of these topics (1).

RATIONALE FOR XENOTRANSPLANTATION

The field of transplantation began during the early years of the twentieth century, when the development of the vascular anastomosis, a surgical technique allowing the suturing of the cut ends of blood vessels, provided the means for transplanting organs from one individual to another (2,3). The transplanting of human organs, also known as *allotransplantation*, was not being undertaken at that time, however, because it was not clear how human organs could be obtained in an ethical manner in the absence of brain death laws that are now in place today, and allow for the harvesting of human organs for transplantation. For instance, because the organs of the recently deceased contained living cells, the organs could be argued to be living, and thus surgical removal was considered unethical. Accordingly, the first efforts to transplant organs into humans were xenotransplants, the organs originating from sheep and pigs, thereby avoiding that particular ethical conflict (4).

Today there are at least three reasons for interest in transplanting animal organs into humans. The first reason is that animal organs would be able to supplement the very limited supply of human organs available for transplantation. Indeed, the shortage of human organs for transplantation is widely seen as the most urgent problem in the clinical treatment of patients with organ failure. So great is the shortage of donor organs, that in the United States, as few as 5 percent of the organs needed ever become available (5), with the problem being even worse in countries where use of human organs is discouraged by law or custom. Furthermore, the shortage of human organs for transplantation sometimes forces the allocation of organs based on social rather than purely medical criteria. While improvements in immunosuppressive therapy and the introduction of immunological tolerance might lessen the shortage of organs by increasing the duration of organ graft survival, this advantage will likely be stifled by the ever-increasing demand for organs and tissues.

A second reason for interest in xenotransplantation stems from the possibility that an animal organ might offer some advantages over the use of a human organ for transplantation. For example, use of an animal organ would allow the transplant procedure to be planned well in advance. The advance planning of a transplant, in turn, allows for the pre-treatment of the patient in ways that more effectively suppress immune responses. Advance planning also allows screening of the donor to minimize infectious risks, and it reduces injury to the graft. The use of animals as a source of organ transplants may also allow the application of transplantation in parts of the world such as Asia, where the use of human tissue is discouraged for cultural reasons. Finally, the use of animal tissues may avert infection of the transplant with human viruses. Such was the motivation behind recent attempts to transplant baboon livers into patients with hepatitis B infection, and baboon bone marrow into an individual with AIDS.

A third reason for interest in xenotransplantation is the possibility that the use of animals would allow for the genetic or biochemical manipulation of the donor, so as to lower the risk of rejection of the organ, or even to express new genes or biochemical processes that benefit the patient. Genetic engineering of large animals such as pigs, to alleviate rejection, has already been undertaken. As a further step, one can envision genetic engineering being carried out to improve the function of an organ transplant, or even to impart new and novel functions on the transplant, such as the secretion of a needed protein, by a patient, to treat a disease.

SELECTION OF ANIMAL SOURCES FOR XENOTRANSPLANTATION

The most important characteristics of the animals used as sources of xenografts are summarized in Table 1. While it might seem intuitive that animals genetically similar to humans, such as chimpanzees and baboons, would be preferable as sources of xenografts, there are a number of factors that argue against this approach and suggest that nonprimates, especially pigs, are preferable. First, and most important, the animals most genetically similar

Table 1. Preferred Characteristics of Animals Used as Sources of Xenografts

Characteristic	Goal	Example
Available in large numbers	Unlimited source	Pig
Size suitable for human	Organs large enough for full sized adults	Pig
Genetically close to human	Minimize immune and biochemical incompatibility	Nonhuman primate
Minimal risk of zoonosis	Ideal donor is from species in regular contact with humans and with well characterized flora	Pig
Ease of breeding	Short gestation time and large litters	Pig
Ease of genetic engineering	Ability to introduce transgenes	Pig

to humans, nonhuman primates, such as baboons, are not available in large numbers, and their relatively small size may prevent the use of some organs in humans. In contrast, the availability of pigs is unlimited, and the size of pig organs is appropriate for transplantation into humans. Second, nonhuman primates frequently harbor viruses, such as the herpes B virus, that are difficult to detect and can be fatal if transferred to humans. Pigs harbor relatively few infectious agents that are harmful to humans, and these agents can generally be eliminated from herds. Third, owing to short gestation time and large litters, pigs can be easily bred, and genes can be introduced into lines of pigs using transgenic techniques. Nonhuman primates are difficult to breed and cannot be engineered genetically with such ease.

THE HURDLES TO XENOTRANSPLANTATION

While various factors favor the use of lower animals, such as pigs, as sources of organs for transplantation into humans, there are still formidable hurdles to carrying out such transplants. The most daunting hurdle to the clinical application of xenotransplantation is the immune response of the recipient against the graft, leading to graft rejection. Recent years have brought about new information regarding the immunology of xenograft rejection, as well as highly specific techniques, such as the genetic engineering of transplant donors, to deal with rejection. In fact, recent successes in dealing with the immune mechanisms of rejection have given rise to two other issues, the physiological limitations of a xenogeneic organ transplant and the possibility that a xenogeneic organ transplant might serve as a vector for introducing novel pathogens into human populations, a process known as *zoonosis*.

IMMUNE RESPONSES TO XENOTRANSPLANTATION

The immune responses to an organ xenograft consist of both "natural" (or innate) immunity, and elicited immune responses. Natural immunity, primarily consisting of natural antibodies and complement but also, in some cases, of natural killer cells, exists without prior exposure to foreign cells or substances. It normally provides an immediate, highly potent defense against extracellular microorganisms. Elicited immune responses, including T cell mediated responses and production of antibodies, following immunization with a foreign cell or substance, provide a defense against intracellular organisms, or organisms less virulent than those targeted by natural immunity.

Natural and elicited immune responses have important roles in the rejection of different types of transplants. While natural immunity is not generally involved in the rejection of organs transplanted between individuals of the same species, that is, allografts (except when the transplants are carried out across blood groups), it is the first and most severe type of immunity directed against organ xenografts. Elicited immune responses contribute to the rejection of both allografts and xenografts.

Hyperacute Rejection

An organ transplanted from a pig into a human would first be subject to hyperacute rejection (6), the most severe and violent immune response known. Hyperacute rejection begins almost immediately upon perfusion of the graft with the blood of the recipient. The graft, initially appearing pink and normal, becomes mottled, then deep red. The flow of blood to the graft declines sharply, then ceases, and the graft is destroyed over a period of minutes to hours. Microscopic analysis of the graft reveals that blood has leaked through small blood vessels, and clots, consisting predominantly of platelets, are formed. Hyperacute rejection of porcine xenografts by primates is mediated by xenoreactive natural antibodies and the complement system of the recipient. Research within the past decade has shed light on the molecular mechanisms by which natural antibodies and complement cause hyperacute rejection, with this knowledge leading to the development of new and insightful therapeutic approaches to this problem.

Xenoreactive Natural Antibodies. Natural antibodies exist in the circulation without a known history of immunization with a foreign substance or organism. For example, some natural antibodies recognize the blood group A or B substances and, thus, define the human blood groups. Some other natural antibodies recognize the cells of foreign species and, as such, are referred to as xenoreactive natural antibodies.

Hyperacute rejection of porcine organ xenografts by primates is initiated by the binding of xenoreactive natural antibodies to the endothelial cells that line blood vessels in the newly transplanted xenograft (7,8). Xenoreactive natural antibodies are present in the circulation of all normal individuals without a known history of exposure to animal cells (9,10). Xenoreactive natural antibodies in humans are mainly directed against a saccharide, Galα1-3Gal (11,12). The importance of Galα1-3Gal as the primary barrier to xenotransplantation was recently demonstrated by the observation that removal

of anti-Galα1-3Gal antibodies from baboons prevents the hyperacute rejection of the pig organs transplanted into the treated baboons (13).

Although the identification of the relevant antigen for pig-to-primate xenotransplantation allows specific depletion of the offending antibodies, more enduring and less intrusive forms of therapy would be preferable. One approach to overcoming the antibody–antigen reaction is to develop lines of pigs with low levels of expression of Galα1-3Gal (14,15). The most obvious approach to developing such lines of xenograft donors might be to "knock out" the enzyme that synthesizes the critical sugar, α1,3-galactosyl transferase. Unfortunately, embryonic stem cells are not available for pigs, so this approach cannot be applied with current technology. The first tested method for modifying expression of α1,3-galactosyl transferase involved expression of a glycosyl transferase, which would terminate sugar chains with a sugar other than Galα1-3Gal (16,17). Transgenic mice and pigs expressing the H-transferase have increased expression of H antigen, as expected, and decreased expression of Galα1-3Gal. Whether the expression of a glycosyl transferase, such as H transferase, will prove sufficient to eliminate hyperacute rejection is yet uncertain, but studies using isolated porcine cells have demonstrated that with H-transferase expression, complement mediated lysis significantly decreases.

Complement Activation. The complement system consists of more than 20 proteins that can assemble into complexes that can facilitate the engulfing of foreign organisms by inflammatory cells and cause the death of foreign organisms and foreign cells. Through these functions, the complement system provides the most potent line of defense against severe infections. In addition to helping in defense against infection, the complement system is involved in the development of various immunological diseases, among which is the rejection of xenografts.

Hyperacute rejection is caused by the activation of the recipient's complement system on donor blood vessels (18,19). Triggered by the binding of xenoreactive antibodies to graft endothelium (6,18), complement activation causes loss of the integrity of the endothelial lining of blood vessels and induces abnormal functions in blood vessels.

Xenografts are especially susceptible to complement-mediated injury. The basis for this susceptibility is an intrinsic incompatibility of the recipient's complement system with complement regulatory proteins expressed in the graft. Complement activation is amplified in a xenograft because the mechanisms that prevent complement activation from causing inadvertent injury to normal cells fail to protect the xenograft from the foreign complement of the recipient (20).

Under normal conditions the complement system is inhibited by proteins in the plasma and on the surface of cells (21). These proteins, called *complement regulatory proteins*, protect normal cells from inadvertent injury during the activation of complement (some of the reactions of the complement cascade occur in the complement deposition of cells residing in the vicinity of infectious organisms). Complement regulatory proteins function in a species-restricted fashion, which is to say, they inhibit complement of the same species far more effectively than they inhibit complement of foreign species. Accordingly, the complement regulatory proteins expressed in a xenograft are ineffective at controlling the complement cascade of the recipient, and thus the graft is subject to severe complement-mediated injury (20,22).

Therapeutic Considerations. One approach to preventing hyperacute rejection of xenografts is to administer complement inhibitors to the recipient, such as cobra venom factor, a complement inhibitory protein found in the venom of cobra snakes, or soluble complement receptor type 1, a recombinant protein that inhibits complement at the level of C3 (23), both of which have been found highly effective. One impediment to using complement inhibitors such as these for the prevention and treatment of xenograft rejection is that the inhibitors also prevent the utilization of the complement system to protect the recipient against infectious organisms.

To address this problem, and to overcome the incompatibility of complement regulatory proteins discussed above, several biotechnology groups have developed lines of pigs transgenic for human complement regulatory proteins and, thus, are able to control human complement reactions occurring in transplanted organs (20,24,25). These efforts have focused on expression of human decay accelerating factor, which regulates complement at the level of C3, and CD59, which regulates complement at the level of C8 and C9 (25,26) or decay accelerating factor alone (24). Recent studies have demonstrated that the expression of even low levels of human decay accelerating factor and CD59 in porcine organs prevents hyperacute rejection by primates (25,26). These results, and the dramatic prolongation of xenograft survival brought about by expressing higher levels of human decay accelerating factor in the pig (27), underscore the importance of complement regulation as a determinant of xenograft outcome.

While hyperacute rejection was once considered to be the most daunting hurdle to the clinical application of xenotransplantation, the development of methods for the specific depletion of xenoreactive antibodies and the inhibition of complement have shown that hyperacute rejection can be reliably prevented. This progress has also disclosed another type of rejection — acute vascular rejection — that may occur when hyperacute rejection is prevented.

Acute Vascular Rejection

If hyperacute rejection of a xenograft is averted, a xenograft is subject to the development of acute vascular rejection (28,29). Acute vascular rejection (sometimes called *delayed xenograft rejection*) may begin within 24 hours of connection to the recipient's circulation, and it leads to graft destruction over a period of days to weeks (28,30). It is characterized by injury to endothelial cells and widespread formation of intravascular clots. Although the cause of acute vascular rejection is not completely understood, there is growing evidence that it is triggered, at least in part, by the binding of xenoreactive

antibodies to the graft. The importance of xenoreactive antibodies in triggering acute vascular rejection is suggested by the findings that (*1*) anti-donor antibodies are present in the circulation of allograft recipients with acute vascular rejection, (*2*) depletion of anti-donor antibodies delays or prevents the occurrence of acute vascular rejection (31), and (*3*) administration of drugs that inhibit the synthesis of anti-donor antibodies delays or prevents the onset of acute vascular rejection (27). Recent studies suggest that the antibodies that trigger acute vascular rejection might include antibodies directed against Galα1-3Gal. Other factors that may contribute to the development of acute vascular rejection include complement, endothelial cell activation, and natural killer cells.

Therapeutic Considerations. The physical removal of anti-donor antibodies from the xenograft recipient, or inhibition of antibody synthesis, as described above, may effectively prevent acute vascular rejection of experimental xenografts. However, both types of treatments expose the graft recipient to potential complications. Furthermore, while hyperacute rejection can be prevented by temporary treatments, acute vascular rejection poses a more significant hurdle because therapies need to be provided on an ongoing basis. Accordingly, it would be desirable to address this problem through genetic engineering. The various approaches which might be used to deal with acute vascular rejection include removal of xenoreactive antibodies combined with administration of immunosuppressive drugs to limit synthesis of new antibody, induction of immunologic tolerance, and genetic engineering to decrease expression of Galα1-3Gal. In addition to lowering antigen expression, it is likely that expression of human complement regulatory proteins will be helpful in preventing complement mediated injury from contributing to graft injury. Also under consideration is the expression of molecules that inhibit either endothelial injury or the abnormal functions exhibited by injured endothelium.

Accommodation. Fortunately, the presence of anti-donor antibodies in the circulation of a graft recipient does not inevitably trigger acute vascular rejection. Some years ago it was found that if anti-donor antibodies are temporarily depleted from a recipient, an organ transplant may be established so that rejection does not ensue when the anti-donor antibodies return to the circulation (20). This phenomenon was referred to as "accommodation" (20). Accommodation, if it can be established, may be especially important in xenotransplantation, as it would obviate the need for certain ongoing interventions. One potential approach to accommodation may be the use of genetic engineering to reduce the susceptibility of an organ transplant to acute vascular rejection and endothelial cell activation associated with it (32).

Elicited Immune Responses. Organ transplants are subject to elicited immune responses leading to rejection. In contrast to the natural immune response, which exists without prior exposure to foreign cells or antigens, elicited immune responses are brought about by exposure to these entities. For example, the immune responses engendered by administration of vaccines are "elicited" responses.

The elicited immune responses that cause the rejection of transplants between individuals of the same species (allotransplants) can be effectively controlled by conventional immunosuppressive therapy, using drugs such as cyclosporine. There is concern, however, that the immune responses elicited by xenotransplantation will be more severe than the immune responses elicited in allotransplantation and that, accordingly, the responses to xenografts may not be agreeable with conventional immunosuppressive therapy. One reason the immune response to a xenograft may be especially intense is that xenografts have a great variety of antigenic proteins, and that introduction of these proteins may lead to recruitment of a diverse set of "xenoreactive" T lymphocytes. Another reason is that the binding of xenoreactive antibodies to the graft, and activation of the complement system, may lead to amplification of elicited immune responses (33). For example, activation of complement in a graft may cause activation of antigen presenting cells, in turn, stimulating the T cell responses that lead to cellular rejection.

Xenografts may be especially susceptible to immune responses mediated by natural killer cells. Natural killer cells are lymphocytes that recognize and kill tumor and virus-infected cells. The recognition of abnormal cells occurs through the function of two types of cellular receptors. One type of receptor recognizes abnormal carbohydrates and, upon doing so, up regulates natural killer cells' activities. The other type of receptor recognizes major histocompatibility antigens and down regulates natural killer activity (major histocompatibility antigens may be expressed at a decreased level or in an abnormal way by tumor cells or cells infected with viruses). The reason natural killer cells may be active against xenografts is that their carbohydrates recognize Galα1-3Gal (34) and their major histocompatibility receptors may fail to recognize the major histocompatibility antigens of the donor species (35). Thus, human natural killer cells may be especially active against xenogeneic cells because of stimulation and a failure to down regulate natural killer cell functions.

Still another type of elicited immune response to a xenograft might be the production of antibodies against foreign substances in the graft. As already mentioned, natural antibodies are specific for Galα1-3Gal and are produced in increased amounts after xenotransplantation. In addition to these antibodies, there also occurs the production of antibodies against "new" antigens (36). The identity of these antigens and the importance of the antibodies directed against them are still unclear. However, experience with experimental xenografts would suggest that the production of these antibodies can be suppressed by measures that inhibit T lymphocytes, an observation that is explained by the likelihood that antibody secretion by B lymphocytes depends on the function of T lymphocytes.

Another significant hurdle may be the elicited humoral response to xenotransplantation. Such a response in allotransplants, occurring over a period of months, is typically

directed against major histocompatibility antigens and has been thought to cause acute vascular rejection. There is every reason to believe that this will be true of xenografts as well. Thus far, studies in xenografts over periods of a few weeks to months suggest that the major immune responses in recipients given immunosuppressive therapy are directed against Galα1-3Gal (36,37), appearing to be an enhancement of the natural immune response. However, humoral responses against other antigens might be elicited by xenotransplantation. If elicited humoral immune responses occur, the responses may be addressed by immunosuppressive therapies or by the development of immunological tolerance. On the other hand, development of genetic approaches to dealing with elicited humoral immune responses must await the identification of the antigens recognized.

Therapeutic Considerations. One major question in xenotransplantation is whether or not the immunosuppressive drugs currently available which so effectively hold the rejection of allografts in check, or those in development, would be able to control the elicited immune responses to xenografts. Yet another question is whether the xenograft can be genetically modified to delimit the elicited immune response. Efforts to control the natural immune barriers to xenotransplantation will likely contribute to limiting the elicited immune response. How a xenogeneic donor could be further modified to limit elicited immune responses is still uncertain but an important area of investigation.

PHYSIOLOGICAL HURDLES TO XENOTRANSPLANTATION

Progress in addressing some of the immune hurdles to xenotransplantation has brought into focus the question of the extent to which a xenotransplant would function optimally in a foreign host. It is known that pig kidneys and pig lungs can replace the most important functions of the primate kidney and primate lung, respectively (38–41). However, subtle defects in physiology across species may exist. Furthermore, organs such as the liver that secrete a variety of proteins and depend on complex enzymatic cascades may prove incompatible with a primate host. Accordingly one important application of genetic engineering in xenotransplantation may be the amplification or modulation of xenograft function to allow for more complete establishment of physiologic function or to overcome critical defects. For example, recent studies by Akhter et al. (42) and Kypson et al. (43) showed that the function of cardiac allografts can be improved by expression of genes encoding beta-adrenergic receptors. Expression of these genes could be adapted to the xenotransplant in order to improve cardiac function. The key question, then, is which of the many potential defects actually need to be repaired.

Another potential hurdle is the possibility that the xenograft may disturb normal metabolic and physiologic functions in the recipient. For example, Lawson has shown that porcine thrombomodulin fails to interact adequately with human thrombin and Protein C to generate activated Protein C (44), a molecule that is critical for the control of coagulation. This defect could lead to a propensity for formation of blood clots in the graft. Of even greater concern is the possibility that a xenogeneic organ, such as the liver, might release substances that would promote abnormal clotting of the blood of the recipient. While a great many physiologic defects can be detected at a molecular level, the critical question will be which of these defects are of importance for the well-being of the recipient, and must therefore be repaired by pharmaceutical or genetic means, and which are innocent defects.

In addition to using genetic engineering to overcome physiological incompatibilities, there is the possibility that genetic engineering might be used to impart novel functions on the xenograft. The animal organ or tissue would then be used as a vehicle for introducing a foreign gene. The genetic engineering of pigs, through transgenic techniques, and the use of these transgenic tissues and organs, as grafts, has certain advantages over conventional gene therapy. First, the genetic material introduced into the genome of the pig can be expressed at high levels in all cells of a given type. Second, since the genetic material is expressed in stem cells, it is passed on to subsequent generations, and therefore the genetic manipulation need not be carried out repeatedly.

ZOONOSIS

The emerging success of experimental xenotransplants and the impending therapeutic trials also bring to the fore the question of zoonosis, that is, infectious disease introduced from the graft into the recipient. The transfer of infectious agents from the graft to the recipient is a well-known complication of allotransplantation. However, the extent of this particular risk of transplantation can generally be estimated so that a decision can be made based on the eventual risk versus the potential benefits conferred by the transplant. The greatest concern of zoonosis in xenotransplantation is not so much the risk to the transplant recipient as it is the risk that an infectious agent will be transferred from the recipient to the population at large. Fortunately, all microbial agents known to infect the pig can be detected by screening and can be potentially eliminated from a population of xenotransplant donors. There is concern, however, that the pig may harbor viruses, such as retroviruses, that might become activated and transferred to the cells of the recipient. For example, Patience et al. recently reported that a C type retrovirus, endogenous to the pig, could be activated in pig cells, leading to release of particles that can infect human cell lines (45). Whether or not this virus or other endogenous viruses can actually infect across species, and whether or not such infection would lead to disease, is still not certain, but recent studies suggest that the risk of infection may be low (46).

If cross-species infection does prove to be a significant hurdle, genetic therapies might also be used to address this problem with further possibility of introducing genes that might target the infectious agent. The simplest genetic therapy would involve "breeding out" the organism, but this approach might fail if the organism were widespread in the pig populations or integrated at multiple loci. Some genetic therapies have been developed to potentially

control HIV (47). While these therapies have not generally succeeded, because it has been difficult or impossible to gain expression of the transferred genes in stem cells and at levels sufficient to deal with high viral loads, the application of such therapies might be much easier in xenotransplantation because the therapeutic genes could be carried by all of the animal cells.

Another concern is that infection of the xenograft might fail to engender an immune response in the recipient that would otherwise lead to control of that infection, because antigenic peptides of the infecting organism would be presented in association with xenogeneic MHC molecules. If this problem proves limiting in xenotransplantation, the problem could potentially be overcome by eliminating those microorganisms through mechanisms described above, or by the introduction of "generic" MHC molecules, which the host might recognize.

A SCENARIO FOR THE CLINICAL APPLICATION OF XENOTRANSPLANTATION

The past few years have brought significant progress in defining the immunological, physiological, and infectious disease hurdles to xenotransplantation. This information has been exploited in the development of incisive new therapies for overcoming these hurdles. Perhaps the most exciting of these therapeutic strategies is the development of genetically engineered animals for use as a source of organs for transplantation. To this point, research on genetic manipulation has focused on the most severe of the immune hurdles, the action of complement on graft blood vessels. This effort will no doubt continue while new work will seek to introduce genes for dealing with the effects of cellular immunity or physiological defects or infectious agents. With these, it can be envisioned that xenotransplantation will enter the clinical arena in a stepwise fashion. First, there will occur free tissue xenografts and extracorporeal use of xenogeneic organs. Limited clinical trials of this sort are in progress (48–50) and there is encouraging early evidence that neural cells may endure in a human recipient (50). Next, xenogeneic organs will probably be used as "bridge" or temporary transplants. Bridge transplants will not solve the problem of organ shortage, but the transplants will allow the gathering of vital information regarding the remaining immune and biological hurdles. Third, there will be the use of porcine organs as permanent replacements, but this use will probably be restricted to patients who cannot receive a human organ allograft. Only with further refinements may there eventually be a fourth and final step, in which xenotransplantation is used as an alternative to allotransplantation.

ETHICAL ISSUES IN XENOTRANSPLANTATION

The ethical issues as related to xenotransplantation can be considered to be of three categories: (*1*) animal welfare, (*2*) clinical experimentation, and (*3*) societal implications. With the exception of the third category, the issues are, for the most part, shared with many fields of experimental medicine and are not unique to xenotransplantation. Some of the ethical aspects of xenotransplantation have been reviewed (51–59).

Animal Welfare

One aspect of the ethics of xenotransplantation relates to the use of animals as a source of organs and tissues for transplantation into humans. The ethical question is whether or not it is proper or just to "use" animals for the benefits of human beings. Some believe that animals should not be used for such purposes, while others believe such use of animals should be permissible. The question, then, is whether society should regulate the use of animals according to the wishes of one group or to the other. One approach to this question has been to reason that if it is permissible to "use" animals for food, it should be permissible to "use" animals as a source of organs or tissues for transplantation. Thus, to the extent that society is willing to countenance the use of animals for food, it would seem difficult to justify development of laws prohibiting the "use" of animals for other purposes that are beneficial to humans provided that the generally accepted facets of animals welfare are addressed.

If it is not unethical to use animals for organ transplantation, ethical issues still remain pertaining to animal welfare or treatment. Similar to the broader subject of animal welfare, these issues include how the source animals are raised, how they will be handled, the potential occurrence of suffering, and the subjection of animals to genetic engineering. For example, to keep animals as free as possible from infectious agents, it will likely be necessary to raise them in isolation. Is this unethical, and, if so, how might this issue be addressed? As another example, it is possible that genetic modification may impair an animal's health and cause suffering.

Of the issues discussed above, one of particular interest is genetic engineering. Genetic engineering involves inserting genes into the germ line (accomplished by various means, e.g., a direct injection of genetic material into a fertilized egg), thereby passing that gene on to subsequent generations of animals. There seems to be a fear, among some, of the changes that may occur as a result of this type of deliberate introduction of genetic material, for instance, in the case of genetic engineering of plants and livestock. Some question whether or not the species will change as a result of genetic manipulation or whether the animal will become more "human." Perhaps some of this fear stems from the lack of knowledge that genetic material can be naturally exchanged between species as a result of infection by retroviruses.

Clinical Experimentation

A second category of ethical issues relates to the use of human subjects in clinical xenotransplantation trials. The recipient of a xenotransplant, as the subject of an experimental procedure, must be adequately informed about that procedure, including the anticipated risks and the limitations of knowledge regarding the anticipated benefits. However, these aspects of the use of human subjects do not differ in substantive ways from any use of human subjects in medical research. For example, similar

weighing of uncertain risks and benefits might also occur with the use of an experimental device or an experimental device or an experimental human-to-human transplant. Some of the matters of ethics to be addressed include (*1*) how to obtain consent for an experimental procedure that will surely require long-term follow-up or monitoring, (*2*) how the right of the patient to withdraw will impact on the individual, (*3*) how to avoid encouraging patients to accept greater-than-usual risks for the benefit of future patients, (*4*) how the experiment might affect quality of life — as an increase in risk of organ rejection means a need for greater immunosuppressive therapies to an unknown end or risk, and (*5*) how to forecast what the possible psychological effects of such an experiment might be.

Ethics of Society

While the ethics of animal use and human experimentation are not unique to xenotransplantation, the social implications may be so. The most important ethical question for society would seem to relate to the possibility that the recipient of xenotransplant might become infected by microbial agents contained in the transplant and that agent might then be spread more broadly in the population. The possibility that a medical or surgical procedure could have implications for the broader community is a relatively novel aspect of medical ethics, but one that does exist outside of xenotransplantation. For example, the use of antibiotics by an individual patient may alter the types of organisms present in a hospital and, thus, impose a risk on other individuals. Similarly, the use of inhaled substances could impact on those in contact with a given patient. While it is now thought by many experts that the risk of an epidemic being caused by a xenotransplant is very low, the potential impact of an epidemic keeps this issue at the fore.

There are other ethical questions, the impact or scope of which might extend beyond the recipient of a xenotransplant. For example, it is thought that family and/or sexual contacts of the recipient may have to provide consent, since they may be at risk. As another example, there is the question of whether or how the recipient of a xenotransplant might withdraw consent after a xenotransplant is in place; such withdrawal would lead to cessation of monitoring for potential infectious agents, which would, in turn, place society at risk. Yet another issue involves the possible costs of xenotransplantation to society, for, although transplantation has proven to be less costly than other chronic care for organ failure, it is uncertain whether this advantage will apply to xenotransplantation. A further ethical question includes how to achieve justice in distribution of human versus animal organs, assuming human organs are preferred. Furthermore, this preference may result from another issue, that of the possibility of a low level of public acceptance, at least at the outset, as many may perceive xenotransplantation as "unnatural" or tampering with nature or God. Important questions then become: Does the use of animal organs change how we define ourselves as human? What is or isn't "natural"? Is it not natural that we evolve and obtain the knowledge that enables us to move forward in ways such as this? Finally, to what extent should the advantages and risks of xenotransplantation be subject to public discussion?

In addition, there are some issues that may be more political than ethical. One such issue is whether or not regulation occurring in one country, though not in another, would be meaningful. Another issue involves the role of the private sector, which will likely play a major role in the research and creation of transgenic breeds of porcine. Relative questions become: What is the role of the private sector in support of research or commercialization? Will patents be given for transgenic breeds and, if so, to whom? Furthermore, are the ethics of business too vastly different from those of science and medicine, and while their motivation is surely different, does this necessarily pose any ethical dilemmas or risk to the public? It would seem logical that though their motivation is, in fact, different, they would have a vested interest in ensuring a safe, quality "product," if not simply to ensure their own survival.

CONCLUSIONS

While ethical issues are salient, many do not differ greatly from the ethical issues raised in any experiments involving animal or human subjects. What makes xenotransplantation somewhat unique is that transplantation has a high profile, and it raises issues for society, although it is far from the only human endeavor to do so. Throughout history (e.g., with the advent of the railroad train) there has been a quandary about how advances in science will affect society and the state. Clearly, some of these issues relating to xenotransplantation will most likely remain unsettled, raising one further question: To what extent can one proceed, in activities that impact on society, without the existence of consensus?

ACKNOWLEDGMENT

I thank the National Institutes of Health for supporting my work in this area.

BIBLIOGRAPHY

1. J.L. Platt, *Nature* **392**(suppl.), 11 (1998).
2. A. Carrel, *Lyon Med.* **98**, 858 (1902).
3. C.C. Guthrie, *Blood-Vessel Surgery and Its Applications*, Longmans, Green & Co., New York, 1912.
4. E. Ullman, *Ann. Surg.* **60**, 195 (1914).
5. R.W. Evans, C.E. Orians, and N.L. Ascher, *J. Am. Med. Assoc.* **267**, 239 (1992).
6. J.L. Platt, *Hyperacute Xenograft Rejection*, CRC Press, Austin, TX, 1995.
7. R.J. Perper and J.S. Najarian, *Transplantation* **4**, 377 (1966).
8. G.R. Giles et al., *Transpl. Proc.* **2**, 522 (1970).
9. S.V. Boyden, *Adv. Immunol.* **5**, 1 (1964).
10. W. Parker et al., *J. Clin. Immunol.* **17**, 311 (1997).
11. A.H. Good et al., *Transpl. Proc.* **24**, 559 (1992).
12. M.S. Sandrin et al., *Proc. Natl. Acad. Sci. USA* **90**, 11391 (1993).
13. S.S. Lin et al., *Transpl. Immunol.* **5**, 212 (1997).

14. R.L. Geller, P. Rubinstein, and J.L. Platt, *Transplantation* **58**, 272 (1994).
15. C.G. Alvarado et al., *Transplantation* **59**, 1589 (1995).
16. M.S. Sandrin et al., *Nat. Med.* **1**, 1261 (1995).
17. A. Sharma et al., *Proc. Natl. Acad. Sci. USA* **93**, 7190 (1996).
18. J.L. Platt et al., *Transplantation* **52**, 214 (1991).
19. A.P. Dalmasso et al., *Am. J. Pathol.* **140**, 1157 (1992).
20. J.L. Platt et al., *Immunol. Today* **11**, 450 (1990).
21. D. Hourcade, V.M. Holers, and J.P. Atkinson, *Adv. Immunol.* **45**, 381 (1989).
22. A.P. Dalmasso et al., *Transplantation* **52**, 530 (1991).
23. H.F. Weisman et al., *Science* **249**, 146 (1990).
24. D. White and J. Wallwork, *Lancet* **342**, 879 (1993).
25. K.R. McCurry et al., *Nat. Med.* **1**, 423 (1995).
26. G.W. Byrne et al., *Transplantation* **63**, 149 (1997).
27. E. Cozzi et al., in D.K.C. Cooper et al., eds., *Xenotransplantation: The Transplantation of Organs and Tissues between Species*, Springer, Berlin, Germany, 1997, p. 665.
28. J.R. Leventhal et al., *Transplantation* **56**, 1 (1993).
29. W. Parker et al., *Immunol. Today* **17**, 373 (1996).
30. J.C. Magee et al., *J. Clin. Invest.* **96**, 2404 (1995).
31. S.S. Lin et al., *J. Clin. Invest.* **101**, 1745 (1998).
32. F.H. Bach et al., *Nat. Med.* **3**, 196 (1997).
33. D.T. Fearon and R.M. Locksley, *Science* **272**, 50 (1996).
34. L. Inverardi et al., *Transpl. Proc.* **28**, 552 (1996).
35. L.L. Lanier and J.H. Phillips, *Immunol. Today* **17**, 86 (1996).
36. K.R. McCurry et al., *Hum. Immunol.* **58**, 91 (1997).
37. A.H. Cotterell et al., *Transplantation* **60**, 861 (1995).
38. G.P.J. Alexandre, in M.A. Hardy, ed., *Xenograft 25*, Elsevier Science Publishers, New York, 1989, p. 259.
39. T. Sablinski et al., *Xenotransplantation* **2**, 264 (1995).
40. C.W. Daggett et al., *Surg. Forum* **47**, 413 (1996).
41. C.W. Daggett et al., *J. Thor. Cardiovas. Surg.* **115**, 19 (1998).
42. S.A. Akhter et al., *Proc. Natl. Acad. Sci. USA* **94**, 12100 (1997).
43. A.P. Kypson et al., *J. Thor. Cardiovas. Surg.* **115**, 623 (1998).
44. J.H. Lawson and J.L. Platt, *Transplantation* **62**, 303 (1996).
45. C. Patience, Y. Takeuchi, and R.A. Weiss, *Nat. Med.* **3**, 282 (1997).
46. K. Paradis et al., *Science* **285**, 1236 (1999).
47. B.A. Sullenger et al., *Cell* **63**, 601 (1990).
48. C.G. Groth et al., *Lancet* **344**, 1402 (1994).
49. R.S. Chari et al., *N. Engl. J. Med.* **331**, 234 (1994).
50. T. Deacon et al., *Nat. Med.* **3**, 350 (1997).
51. The Nuffield Council on Bioethics, London, 1996.
52. R. Downie, *J. Med. Ethics* **23**, 205 (1997).
53. M. First, *Kidney Int.* **58**, S46 (1997).
54. A. Daar and D. Phil, *World J. Surg.* **21**, 975 (1997).
55. J. Moatti, *An. N.Y. Acad. Sci.* **862**, 188 (1998).
56. B. Miranda and R. Matesanz, *An. N.Y. Acad. Sci.* **862**, 129 (1998).
57. J. Hughes, *J. Med. Ethics* **24**, 18 (1998).
58. A. Daar, *Bull. World Health Organ* **77**, 54 (1999).
59. G. Beauchamp, *Can. J. Surg.* **42**, 5 (1999).

See other entries ANIMAL, MEDICAL BIOTECHNOLOGY, LEGAL, LAWS AND REGULATIONS GOVERNING ANIMALS AS SOURCES OF HUMAN ORGANS; FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE; see also RESEARCH ON ANIMALS entries; TRANSGENIC ANIMALS: AN OVERVIEW.

B

BEHAVIORAL GENETICS, HUMAN

DAVID T. WASSERMAN
University of Maryland
College Park, Maryland

OUTLINE

INTRODUCTION

A variety of contemporary research programs in biology, psychology, and medicine investigate the magnitude and character of genetic influence on human behavior, affect, and cognition. Heritability studies attempt to tease out genetic from environmental effects, largely by comparisons involving twins and adoptees; association and linkage studies seek markers, and ultimately genes, associated with behaviors or traits of interest; neurobiological research explores causal pathways by which genetic variations may affect behavior. While specific research programs differ greatly in method and ambition, they share the assumption that it makes scientific sense to look for genetic contributions to important mental and behavioral differences among people.

Statistical techniques for teasing apart genetic and environmental influence date back to Francis Galton, who was — not coincidentally, critics insist — the founder of eugenics. In the half-century after Galton proposed the systematic study of heritable differences among people and the improvement of the genetic stock of humanity, heritability research evolved in close association with eugenic policies. The research, while not devoid of scientific achievement, proceeded in ignorance of the complex patterns and actual mechanisms of inheritance, and produced such retrospective embarrassments as Charles Davenport's work on the genetics of seafaring, which argued, in effect, that the love of the sea was a simple Mendelian trait. The policies, while supported in part by liberals as well as conservatives, included the restrictive immigration laws and sterilization campaigns in the United States and led, indirectly, to the programs of mass sterilization and "euthanasia" of the unfit in Nazi Germany (1,2).

Research on the genetics of human behavior and psychology underwent a period of understandable quiescence after the Second World War. In the past two decades, though, it has made a dramatic comeback. Two developments have arguably contributed to its resurgence: first, the advent of sophisticated techniques for isolating and manipulating genetic material; second, increasing public dissatisfaction with the optimistic environmentalism that supposedly dominated the social policy of the postwar era.

Much of the debate between critics and proponents of genetic research centers on the interpretation of these two trends. Proponents see the resurgence of genetic research on human behavior as driven by its scientific progress and a growing appreciation of the complexity of human behavior; they see the public as rejecting the environmental determinism of the postwar era much as it had rejected the genetic determinism of the pre-war era. They argue that judging contemporary behavioral human genetics in terms of the genetic research and eugenic policies of the first half of this century is "somewhat akin to attempting to explain the behavior of a butterfly by studying the caterpillar (or for that matter, understanding the fruitfly by studying the maggot)" (3).

Critics of behavioral genetics doubt the claimed metamorphosis; they are less inclined to see it as emerging from a cocoon than crawling out of an overturned rock. They see the research as exploiting the general retreat from an environmentalism that policy makers never seriously pursued; they regard the advances in molecular genetics, however useful in other areas of research, as serving a largely cosmetic role in behavioral genetics, lending a veneer of hard science to an enterprise that remains essentially confused and speculative.

This clash of interpretations helps frame the central issue this article will address: Do advances in molecular biology make it significantly more likely that behavioral genetics will avoid the scientific pitfalls and social abuses of earlier human genetic research? Or do behavioral geneticists misunderstand the scientific and social lessons of the past, making them likely to repeat it?

This article will not attempt to adjudicate the conflict between behavioral genetics and its critics so much as to clarify the sources of disagreement between them. Their disagreement, it will suggest, does not concern the specific methods, assumptions or findings of the research so much as its explanatory and practical value. Researchers believe that much can be learned about the causes and control of significant human traits and behavior by isolating genetic features associated with individual differences in those traits and behaviors, while critics believe that even if such features can be isolated, they will yield limited insight into causal and developmental processes, while creating a substantial risk of scientific oversimplification and social abuse.

SCOPE OF THE ARGUMENT

Some objections to human behavioral genetics are based on a denial that human conduct is subject to, or can be described by, the kind of law-like generalizations found in other sciences. These objections are directed against the generalizations offered by sociology no less than those offered by behavioral genetics. Other objections to behavioral genetics are to biological, as opposed to social or psychological explanations of behavior; still others are to individual, as opposed to situational explanations. Some of these more general objections, e.g., to the attempt to reduce human behavior and cognition to biology, are addressed in other articles in this Encyclopedia. This article will discuss these general objections only to the extent that they figure in the debate over genetics and behavior.

One such general objection is that biological accounts of mental traits and social behavior inevitably oversimplify or distort the phenomena they seek to explain. Thus, for example, a vast array of social behavior and interaction is lumped together as "criminality," a classification that assumes a common trait or disposition underlying the myriad of criminal activities and styles of transgression. Similarly, the richly varied forms of sexual attraction, play, and intimacy are dichotomized as hetero- and homosexuality. Researchers would respond that simplification and abstraction are essential to scientific progress. Even if the initial categories employed by the researchers are vague or coarse, they will be refined by additional research, and they may have significant heuristic value in generating hypotheses and developing theories. Simplification and abstraction are critical for sociological as biological generalizations about behavior.

Another general objection is that explanations based on individual differences divert attention from broader structural or institutional causes. For example, it is claimed that the attempt to attribute learning failures to individual factors — whether biological, like Attention Deficit Disorder, or social, like family dysfunction — diverts attention from the appalling state of public schools and public education. The attribution to biological differences may appear to be especially diversionary, since those differences may appear, unlike family dysfunction, to have no relation to structural or institutional causes. Researchers may insist in good faith that biological or genetic variations are only one type of causal factor among many, and that the effect of those variations is mediated by broader social factors. But policy makers will inevitably single out biological or genetic factors, since it is far easier to identify "bad apples" than to change a rotten system.

Other objections are specific to genetic as opposed to other biological causes — to genetic variations as opposed, say, to birth trauma. Those who regard genetic explanations as especially reductive or prone to abuse often assume that genetic causes have less mutable effects than other biological causes, and that the effects of genetic causes are essential to, or constitutive of, an individual in a way that the effects of trauma or insult are not. Both these assumptions, however, are mistaken. The effects of many genetic conditions are, or may be, largely or completely preventable, such as PKU, a single-gene disorder, causes severe retardation, but that effect can be avoided or mitigated by a modified diet. And if genetic effects are not immutable, neither are they essential or constitutive. Even if possessing all or most of one's genome is a necessary condition for personal identity, acquiring all or most of the traits associated with that genome in standard environments is not. We hardly think that a phenolallenine-free diet alters the identity of a child with PKU, even if the absence of the PKU mutation would (and that is debatable). But these beliefs may be as ingrained and recalcitrant as they are mistaken, and they may have a profound effect on the social reception of behavioral genetic research.

Behavioral genetics confronts all of the above objections — to individual, biological, and genetic explanations — because its defining feature is its focus on the effects of individual differences in genetic constitution. Humans beings share 99 percent of their genome with chimpanzees and more with each other; behavioral genetics is concerned with the relatively small portion that differs from person to person. Many critics believe this emphasis is misplaced; they question whether attempts to explain differences in human behavior and personality in terms of genetic differences are scientifically promising or socially beneficial. Some critics argue from an evolutionary perspective that while all human beings, or human males, may have genetic propensities for, say, violent or antisocial behavior, individual differences in the strength of those propensities are unlikely to have a genetic source. Even if genetic differences among individuals made them more or less prone to violent or antisocial behavior, they are likely to contribute far less to the understanding of human violence than genetic commonalities. Other critics fear that the search for individual genetic differences, whether or not scientifically justified, will stigmatize those individuals thought to be genetically predisposed, and the social groups to which they belong.

While behavioral genetics looks at a wide range of traits and behaviors, this article will focus on predispositions to criminal, violent, aggressive, impulsive and antisocial behavior. There are two reasons for this focus: first, research in these areas has engendered more public attention and controversy over the past decade than research on other traits, even intelligence; second, and more importantly, genetic research on antisocial and criminal behavior has taken much more of a molecular turn than genetic research on intelligence or other traits, bringing to behavioral genetics the powerful hopes and fears raised by the Human Genome Project.

CLASSICAL AND CONTEMPORARY BEHAVIORAL GENETICS

Heritability Research

Until techniques were developed in the 1970s for isolating and manipulating genetic and other molecular material, human behavioral genetics was largely confined to heritability studies. Heritability research takes off from the scarcely debatable observation that many behaviors and psychological traits "travel in families." But parents usually confound the assessment of genetic influence by providing their children with rearing environments as well

as genes; the transmission of behavior from one generation to the next can be attributed to either one.

Behavioral genetics exploits two processes which tend to tease apart these genetic and environmental contributions: twinning and adoption. Although the logic of these comparisons has been understood and debated for over a century, the implementation of careful studies only became possible with the systematic government record-keeping and refined statistical techniques of the twentieth century (4).

Twinning produces offspring that share either half their genes (dizygotic/DZ), the same proportion as in normal siblings, or all their genes (monozygotic/MZ). If the rearing environments of DZ twins can be assumed to be as much alike as those of MZ twins, and if other, less controversial assumptions are satisfied, then any greater similarity (concordance) in the behavior of the MZ twins can be attributed to their greater genetic commonality, and differences in the concordance of MZ and DZ twins can be used to estimate the heritability of the behavior or trait in the general population.

The second process that tends to separate out genetic and environmental contributions is adoption. A true experiment would randomly assign children at birth (or better yet, at conception) to other parents, and compare their traits and behaviors with those of their biological and adopted parents. Social practice very roughly approximates such an experiment by somewhat fortuitously if nonrandomly assigning children at some point after birth to adoptive parents. The greater the similarity of the children to their biological parents in the trait or behavior studied, the stronger the evidence of a genetic contribution. In addition to providing potentially corroborating evidence for twin studies, adoption studies help identify specific environmental factors and provide insight into the role of genetic factors in shaping the rearing environment, such as by evoking parental responses (5). Researchers can combine twin and adoption studies by examining the similarities of MZ twins reared apart — a prized commodity in human behavioral genetics.

Studies conducted over the past 40 years have reported significant heritabilities for a variety of psychiatric and behavioral conditions, including schizophrenia, intelligence, neuroticism, antisocial behavior, and property crimes. Interestingly, they have failed to find significant heritabilities for other behavior, such as violent crime, which the public assumes to have a heavy genetic loading, but such negative findings rarely receive the same press as positive ones (6,7). Among the strongest, and most surprising, results from twin and adoption studies have been negative — the consistent finding that "shared environment," the congeries of parental, domestic, and other local factors affecting all siblings equally, makes virtually no contribution to most behavioral and psychological states that have been investigated. The environment that appears to matter is the idiosyncratic environment of each sibling (8).

There has been much discussion about the validity of the assumptions on which these findings rest: For example, are the rearing environments of DZ twins as much alike as those of MZ twins, or are identical twins environmentally as well as genetically more alike, down to their identical wardrobes? Does the time adoptees spend with their biological parents, or do the adoption agencies' nonrandom placement practices, help explain the behavioral similarities of the adoptees and their biological parents? The researchers themselves are keenly aware of weaknesses in the assumptions and methods of previous studies, although they tend to be more optimistic than the critics about the prospects for dispensing with controversial assumptions or controlling for their violation.

There is a broader, more important debate about what to make of findings that a given trait or behavior has high heritability. Heritability studies do not identify the genetic or chromosomal variations responsible for traits or behavior. Rather, they calculate the population variance on some measure of a trait or behavior (e.g., IQ score, number of arrests), then use the assumptions about genetic similarity referred to earlier to assign percentages of that variance to heredity and environment. Thus a typical finding of an MZ/DZ twin study might be that "60 percent of the variance in IQ score is attributable to heredity." There are several critical points to make about this type of finding, about which researchers and their critics do not disagree.

First, heritability studies assess genetic contributions to differences in traits or behaviors. Features shared by virtually everyone in the population have little or no heritability; for example, the trait of "having two arms" will have very low heritability in a country where almost everyone does, and where most of those who do not are missing limbs because of dismemberment or infection rather than genetic disease or mutation (9). It is important to distinguish a trait's heritability from its genetic basis, a distinction critical to the evolutionary critique of behavioral genetics discussed below.

Second, in human heritability studies, the "environment" covers everything besides the genetic endowment to which variance in a trait or behavior can be attributed, from prenatal stress to adolescent peer pressure. Human studies do not plot the trait or behavior against an array of distinct environments, defined by the presence, absence, or level of quantified variables manipulated by the experimenter — the standard methodology in plant and animal behavioral genetics. The differences among the environments where human genetic variation is expressed are largely unmeasured and often unknown (10,11).

Third, heritability studies attempt to account for population variance in the trait or behavior; they reveal nothing directly about the comparative importance of genetic and biological factors in any individual. A particular level of heritability could reflect the effect of different genetic and environmental factors in different segments of the population, and could arise from the large effects of a small number of genes or the small effects of a large number. The claim, for example, that heredity accounts for 60 percent of the variance in IQ does not mean that 60 percent of an individual's IQ score, or deviation from the mean IQ, is genetically determined (if such claims are even intelligible) (9).

Fourth, heritability studies do not investigate or distinguish among possible causal pathways from gene

to trait or behavior. Thus, to take an example from Christopher Jencks, people with red hair might have below-average IQ solely because they are neglected and mistreated as children throughout the society. Common sense would treat this as an environmental explanation of their lower IQ. Nevertheless, because there might be little or no variation in the treatment or the performance of red-haired children across the environments actually studied, low IQ could be counted as genetic in a heritability study. The extent to which genes lower IQ by evoking discrimination is not reflected in the proportion of the variance assigned to heredity. The example becomes less fanciful if we substitute dark skin for red hair (11).

The indirect manner in which behaviors and mental traits are typically measured gives such alternative explanations further plausibility. If the researcher adopts arrests or convictions as a proxy for criminal behavior or disposition, or score on a written test as a measure of task proficiency, the causal pathway from the genes to the measured variable may not run through the trait or behavior at all: What may be transmitted genetically is not a propensity to commit crimes, but a lack of talent for concealment or evasion; not proficiency in a specific cognitive task, but a general facility with written tests.

Again, behavioral geneticists are well aware that the analysis of variance is quite different from the analysis of causation; that heritability studies face inherent limits in what they can reveal about the ways genes affect traits and behavior; and that observed heritability of a trait or behavior may vary with the measure chosen. Thus the American Society for Human Genetics nicely summarized the limitations of heritability research in a 1997 statement on the state of the research in behavioral genetics:

> The concept of heritability refers to the ratio of the genetic variance to the overall phenotypic variance. It is based on a specific situation involving a particular phenotype in a population with some array of genetic and environmental factors at a given time. It can differ from population to population and from time to time. It can change with age during development. It is important to keep in mind that heritability is a descriptive statistic of a trait in a particular population, not of a trait in an individual (5, p. 1266).

But researchers and critics disagree about the significance of these limitations. For researchers, heritability studies make a valuable contribution despite these limitations, by identifying traits and behaviors that are good candidates for genetic influence in light of the strong evidence that something is being genetically transmitted. Critics, in contrast, argue that the very attempt to partition causes of variance into environmental and genetic factors is misleading, if not meaningless. Given the indefinite variety of ways in which causally inert genetic traits such as skin color interact with powerful social forces such as racism to produce destructive social behavior, there is no reason to think that interesting or informative genetic explanations lie behind findings of high heritability even if they are valid on their own terms (5,10).

This difference in outlook is illustrated in the response to reports of high heritabilities for traits such as religious affiliation, where any genetic effect is obviously mediated by a vast array of cellular, somatic, and social variables (12). While researchers see such findings as intriguing (though clearly in need of a more complete causal account), critics regard them as a reductio ad absurdum: if religious affiliation is heritable, then heritability can have only the most tenuous relation to genetic causation.

Twin and adoption studies were not the only tools in the behavioral geneticist's armamentarium before the introduction of molecular genetics. During the 1970s some sophisticated mathematical techniques were developed for teasing out genetic and environmental contributions to behavior from a wider range of family data. These "model-fitting" techniques employ data from a variety of family relationships and a variety of populations. They have proved useful in overcoming one of the principal limitations of heritability estimates based on standard twin and adoption studies — their restriction to a particular population (5). These techniques have also been useful in incorporating the associations found between behavioral and psychological traits and specific genetic locations into models of gene–environment interaction, integrating the old behavioral genetics and the new.

Out with Heritability, in with Molecular Genetics

The clearest difference in outlook between behavioral geneticists and their critics is not in their evaluation of past or present research but in their assessment of the prospects for future research. Researchers and critics agree that heritability research will play a diminishing role in the behavioral genetics of the next century. But they disagree about the scientific potential of the techniques that are superseding it. These techniques are designed to link behavioral and mental traits to specific genes and to move beyond heritability research in several respects: to identify what is transmitted genetically, trace causal pathways from genes to traits or behavior, and thereby move partway from the population to the individual. But they also maintain the focus of heritability research on the genetic basis for differences among individuals.

Researchers expect that molecular and neurogenetic research will resolve the ambiguities about causation by tracing the complex pathways through which specific genes affect traits or behavior. This work, they claim, will place behavioral genetics on a solid biological foundation, and it will be less susceptible to oversimplification than prior research, since it will replace global estimates of heritability with narrow and testable causal hypotheses. They point to the progress made by medical genetics in identifying genes associated with a variety of disorders, many of which had not previously been regarded as "genetic." At the same time they recognize that even with the new molecular techniques, genetic contributions will be harder to identify, in part because genes may make a more modest and complex contribution to behavioral and psychological disorders than somatic disorders.

While researchers expect genetic effects on psychology and behavior to be modest and complex, they also expect them to be sufficiently strong and interpretable to yield insight into the causes of important human traits

and behavior. For example, Gregory Carey and Irving Gottesman speculate:

> The current generation of molecular genetic research is likely to uncover polymorphisms associated with behavior, and some of these loci will probably be correlated with antisocial behavior. No one is banking on a major "crime gene." Instead, most suspect that there will be a number of loci of small effect that partly influence temperament, motivation, and cognition. . . .The statistical predictability of these loci may be quite small. They may, however, prove quite important for unraveling the complicated psychology and neurobiology behind behavior (13, p. 89).

Moreover researchers expect that the identification of specific genes, and of the cellular, somatic and social variables that mediate between those genes and behavior, will increase the possibilities for humane and effective intervention, and put to rest the public misconception that genetic causes have immutable effects.

Critics, on the other hand, think that molecular genetics will help to sustain the false hopes that they believe have always informed behavioral genetics. They fear that neurogenetic research focused on individual differences is unlikely to yield much understanding of the causes of antisocial and violent behavior. They agree that the research is likely to find markers and genes that are loosely associated with that behavior. But they expect that most genetic contributions to differences in mental traits and behavior will be slight and oblique — all interaction and no main effects — difficult to interpret and of very limited theoretical interest. On the rare occasions when genetic contributions are significant, they are likely to be the work of mutations that cause major dysfunctions in small numbers of people. On this point, Evan Balaban contends that "the biochemical equivalent of hitting a subject on the head with a club may explain a pattern of pathology in a small number of individuals but will not be very enlightening for either the scientific study of behavioral biology or for general societal problems involving crime and violence" (14, p. 87).

Critics fear that the discovery of molecular markers for behavior will be less to advance scientific understanding than to increase social control. The danger of abuse is much greater for research employing sophisticated genomic technologies and enjoying the cachet of molecular genetics. Because the markers and genes will be easy to detect, and will have the appearance of hard scientific data, they are likely to be employed in programs of screening and preemptive intervention.

From XYY to MAO

These conflicting expectations about the scientific and social value of molecular genetic research are reflected in the divergent lessons that critics and researchers draw from the history of the first microbiological marker linked with human behavior, the XYY karyotype. In 1965 researchers found an apparently high incidence of that karyotype among prison inmates in Britain (15). They assumed that this incidence was higher than in the general population, an assumption since confirmed, and they speculated that men with an extra Y chromosome tended to be hyperaggressive, a hypothesis that subsequent research failed to support. It is now widely believed that if an extra Y chromosome leads to prison, it is by an indirect route. XYY individuals do not appear to be more aggressive or violent than average, but they may be taller, less intelligent, and more impulsive (16,17). Their increased risk of arrest may reflect a greater likelihood of getting caught, not heightened aggressiveness or a greater disregard of social norms.

Both researchers and critics regard the XYY story as a cautionary tale. But where critics see an illustration of the risks inherent in any inquiry into biological markers for social behavior, researchers see a modest triumph of scientific self-correction. Critics observe that the early XYY investigators, in their rush to find a direct link between genes and behavior, assumed that an extra "male" chromosome would make a specific contribution to violence or aggression, instead of having the generally impairing effects typically associated with an extra chromosome. Researchers, on the other hand, note that it was behavioral geneticists who ruled out any association between the XYY karyotype and violence or aggression (while confirming the high incidence of XYY in prisons and other institutions).

In the 25 years since the XYY controversy, the techniques for identifying biological markers for behavior have changed far more than the issues concerning their interpretation. With the development of recombinant DNA technology in the late 1970s, researchers were able to identify and manipulate individual genes and genetic material. That technology became directly relevant to behavioral genetics with the discovery of genetic "markers" for a variety of diseases and traits — highly variable ("polymorphic") but functionally inert DNA segments found to be associated with various phenotypic traits, presumably because they were located in close proximity to genes that actually contributed to those traits.

In the late 1980s and early 1990s, markers, and in some cases genes, were identified for a number of diseases known or suspected by their inheritance patterns to have a significant genetic component. Behavioral geneticists were quick to adopt the same methods, hoping to replicate the dramatic successes of medical genetics. They were soon reporting markers for a number of psychiatric and behavioral conditions, including bipolar disorder, schizophrenia, and alcoholism. Early findings in the first two areas had to be retracted, however, and findings in the third remain mired in controversy (5,18). A decade after the first application of molecular genetic techniques to psychiatric and behavioral disorders, there has yet to be a single consistently replicated, generally accepted association between a specific gene or marker and a common behavioral or psychiatric disorder. Although the yield of medical genetic research has also been more modest than its enthusiasts had predicted, the contrast between the two areas of research is still striking.

In 1993 researchers did find a marker, then a gene, associated with violence and aggression, in the male members of a Dutch family (19). This was an unexpected and somewhat awkward finding; it looked like the kind of "major 'crime gene'" that researchers had not been expecting to discover. While the family studied was

atypical in several relevant respects, the study had enormous impact, in part because the affected gene was known to produce a protein, MAO, involved in regulating the metabolism of serotonin, a neurotransmitter thought to play a critical role in mediating between genes and behavior: Serotonin has been implicated in psychiatric and behavioral conditions ranging from depression to impulsive violence.

A comparison of the MAO and XYY studies suggests both significant advances in scientific technique and persisting issues of interpretation. The connection between genotype and phenotype was closer in several respects for MAO than for XYY. First, a statistically significant association was found between MAO and aggressive and violent behavior in one family; in contrast, no association was established between XYY and any form of criminal behavior until a decade after the karyotype was identified (17). Second, there appeared to be better prospects for finding a specific causal pathway to antisocial behavior from a mutation in the MAO gene, which helps regulate the inhibitory mechanisms of the central nervous system, than from an extra copy of the entire chromosome linked with male gender. Over 50 studies have found an association between aggressive, violent, antisocial, or suicidal behavior and the serotonin metabolite, CSF 5-HIAA, that MAO helps to produce (20). Third, the measure of behavior was more direct in the MAO study than in the original XYY study — observation by the researchers or reports from close relatives, as opposed to inferences from official records.

And yet critics have argued that the link between MAO and violence and aggression, even in the one family studied, is much more tenuous than the researchers suggest, and that their work reveals some of the same inferential leaps that characterized early XYY research. In seeking a genetic cause for the high incidence of violent and aggressive behavior among male family members, the researchers may have paid insufficient attention to more global effects of the MAO mutation:

> Since a primary characteristic of the affected subjects was lowered IQ, it is unclear why the subjects' aggression received more emphasis than their cognitive deficits, and why there was no mention of the possibility that these cognitive deficits may have contributed to behavioral pathologies. . . .Perhaps these acts of violence are secondary to some more widespread defect in affect or cognition (20, p. 18).

Although some of the men in the family studied engaged in clearly violent and antisocial acts, that conduct, like the convictions on which the XYY researchers relied, may well have reflected a more general deficit. The mutant-MAO males, like the XYY males, may not have been more aggressive but less intelligent, lacking constructive outlets for their aggression or clever ways of concealing it.

The MAO researchers acknowledged that their findings did not support "a simple causal relationship between the metabolic abnormality and the behavioral disturbance" they observed, and that "borderline mental retardation" was also associated with the MAO mutation (21). Critics, however, complained that they did not fully investigate alternative explanations for the observed effects of MAO on behavior, nor adequately examined the range of behavior that might have been affected. Some critics (10) also question the claimed link between low serotonin and aggressive and antisocial behavior, which gives the MAO finding its threshold plausibility. These critics claim that the misconduct of serotonin-deficient individuals may reflect little more than the broadly debilitating impact of low serotonin levels on mental function, a breadth suggested by the sheer range of conditions in which serotonin deficits are implicated. They suggest that the fixation of serotonin researchers on specific psychiatric or behavioral pathologies has obscured such more general effects.

Researchers might respond by citing other evidence, or other studies, which tend to rule out more global explanations of the association between serotonin and violence, or which support a more direct causal pathway. Or they might regard this challenge merely as posing a legitimate question for further research. But they might also insist that these concerns, however reasonable, are hardly unique to behavioral genetics. If medical genetics has enjoyed greater success, it has also been beset by many of the same interpretive difficulties.

Thus medical geneticists also confront the problem of genetic pleiotropy — the multiple effects of a single gene and the multiplicity of possible causal pathways from gene to trait. The same kind of ambiguity arises in studies that search for the genes associated with various physiological effects, and in studies that probe the physiological effects of "candidate" genes (22). The association between a gene mutation and a disease revealed in a linkage or targeting study may arise indirectly, from more global effects of the gene on the organism or from the attempt of other genes or bodily systems to compensate for the loss of the gene's standard function. Because researchers will rarely be able to track the full range of functions and interactions a gene can have, they will rarely be warranted in claiming a direct causal relationship between that gene and a disease or other condition on the basis of a linkage or targeting study. Such claims will only be justified when researchers have acquired enough knowledge about developmental and physiological processes to narrow the range of plausible causal pathways. But that knowledge cannot come from molecular genetics alone.

If similar inferential difficulties confront medical and behavioral genetics, however, why should critics be so much more skeptical about the findings of the latter than of the former? One reason is that they may doubt the possibility of any causal generalizations concerning human behavior (a skepticism briefly discussed at the outset of this article). This blanket skepticism would give explanations based on cognitive limitations greater threshold plausibility than explanations based on behavioral predispositions, such as heightened aggression. A second reason is that critics may believe that genes contribute less, and less directly, to behavior than to physiology, and that the ways in which human behavior is defined and classified compound the difficulties of finding causal relationships. Because of the vast array of social and environmental forces shaping variations in human behavior and psychology, any genetic effect on

behavior or psychology will not only be modest but very difficult to track. And because many important behavioral and psychological categories are socially constructed or imposed, they may not be amenable to genetic or other biological explanation.

Thus researchers looking for genetic contributions to criminal behavior must confront the fact that such behavior is defined by legislators and "ascertained" by police, prosecutors, judges, and juries. The problem is not that genes cannot affect voluntary behavior — scientific critics readily concede that they can — but that social types may not correspond with biological types, so there may be no one type of behavior to be explained. We should not expect much in common psychologically or neurobiologically, let alone genetically, between a child-abuser, a pickpocket, a mob boss, and a political terrorist; between the bank robber prosecuted by one regime and the bank founder prosecuted by another. It is unlikely that any genetic feature distinguishes the members of such an eclectic rogues' gallery from the general population, and if there were, it would be unlikely to have much explanatory value.

Researchers acknowledge that the genetic contribution to behavior may well be more subtle and elusive than the genetic contribution to physiology, but they deny that there is less scientific value in finding smaller or less direct effects — Carey and Gottesman, for example, concede that "some of these loci may have little to do with an internal biology of antisocial behavior" as opposed to a general (in)sensitivity to environmental influences (13). Researchers also regard the heterogeneity and social construction of human behavior as part of the challenge of their work: either to discover unity in heterogeneity, through such underlying traits as impulsivity, antisocial personality, or novelty-seeking (the general tendency toward disobedience or risk) or to develop refined typologies of criminal behavior and look for different kinds of genetic influence on different types of crime (serial killing, leadership of an urban drug ring, embezzlement or tax evasion to finance a second home or third car) (23).

A final reason for the divergent expectations of researchers and critics is that the latter doubt the potential of genetic differences to explain behavioral differences. Behavioral geneticists using a wide variety of methods are united in the belief that small difference in genetic constitution can cause large differences in behavior in a fairly direct way. This belief is challenged by critics convinced that genes contribute little or nothing to behavior except its neuromuscular, sensory, and cognitive prerequisites. But it is also challenged by evolutionary biologists and psychologists who believe that genetic commonalities explain more than genetic differences, that differences in human behavior are generally explained by complex contingency plans encoded in genetic material almost all of us share. From this evolutionary perspective the genetic variations in important traits are likely to be preserved over time only if they are adaptive, that is, if they enhance reproductive fitness, thereby increasing their representation in subsequent generations. It is not clear, however, that genetic differences in consequential behavioral traits such as aggressiveness would be adaptive. Thus Allan Gibbard contrasts genetic differences affecting skin color from genetic differences affecting behavior:

> Survival and vitality are reproduction-enhancing, other things equal, and when there is much sun, dark pigment promotes survival and vitality, whereas with little sun, survival and vitality are promoted by having low levels of skin pigment. One could imagine that the same kind of pattern could hold for certain kinds of behavior.
>
> Might such effects be substantial, though? There are grounds for thinking not. The pattern requires genetic selection of different characteristics in different environments — as opposed to contingency plans for the same individual's having one set of characteristics if in one kind of environment and another set in another. This requires special conditions: the difference in which characteristic is more advantageous must last, on average, over many generations. Climate can make such stable differences, and account for such things as differences in skin color. With violence, relevant differences in life circumstances are likely to be much more volatile. When violence "pays" reproductively and when it doesn't will depend chiefly on characteristics of one's society and one's position in it. Whether social differences will be stable enough to produce different selection pressures is doubtful (24).

Gibbard, somewhat like Balaban, suspects that the genetic differences most likely to increase violence are maladaptive mutations that affect violence only indirectly (10).

Although these differences in expectation are in a broad sense empirical, they are unlikely to be resolved by the findings of one study, or even a series of studies. Behavioral geneticists and their critics are likely to interpret the findings of specific studies quite differently, and to have sharply divergent views about the comparative plausibility of alternative hypotheses.

There is also likely to be disagreement about whether the kind of genetic differences most likely to be identified could be said to predispose a person to a particular type of behavior. Can a genetic constitution that makes a person less intelligent be said to predispose him to crime or violence, if he is more likely to engage in crime or violence only because of more limited opportunities for education and employment? Can a person be said to be predisposed to violence if he acts violently only in extreme enviroments, such as a concentration camp, a space station, or a burnt-out urban neighborhood? Does it depend on how common such environments are? These are not narrowly scientific questions; they concern our understanding of causation and ascription of causal responsibility.

SOCIAL MEANING AND IMPACT

It is ironic that one of the strongest challenges to behavioral genetics comes from evolutionary biology and psychology. Those fields are only recently descended from sociobiology, a discipline that has been attacked by many critics of behavioral genetics for the same offense: that it "naturalizes" social injustice by attempting to ground it in biological reality. Although an evolutionary perspective may appear to rationalize patriarchy, infidelity, and other

objectionable social practices, in suggesting a genetically based predisposition to engage in them, it regards all people, or at least all males, as created equally predisposed. In contrast, behavioral genetics is dedicated to the proposition that some people are more predisposed genetically than others to destructive behavior. It is this controversial proposition that sets behavioral genetics apart from other disciplines committed to finding biological explanations for social phenomena. There has been more debate over the social implications of this proposition than over its scientific plausibility.

The debate has focused on the historical legacy of genetic-difference research and the social context in which it now takes place. Critics of behavioral genetics not only see more continuity between the old genetic-difference research and the new, they also see more continuity in their social settings; they do not believe that the "bad old days" are over. In light of the role that research on human genetic differences has played in justifying racism and inequality, they fear that even research focused solely on individual differences, and apolitical on its face, will be taken to justify racial stereotypes and used to justify coercive policies.

Behavioral geneticists see their research as having made a more complete break with the past socially as well as scientifically. They argue that contemporary society has a far greater capacity to use the results of their research humanely or at least to protect it from abuses. They see in their work the promise of more effective and less coercive solutions to recalcitrant social problems, and deny that it is a vehicle for a repressive political agenda or the latest incarnation of scientific racism.

We can conveniently divide the conflicts over the social context and impact of the research into issues concerning the past and the present.

The Eugenic Legacy

Behavioral geneticists tend to identify the cardinal sins of eugenics as coercion and group discrimination. Programs of forced sterilization were objectionable primarily because they were *forced*, employing physical compulsion, threats, or deception to induce people to limit their fertility. They were further objectionable because they relied on crude racial, ethnic, and class generalizations to decide whose fertility to limit. Most contemporary researchers emphatically reject both features of the old eugenics. They intend their findings to be used only by individuals for their own benefit or the benefit of their families, they see the added information they provide as enhancing rather than limiting reproductive choice, and they eschew generalizations for (nonfamily) groups. That is, they do not investigate, and do not expect to find, racial or ethic differences in significant behavioral and mental traits. Most researchers also oppose restrictions on reproductive freedom and support restrictions on third-party access to, or use of, behavioral genetic information.

Critics respond to these avowals in two ways. First, they think that the specter of coercion and discrimination is still present, albeit in subtler form. They fear that eugenic values have become internalized in the decisions made by individual reproductive agents; that our society is embracing a "backdoor" eugenics more insidious, if less oppressive, than the old front-door variety (25). The social and economic pressures to prevent the birth of "defective" fetuses are very strong, and they are felt across the whole population, not merely by the marginal groups once targeted for "improvement" or elimination. And while behavioral geneticists, with a few notorious exceptions, are not interested in investigating genetic differences among racial, ethnic, or other social groups, they are widely perceived to be, and their research methods can be easily adapted to that purpose.

Critics also point out that coercion and discrimination were not the only objectionable features of the old eugenics. Equally central was the belief that social problems such as poverty and crime persist in part because of the failure to design social institutions to take account of constitutional differences among individuals. However emphatically current genetic-difference researchers reject the racism or the political conservatism of their predecessors, they share that underlying conviction. Critics argue that that conviction also lies behind the public enthusiasm for the resurgence of behavioral genetics (26).

The Current Climate

Critics of behavioral genetics attribute its renewed popularity to the perceived failure of the movements of social reform that began with the New Deal and ended with the Great Society. Having made desultory efforts to improve conditions for the worst-off members of society and witnessed few dramatic successes, our society has been all too eager to return to policies that blame "defective" individuals and groups for the social problems that persist. Behavioral genetics provides a scientific pretext for this retrenchment.

For contemporary behavioral geneticists, the apparent failure of many of the social reforms of the post–World War II reflects scientific ignorance more than excessive optimism. The failure of public schools to educate or prisons to reform does not mean that students or immates were stupid or incorrigible, but that policy makers did not know how to teach or rehabilitate them. They failed to recognize individual differences in cognitive and behavioral disposition and to tailor their interventions accordingly. In offering one-size-fits-all programs, they slighted the needs of those with the greatest difficulties in learning or conforming. Behavioral genetics suggests ways of customizing interventions that actually work. Any stigma associated with the identification of individuals requiring special intervention is far milder than the stigma of failure or recalcitrance that those individuals would otherwise bear.

Deficient Serotonin and Heightened Aggression

Long before the advent of molecular genetics, it was widely suspected that the most prolific and recalcitrant offenders differed genetically from the rest of us. That suspicion was memorably captured in Maxwell Anderson's play *Bad Seed*, written in the 1950s, a time when the crime rate was relatively low and public confidence in environmental interventions relatively high. Almost fifty years later

those suspicions are far stronger. Although the unprecedented rise in violent crime after 1960 is often attributed to broad social conditions, from the decline of traditional families to the increase in economic and social inequality, there has been a growing conviction among researchers and policy makers that a large share of the increase is due to a small number of individuals, whose identification and incapacitation will go a long way toward reducing violent crime (27). While the primary means of identifying such individuals has been by their recorded deeds — arrests and convictions — some researchers expect to find chemical and genetic "indicators" that would permit early identification and preemptive intervention. This prospect places the hopes and fears about the social impact of the new behavioral genetics into sharp focus.

Researchers see the promise of more humane and effective treatment: "prenatal and postnatal care, rehabilitation for pregnant drug abusers, educational enrichment programs, parenting education, conflict resolution tactics, media restrictions for violent programming, and gun control" (28). Critics see the threat of further racial polarization and social control. They argue that it would be reckless to develop biological techniques for predicting and "treating" juvenile delinquency or adult criminality, when those techniques would be placed in the hands of educational and criminal justice institutions that have shown so little capacity to serve their clients humanely. They imagine genetic testing and medical treatment being made a condition of probation or parole, or "offered" to disruptive or unreliable employees as an alternative to termination. Other pharmacological treatments, from anabuse to lithium, have been introduced in this way, setting a strong and ominously seductive precedent.

Individual Differences and Racial Generalizations

Researchers and critics disagree about the *racial implications* of individual-differences research. Researchers insist that their work is free of the racial taint of previous research, since it does not even look for group differences and studies mainly white population subgroups with large families or good archives, from Finns to Mormons. Critics, however, argue that the distinction between individual and group differences will prove to be untenable because claims of individual differences in genetic predisposition will almost inevitably lend support to claims of group differences. There are several ways in which this might happen.

First, if some researchers find genes or markers associated with criminal behavior in individuals, other researchers may try to compare the incidence of those genes or markers in racial and ethnic groups. (The data for such comparisons will be found in studies on individual predispositions if those studies involve multiracial populations and code subjects by race.) There are good reasons to doubt that researchers would find significant group differences, or that any differences they found would correspond to differences in present offense rates. Still the discovery of individual differences in genetic predisposition could have an adverse impact on minority groups even if, as seems likely, those groups are *not* found to have a higher incidence of the relevant genes or markers. There will be considerable pressure to use those genes or markers for detecting criminal tendencies in young children and assessing dangerousness in convicted offenders. Universal screening is very unlikely, but the selective screening of those who look "vulnerable" by dint of misbehavior is well within the realm of social possibility. Those screened are likely to be drawn disproportionately from the predominantly black and Hispanic inner cities, since these are the areas in which violent crime is concentrated (29).

Even without population research or screening programs, evidence of differences in genetic predisposition within groups will be widely taken as evidence of differences between groups, however unwarranted that inference may be. The discovery of genes or markers associated with criminal behavior may be publicly perceived as implicating the black community, at a time when an alarmingly high proportion of African-American males are involved in the criminal justice system (30).

Whether or not the critics' apprehensions are more realistic than the researchers' hopes, these different expectations about the social impact of research on genetics and crime reflect broader differences in outlook, about the fairness and reliability of our institutions of social control, and about the pervasiveness and recalcitrance of racial bias in contemporary society. Defenders and critics of genetic research on crime may differ less in their political values than in their levels of trust and optimism.

A Marker for Homosexuality?

The same year the Dutch study on MAO and aggression was published, researchers claimed to find a marker associated with homosexual orientation (31). Although there has been an unusual degree of controversy over the methods used to ascertain the phenotype of sexual orientation, as well as over the statistical design that yielded the linkage, what was most striking about the study was the enthusiasm with which it was received by some members of the gay community: It was the first time that a large segment of a minority community had warmly embraced a claim of genetic predisposition (32).

For the researchers who located the marker, the discovery was important socially as well as scientifically. It showed that behavioral genetics could be liberating rather than oppressive, legitimizing socially deviant conditions by revealing that they had a biological basis (33). This optimistic view was shared by some members of the gay community, who saw the research as helping to place sexual orientation in the constitutionally protected category of "immutable characteristics," thereby giving homosexuals a legal defense against many forms of discrimination to which they were currently subject.

For critics of behavioral genetics, however, as well as for other representatives of the gay community, the research was simplistic empirically and naive politically. It distorted the complexity and plasticity of sexual attraction and activity, transforming shifting "preferences" into deep, immutable "orientations," dichotomizing a multidimensional array of preferences and behaviors into hetero and homosexuality, and falsely associating effeminacy with

homosexuality (34). Moreover critics regard genetic vindication as a Faustian bargain. To embrace the claim that sexual orientation is genetically influenced is to treat it as something that needs to be excused. At the same time genetic predisposition would furnish at most an incomplete excuse: No researcher would claim that sexual orientation, let alone sexual conduct, is fully determined genetically. And even if sexual orientation were seen as genetically determined, it would be unlikely to furnish an excuse the homophobic public would accept — someone born gay might, if anything, be seen as more deeply flawed than someone made gay by a domineering mother — the most popular "scientific" explanation a generation ago. As Diane Paul argues:

> [T]here is little reason to think that social policy would change even if it were generally agreed that homosexuals were in no way blameworthy. Those who consider genetic (or other biological) explanations necessarily liberating are, in my view, naive. Research results acquire social meaning only in the context of other assumptions, for example, how repugnant and threatening others find the behavior. . . .The history of eugenics provides a warning to those with a socially progressive agenda: to stress that our fate is in our genes is a very risky strategy (34, p. 99).

Finally, the mitigating effects of a genetic excuse might well be undermined or offset by the use of prenatal testing for suspected markers or genes to prevent the birth of gay children. While it is unlikely that such testing will ever become accurate enough to substantially reduce the incidence of homosexuality, many parents would seek testing as a way to avoid the frustration and embarrassment of having a gay child (35). The mere availability of testing would contribute to the perception of homosexuality as a "preventable" disorder, like Down syndrome or Tay Sachs disease. Life would be especially difficult for the "false negatives" — those children who tested straight but became gay — who would risk condemnation for embracing a deviant orientation against their genetic grain. Matters might be even worse if genetically based treatments for homosexuality became available: Those who declined treatment would be seen as recalcitrant as well as deviant, ratifying the sexual orientation that they had been acquitted of choosing in the first place. What genetic diagnosis would give by way of excuse, genetic therapy would take away.

Researchers might respond that a deeper understanding of the biological roots of sexual orientation would increase tolerance by making differences in orientation appear not only involuntary but natural. Instead of being seen as a disease or disorder, homosexuality might be seen as a normal variation, such as red hair. They could cite historical precedent: The very first scientific attempts to explain homosexuality in biological terms were made by researchers who regarded all combinations of male and female attributes as natural and sought to decriminalize same-sex sexual behavior. Critics, however, would see this precedent as double-edged. The work that sought to normalize homosexuality was soon built upon by researchers seeking to pathologize it; the hypothesis that homosexuality had a biological or genetic basis became the cornerstone of eugenic policies against homosexuals (31).

The researchers might respond that the fact that biological research had been misappropriated in this way hardly meant that it would inevitably be turned to eugenic or repressive purposes. They might agree with Diane Paul that the way genetic explanations are received "depends on how repugnant and threatening others find the behavior," but insist that public attitudes toward homosexuality in Western societies have become increasingly accepting, and that genetic explanations will only accelerate the trend (34). As in the case of criminal behavior, divergent expectations about the social impact of genetic research into sexual orientation appear to arise from broader differences in outlook about the social maturity and political fairness of the societies in which the research takes place.

CONCLUSION

It would be naive to expect a single finding, or series of findings, to resolve doubts about the scientific value of human behavioral genetics, or to expect a single application of the research, or series of applications, to resolve doubts about its social utility. At the same time it would be dogmatic to insist that no conceivable findings or applications could vindicate the researchers' confidence and optimism, or their critics' doubts and pessimism. It is, however, quite likely that an article on human behavioral genetics written for the next edition of this Encyclopedia will reflect some of the same uncertainty and ambivalence as this one.

BIBLIOGRAPHY

1. D.J. Kevles, *In the Name of Eugenics: Genetics and the Uses of Human Heredity*, Knopf, New York, 1985.
2. D. Paul, *Controlling Human Heredity: 1865 to the Present*, Humanities Press, Atlantic Highlands, NJ, 1995.
3. D. Goldman, *Politics Life Sci.* **15**(1), 97–98 (1996).
4. R. Plomin et al., *Behavioral Genetics*, 3rd ed., W.F. Freeman, New York, 1997.
5. S.L. Sherman et al., *Am. J. Hum. Genet.* **60**, 1265–1275 (1997).
6. R. Plomin, M.J. Owen, and P. McGuffin, *Science* **264**, 1733–1739 (1994).
7. G. Carey, in A.J. Reiss, Jr. and J.A. Roth, eds., *Understanding and Preventing Violence: Biobehavioral Influences on Violence*, vol. 2, National Academy Press, Washington, DC, 1994.
8. D.C. Rowe and K.C. Jacobson, in R.A. Carson and M.A. Rothstein, eds., *Behavioral Genetics: The Clash of Biology and Culture*, Johns Hopkins University Press, Baltimore, MD, 1999, pp. 12–34.
9. E. Sober, in D. Wasserman and R. Wachbroit, eds., *Genetics and Criminal Behavior: Methods, Meanings, and Morals*, Cambridge University Press, New York, forthcoming.
10. E. Balaban, in Singh et al., eds., *Thinking about Evolution: Historical, Philosophical, and Political Perspectives*, Cambridge University Press, Cambridge, UK, forthcoming.
11. C. Jencks, *New York Rev. Books*, February 12, pp. 33–41, 34 (1987).

12. L.J. Eaves, N.G. Martin, and A.C. Heath, *Beh. Genet.* **20**, 1–22 (1990).
13. G. Carey and I. Gottesman, *Politics Life Sci.* **15**(1), 88–90 (1996).
14. E. Balaban, *Politics Life Sci.* **15**(1), 86–88 (1996).
15. P. Jacobs et al., *Nature* **208**, 1351–1352 (1965).
16. H.A. Witkin et al., *Science* **196**, 547–555 (1976).
17. M. Rutter, H. Giller, and A. Hagell, *Antisocial Behavior by Young People*, Cambridge University Press, New York, 1998.
18. I.D. Waldman, in R.A. Carson and M.A. Rothstein, eds., *Behavioral Genetics: The Clash of Biology and Culture*, Johns Hopkins University Press, Baltimore, MD, 1999, pp. 35–59.
19. H.G. Brunner et al., *Science* **262**, 578–580 (1993).
20. E. Balaban, J. Alper, and Y.L. Kasamon, *J. Neurogenet* **11**, 1–43 (1996).
21. H.G. Brunner, in G.R. Bock and J.A. Goode, eds., *Genetics of Criminal and Antisocial Behavior*, Wiley, New York, 1996, pp. 156, 159.
22. S. Culp, *Philos. Sci.* (*Proceedings*) **64**, S268–S278 (1997).
23. I. Gottesman and H.H. Goldsmith, in C.A. Nelson, ed., *Threats to Optimal Development: Integrating Biological, Psychological, and Social Risk Factors*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1994.
24. A. Gibbard, in D. Wasserman and R. Wachbroit, eds., *Genetics and Criminal Behavior: Methods, Meanings, and Morals*, Cambridge University Press, New York, forthcoming.
25. T. Duster, *Backdoor to Eugenics*, Routledge, New York, 1990.
26. R. Lewontin, S. Rose, and L. Kamin, *Not in Our Genes: Biology, Ideology, and Human Nature*, Pantheon, New York, 1984.
27. J.Q. Wilson and R. Herrnstein, *Crime and Human Nature*, Simon & Schuster, New York, 1985.
28. D. Fishbein, *Politics Life Sci.* **15**(1), 91–94, 93 (1996).
29. D. Wasserman, in Robert Post, ed., *Censorship and Silencing: Practices of Cultural Regulation*, Getty Research Institute, Santa Monica, CA, 1998, pp. 169–193.
30. J. Miller, *Search and Destroy: African-American Males in the Criminal Justice System*, Cambridge University Press, New York, 1996.
31. D. Hamer et al., *Science* **261**, 321–327 (1993).
32. D. Hamer, *The Search for the Gay Gene and the Biology of Behavior*, Simon & Schuster, New York, 1994.
33. J. DeCecco and M. Shively, eds., *Sex, Cells, and Same-Sex Desire: The Biology of Sexual Preference*, Haworth, New York, 1995.
34. D. Paul, in *Politics Life Sci.* **15**(1), 99–100 (1996).
35. E. Stein, *Bioethics* **12**(1), 1–48 (1998).

See other entries GENETIC DETERMINISM, GENETIC REDUCTIONISM, AND GENETIC ESSENTIALISM; see also HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY entries; PUBLIC PERCEPTIONS: SURVEYS OF ATTITUDES TOWARD BIOTECHNOLOGY.

C

CLONING, ETHICS

DAN W. BROCK
Brown University
Providence, Rhode Island

OUTLINE

INTRODUCTION

One of the most dramatic recent advances in biotechnology was the successful cloning of a sheep from a single cell of an adult sheep. The world of science and the public at large were both shocked and fascinated by this accomplishment of Ian Wilmut and his colleagues (1). Scientists were in part surprised because many had believed that after the very early stage of embryo development at which differentiation of cell function begins to take place it would not be possible to achieve the cloning of an adult mammal by nuclear transfer. In this process the nucleus from the cell of an adult mammal is inserted into an ennucleated ovum, and the resulting embryo develops following the complete genetic code of the mammal from which the inserted nucleus was obtained. But some scientists and much of the public were troubled or apparently even horrified at the prospect that if adult mammals such as sheep could be cloned, then cloning of adult humans by the same process would likely be possible as well. Of course, the process is far from perfected even with sheep — it took 276 failures by Wilmut and his colleagues to produce Dolly, their one success, and whether the process can be successfully replicated in other mammals, much less in humans, is not now known. But those who were horrified at the prospect of human cloning were not assuaged by the fact that the science with humans is not yet there, for it looked to them now perilously close.

The response of most scientific and political leaders to the prospect of human cloning, indeed of Dr. Wilmut as well in testimony before Congress in March 1997, was immediate and strong condemnation. In the United States President Clinton immediately banned federal financing of human cloning research and asked privately funded scientists to halt such work until the newly formed National Bioethics Advisory Commission could review the "troubling" ethical and legal implications. The Director-General of the World Health Organization characterized human cloning as "ethically unacceptable as it would violate some of the basic principles which govern medically assisted reproduction. These include respect for the dignity of the human being and the protection of the security of human genetic material" (2). Around the world similar immediate condemnation was heard as human cloning was called a violation of human rights and human dignity. Even before Wilmut's announcement, human cloning had been made illegal in nearly all countries in Europe and had been condemned by the Council of Europe (3).

A few more cautious voices were heard both suggesting some possible benefits from the use of human cloning in limited circumstances and questioning its too quick prohibition, but they were a clear minority. In the popular media, nightmare scenarios of laboratory mistakes resulting in monsters, the cloning of armies of Hitlers, the exploitative use of cloning for totalitarian ends as in Huxley's *Brave New World*, and the murderous replicas of the film Blade Runner all fed the public controversy and uneasiness. A striking feature of these early responses was that their strength and intensity seemed far to outrun the arguments and reasons offered in support of them — they seemed often to be "gut level" emotional reactions rather than considered reflections on the issues. Such reactions should not be simply dismissed, both because they may point to important considerations otherwise missed, and not easily articulated, and because they often have a major impact on public policy. But the formation of public policy should not ignore the ethical reasons and arguments that bear on the practice of human cloning — these must be articulated in order to understand and inform people's more immediate emotional responses. This article will evaluate critically the main moral considerations and arguments for and against human cloning. Though many people's religious beliefs inform their views on human cloning, and it is often difficult to separate religious from secular positions, this article is restricted to arguments and reasons that can be given a clear secular formulation and does not take up explicitly religious positions and arguments pro or con. This article is also concerned principally with cloning by nuclear transfer, which permits cloning of an adult, not cloning by embryo splitting, although some of the issues apply to both (4).

I begin by noting that on each side of the issue there are two distinct kinds of moral arguments brought forward. On the one hand, some opponents claim that human cloning would violate fundamental moral or human rights, while some proponents argue that its prohibition would violate such rights. On the other hand, both opponents and proponents also cite the likely harms and benefits, both to individuals and to society, of the practice. While moral and even human rights need not be understood as

absolute, that is, as morally requiring people to respect them no matter how great the costs or bad consequences of doing so, actions that would violate them cannot be justified by a mere balance of benefits over harms. For example, the rights of human subjects in research must be respected even if the result is that some potentially beneficial research is more difficult or cannot be done, and the right of free expression prohibits the silencing of unpopular or even abhorrent views; in Ronald Dworkin's striking formulation, rights trump utility (5). This article will take up the moral rights implicated in human cloning, as well as its more likely significant benefits and harms, because none of the rights as applied to human cloning is sufficiently uncontroversial and strong to settle decisively the morality of the practice one way or the other. But because of their strong moral force, the assessment of the moral rights putatively at stake is especially important. A further complexity here is that it is sometimes controversial whether a particular consideration is merely a matter of benefits and harms, or is instead a matter of moral or human rights. We begin with the arguments in support of permitting human cloning, although with no implication that it is the stronger or weaker position.

MORAL ARGUMENTS IN SUPPORT OF HUMAN CLONING

Is There a Moral Right to Use Human Cloning?

What moral right might protect at least some access to the use of human cloning? Some commentators have argued that a commitment to individual liberty, as defended by J.S. Mill, requires that individuals be left free to use human cloning if they so choose and if their doing so does not cause significant harms to others, but liberty is too broad in scope to be an uncontroversial moral right (6,7). Human cloning is a means of reproduction (in the most literal sense), so the most plausible moral right at stake in its use is a right to reproductive freedom or procreative liberty (8,9). Reproductive freedom includes not only the familiar right to choose not to reproduce, for example, by means of contraception or abortion, but also the right to reproduce. The right to reproductive freedom is properly understood to include as well the use of various assisted reproductive technologies, such as in vitro fertilization (IVF), and oocyte donation. The reproductive right relevant to human cloning is a negative right, that is, a right to use assisted reproductive technologies without interference by the government or others when made available by a willing provider. The choice of an assisted means of reproduction, such as surrogacy, can be defended as included within reproductive freedom even when it is not the only means for individuals to reproduce, just as the choice among different means of preventing conception is protected by reproductive freedom. However, the case for permitting the use of a particular means of reproduction is strongest when that means is necessary for particular individuals to be able to procreate at all. Sometimes human cloning could be the only means for individuals to procreate while retaining a biological tie to the child created, but in other cases different means of procreating would also be possible.

It could be argued that human cloning is not covered by the right to reproductive freedom because, whereas current assisted reproductive technologies and practices covered by that right are remedies for inabilities to reproduce sexually, human cloning is an entirely new means of reproduction; indeed, its critics see it as more a means of manufacturing humans than of reproduction. Human cloning is a different means of reproduction than sexual reproduction, but it is a means that can serve individuals' interest in reproducing. It cannot be excluded from the moral right to reproductive freedom merely because it is a new means of reproducing, but rather only if it has other objectionable moral features, such as eroding human dignity or uniqueness; we will evaluate these other ethical objections to it below.

When individuals have alternative means of procreating, human cloning typically would be chosen because it replicates a particular individual's genome. The reproductive interest in question then is not simply reproduction itself, but a more specific interest in choosing what kind of children to have. The right to reproductive freedom is usually understood to cover at least some choice about the kind of children one will have; for example, genetic testing of an embryo or fetus for genetic disease or abnormality, together with abortion of an affected embryo or fetus, is now used to avoid having a child with that disease or abnormality. Genetic testing of prospective parents before conception to determine the risk of transmitting a genetic disease is also intended to avoid having children with particular diseases. Prospective parents' moral interest in self-determination, which is one of the grounds of a moral right to reproductive freedom, includes the choice about whether to have a child with a condition that is likely to place severe burdens on them, and to cause severe burdens to the child itself.

The more a reproductive choice is not simply the determination of oneself and one's own life but the determination of the nature of another, as in the case of human cloning, the more moral weight the interests of that other person, that is the cloned child, should have in decisions that determine its nature (10). But even then parents are typically taken properly to have substantial, but not unlimited, discretion in shaping the persons their children will become, for example, through education and other child-raising decisions. Even if not part of reproductive freedom, the right to raise one's children as one sees fit, within limits mostly determined by the interests of the children, is also a right to determine within limits what kinds of persons one's children will become. This right includes not just preventing certain diseases or harms to children, but selecting and shaping desirable features and traits in one's children. The use of human cloning is one way to exercise that right.

Its worth pointing out that current public and legal policy permits prospective parents to conceive, or to carry a conception to term, when there is a significant risk, or even certainty, that the child will suffer from a serious genetic disease. Even when others think the risk or presence of genetic disease makes it morally wrong to conceive, or to carry a fetus to term, the parents' right to reproductive freedom permits them to do so. Most possible harms to a

cloned child to be considered below are less serious than the genetic harms with which parents can now permit their offspring to be conceived or born.

To conclude our discussion of a moral right to use human cloning, there is good reason to accept that a right to reproductive freedom presumptively includes both a right to select the means of reproduction, as well as a right to determine what kind of children to have, by use of human cloning. However, the particular reproductive interest of determining what kind of children to have is less weighty than other reproductive interests and choices whose impact falls more directly and exclusively on the parents rather than the child. Accepting a moral right to reproductive freedom that includes the use of human cloning does not settle the moral issue about human cloning, however, since there may be other moral rights in conflict with this right, or serious enough harms from human cloning to override the right to use it; this right can be thought of as establishing a serious moral presumption supporting access to human cloning.

There is a different moral right that might be thought to be at stake in the dispute about human cloning — the right to freedom of scientific inquiry and research in the acquisition of knowledge. If there is such a right, it would presumably be violated by a legal prohibition of research on human cloning, although the government could still permissibly decide not to spend public funds to support such research. Leaving aside for the moment human subject ethical concerns, research on human cloning might provide valuable scientific or medical knowledge beyond simply knowledge about how to carry out human cloning. Whether or not there is a moral right to freedom of scientific inquiry, for example, as part of a right to free expression, prohibiting and stopping scientific research and inquiry is a serious matter and precedent that should only be undertaken when necessary to prevent grave violations of human rights or to protect fundamental human interests. But even for opponents of human cloning the fundamental moral issue is not acquiring the knowledge that would make it possible, but using that knowledge to do human cloning. Since it is possible to prohibit human cloning itself, without prohibiting all research on it, it is not necessary to limit the freedom of scientific inquiry in order to prevent human cloning from taking place. But this means as well that a right to freedom of scientific inquiry could only protect research on human cloning, not the use of human cloning. For this reason the fundamental moral right which provides presumptive moral support for permitting the use of human cloning is the right to reproductive freedom, not the right to freedom of scientific inquiry. In what follows, the discussion will principally concern the moral issues in the use of human cloning, not those restricted to research on it.

What Individual or Social Benefits Might Human Cloning Produce?

Largely Individual Benefits. The literature on human cloning by nuclear transfer, as well as the literature on embryo splitting where it is relevant to the nuclear transfer case, contain a few examples of circumstances in which individuals might have good reasons to want to use human cloning. However, human cloning does not appear to be the unique answer to any great or pressing human need or social problem, and its benefits would likely be at most limited. What are the principal benefits of human cloning that might give persons good reasons to want to use it?

Human Cloning Would be a New Means to Relieve the Infertility Some Persons Now Experience. Human cloning would allow women who have no ova or men who have no sperm to produce an offspring that is biologically related to them (8,11–13). Embryos might also be cloned, either by nuclear transfer or embryo splitting, in order to increase the number of embryos for implantation and improve the chances of successful conception (14). While the moral right to reproductive freedom creates a presumption that individuals should be free to choose the means of reproduction that best serves their interests and desires, the benefits from human cloning to relieve infertility are greater the more persons there are who cannot overcome their infertility by any other means acceptable to them; data are not now available about the numbers of persons who could only relieve their infertility by human cloning.

It is not enough to point to the large number of children throughout the world possibly available for adoption as a solution to infertility, unless we are prepared to discount as illegitimate the strong desire many persons, fertile and infertile, have for the experience of pregnancy and for having and raising a child biologically related to them. While not important to all infertile (or fertile) individuals, it is important to many and is respected and met through other forms of assisted reproduction that maintain a biological connection when that is possible; there seems no good reason to refuse to respect and respond to it when human cloning would be the best or only means of overcoming individuals' infertility.

Human Cloning Would Enable Couples in Which One Party Risks Transmitting a Serious Hereditary Disease, a Serious Risk of Disease, or an Otherwise Harmful Condition to an Offspring, to Reproduce Without Doing So (8). Of course, by using donor sperm or egg donation, such hereditary risks can generally be avoided now without the use of human cloning. These procedures may be unacceptable to some couples, however, or at least considered less desirable than human cloning because they introduce a third party's genes into their reproduction, instead of giving their offspring only the genes of one of them. Thus in some cases human cloning would be a means of preventing genetically transmitted harms to offspring. Here too there are no data on the likely number of persons who would wish to use human cloning for this purpose instead of either using other available means of avoiding the risk of genetic transmission of the harmful condition or accepting the risk of transmitting the harmful condition.

Human Cloning a Later Twin Would Enable a Person to Obtain Needed Organs or Tissues for Transplantation (12,15,16). Human cloning would solve the problem of finding a transplant donor who is an acceptable organ or tissue match and would eliminate, or drastically reduce, the risk of transplant rejection by the host. The availability of human cloning for this purpose would amount to a form of insurance policy to enable treatment of certain kinds of

medical needs. Of course, often the medical need would be too urgent to permit waiting for the cloning, gestation and development of the later twin necessary before tissues or organs for transplant could be obtained. In other cases the need for an organ that the later twin would him or herself need to maintain life, such as a heart or a liver, would preclude cloning and then taking the organ from the later twin.

Such a practice has been criticized on the ground that it treats the later twin not as a person valued and loved for his or her own sake, as an end in itself in Kantian terms, but simply as a means for benefiting another. This criticism assumes, however, that only this one motive would determine the relation of the person to his or her later twin. The well-know case some years ago in California of the Ayala's, who conceived in the hopes of obtaining a source for a bone marrow transplant for their teenage daughter suffering from leukemia illustrates the mistake in this assumption. They argued that whether or not the child they conceived turned out to be a possible donor for their daughter, they would value and love the child for itself, and treat it as they would treat any other member of their family. That one reason it was wanted was as a means to saving their daughter's life did not preclude its also being loved and valued for its own sake; in Kantian terms, it was treated as a possible means to saving their daughter, but not *solely as a means*, which is what the Kantian view proscribes.

Indeed, when people have children, whether by sexual means or with the aid of assisted reproductive technologies, their motives and reasons for doing so are typically many and complex, and include reasons less laudable than obtaining life-saving medical treatment, such as having a companion or someone who needs them, enabling them to live on their own, and qualifying for public or government benefit programs. While these other motives for having children sometimes may not bode well for the child's upbringing and future, public policy does not assess prospective parents motives and reasons for procreating as a condition of their doing so.

One commentator has proposed human cloning for obtaining even life-saving organs (15). After cell differentiation some of the brain cells of the embryo or fetus would be removed so that it could then be grown as a brain dead body for spare parts for its earlier twin. This body clone would be like an anencephalic newborn or presentient fetus, neither of whom arguably can be harmed because of their lack of capacity for consciousness. Most people would likely find this practice appalling and immoral, in part because here the cloned later twin's capacity for conscious life is destroyed *solely as a means* for the benefit of another. Yet, if one pushes what is already science fiction quite a bit further in the direction of science fantasy, and imagines the ability to clone and grow in an artificial environment only the particular life-saving organ a person needed for transplantation, then it is far from clear that it would be morally impermissible to do so.

Human Cloning Would Enable Individuals to Clone Someone Who Had Special Meaning to Them, Such as a Child Who Had Died (12). There is no denying that if human cloning were available, some individuals would want to use it in order to clone someone who had special meaning to them, such as a child who had died, but that desire usually would be based on a deep confusion. Cloning such a child would not replace the child the parents had loved and lost, but rather would create a new different child, though with the same genes. The child they loved and lost was a unique individual who had been shaped by his or her environment and choices, not just his or her genes, and, more important, who had experienced a particular relationship with them. Even if the later cloned child could have not only the same genes but also be subjected to the same environment, which is in fact impossible, it would remain a different child than the one they had loved and lost because it would share a different history with them (17). Cloning the lost child might help the parents accept and move on from their loss, but another already existing sibling or another new child that was not a clone might do this equally well; indeed, it might do so better since the appearance of the cloned later twin would be a constant reminder of the child they had lost. Nevertheless, if human cloning enabled some individuals to clone a person who had special meaning to them and doing so had deep meaning and satisfaction for them, that would be a benefit to them even if their reasons for wanting to do so, and the satisfaction they in turn received, were based on a confusion.

Largely Social Benefits

Human Cloning Would Enable the Duplication of Individuals of Great Talent, Genius, Character, or Other Exemplary Qualities. The first four reasons for human cloning considered above all looked to benefits to specific individuals, usually parents, from being able to reproduce by means of human cloning. This reason looks to benefits to the broader society from being able to replicate extraordinary individuals — a Mozart, Einstein, Gandhi, or Schweitzer (18,19). Much of its appeal, like much thinking both in support of and in opposition to human cloning, rests on a confused and mistaken assumption of genetic determinism, that is, that one's genes fully determine what one will become, do, and accomplish. What made Mozart, Einstein, Gandhi, and Schweitzer the extraordinary individuals they were was the confluence of their particular genetic endowments with the environments in which they were raised and lived and the particular historical moments they in different ways seized. Cloning them would produce individuals with the same genetic inheritances (nuclear transfer does not even produce 100 percent genetic identity, although for the sake of exploring the moral issues we follow here the common assumption that it does), but neither by cloning, nor by any other means, would it be possible to replicate their environments or the historical contexts in which they lived and their greatness flourished. We do not know, either in general or with any particular individual, the degree or specific respects in which their greatness depended on their "nature" or their "nurture," but we do know in all cases that it depended on an interaction of them both. Thus human cloning could never replicate the extraordinary accomplishments for which we admire individuals like Mozart, Einstein, Gandhi, and Schweitzer.

If we make a rough distinction between the extraordinary capabilities of a Mozart or an Einstein and how they used those capabilities in the particular environments and historical settings in which they lived, it would also be a mistake to assume that human cloning could at least replicate their extraordinary capabilities, if not the accomplishments they achieved with them. Their capabilities too were the product of their inherited genes and their environments, not of their genes alone, so it would be a mistake to think that cloning them would produce individuals with the same capabilities, even if they would exercise those capabilities at different times and in different ways. In the case of Gandhi and Schweitzer, whose extraordinary greatness lies more in their moral character and commitments, we understand even less well the extent to which their moral character and greatness was produced by their genes.

None of this is to deny that Mozart's and Einstein's extraordinary musical and intellectual capabilities, nor even Gandhi's and Schweitzer's extraordinary moral greatness, were produced in part by their unique genetic inheritances. Cloning them might well produce individuals with exceptional capacities, but we simply do not know how close their clones would be in capacities or accomplishments to the great individuals from whom they were cloned. Even so, the hope for exceptional, even if less and different, accomplishment from cloning such extraordinary individuals might be a reasonable ground for doing so.

The examples above are of individuals whose greatness is widely appreciated and largely uncontroversial, but if we move away from such cases we encounter the problem of whose standards of greatness would be used to select individuals to be cloned for the benefit of society or humankind in general. This problem inevitably connects with the important issue of who would control access to and use of the technology of human cloning, since those who controlled its use would be in a position to impose their standards of exceptional individuals to be cloned. This issue is especially worrisome if particular groups or segments of society, or if government, controlled the technology for we would then risk its use for the benefit of those groups, segments of society, or governments under the cover of benefiting society or humankind.

Human Cloning and Research on Human Cloning Might Make Possible Important Advances in Scientific Knowledge, for Example About Human Development (20,21). While important potential advances in scientific or medical knowledge from human cloning or human cloning research have frequently been cited in some media responses to Dolly's cloning, there are at least three reasons why these possible benefits are highly uncertain. First, there is always considerable uncertainty about the nature and importance of the new scientific or medical knowledge that a dramatic new technology like human cloning will lead to; the road to that new knowledge is never mapped in advance and takes many unexpected turns. Second, we do not know what new knowledge from human cloning or human cloning research could also be gained by other methods and research that do not have the problematic moral features of human cloning to which its opponents object. Third, what human cloning research would be compatible with ethical and legal requirements for the use of human subjects in research is complex, controversial, and largely unexplored. For example, in what contexts and from whom would it be necessary, and how would it be possible, to secure the informed consent of parties involved in human cloning? No human cloning should ever take place without the consent of the person cloned and the woman receiving a cloned embryo, if they are different. But we could never obtain the consent of the cloned later twin to being cloned, so research on human cloning that produces a cloned individual might be barred by ethical and legal regulations for the use of human subjects in research (22). Moreover, creating human clones solely for the purpose of research would be to use them solely for the benefit of others without their consent, and so unethical. Of course, once human cloning was established to be safe and effective, then new scientific knowledge might be obtained from its use for legitimate, nonresearch reasons. How human subjects regulations would apply to research on human cloning, needs further exploration to help clarify how significant and likely the potential gains are in scientific and medical knowledge from human cloning research and human cloning.

Although there is considerable uncertainty concerning most of the possible individual and social benefits of human cloning discussed above, and although no doubt it may have other benefits or uses that we cannot yet envisage, it does appear reasonable to conclude that human cloning at this time does not promise great benefits or uniquely meet great human or social needs. Nevertheless, a case can be made that scientific freedom supports permitting research on human cloning to go forward and that freedom to use human cloning is protected by the important moral right to reproductive freedom. We must therefore assess what moral rights might be violated, or harms produced, by research on or use of human cloning.

MORAL ARGUMENTS AGAINST HUMAN CLONING

Would the Use of Human Cloning Violate Important Moral Rights?

Many of the immediate condemnations of any possible human cloning following Wilmut's cloning of an adult sheep claimed that it would violate moral or human rights, but it was usually not specified precisely, or often even at all, what the rights were that would be violated. We will consider here two possible candidates for such a right: a right to have a unique identity and a right to ignorance about one's future or to an open future. The former right is cited by many commentators, but even if any such a right exists, it seems not to be violated by human cloning. The latter right has only been explicitly defended by two commentators, and in the context of human cloning, only by Hans Jonas; it supports a more promising, even if ultimately unsuccessful, argument that human cloning would violate an important moral or human right.

Is there a moral or human right to a unique identity, and if so would it be violated by human cloning? For human cloning to violate a right to a unique identity, the relevant sense of identity would have to be genetic identity, that

is a right to a unique unrepeated genome. This would be violated by human cloning, but is there any such right? It might be thought there could not be such a right because it would be violated in all cases of identical twins, yet no one claims in such cases that the moral or human rights of each of the twins have been violated. However, this consideration is not conclusive (14,23). It is commonly held that only deliberate human actions can violate others' rights, but that outcomes that would constitute a rights violation if done by human action are not a rights violation if a result of natural causes; if Arthur deliberately strikes Barry on the head so hard as to cause his death, he violates Barry's right not to be killed, but if lightening strikes Cheryl causing her death, then we would not say that her right not to be killed has been violated. The case of twins does not show there could not be a right to a unique genetic identity.

What is the sense of identity that each person might have a right to have uniquely? What constitutes the special uniqueness of each individual (24,25)? Even with the same genes, two individuals, for example, homozygous twins, are numerically distinct and not identical, so what is intended must be the various properties and characteristics that make each individual qualitatively unique and different than others. Does having the same genome as another person undermine that unique qualitative identity? Only on the crudest genetic determinism, a genetic determinism is possible according to which an individual's genes completely and decisively determine everything else about the individual, all his or her other nongenetic features and properties, together with the entire history or biography that will constitute his or her life. But there is no reason whatever to believe that kind of genetic determinism. Even with the same genes, as we know from the case of genetically identical twins, while there may be many important similarities in the twins' psychological and personal characteristics, differences in these develop over time together with differences in their life histories, personal relationships, and life choices. This is true of identical twins raised together, and the differences are still greater in the case of identical twins raised apart; sharing an identical genome does not prevent twins from each developing a distinct and unique personal identity of their own.

We need not pursue what the basis or argument in support of a moral or human right to a unique identity might be — such a right is not found among typical accounts and enumerations of moral or human rights — because even if we grant that there is such a right, sharing a genome with another individual as a result of human cloning would not violate it. The idea of the uniqueness, or unique identity, of each person historically predates the development of modern genetics and the knowledge that except in the case of homozygous twins, each individual has a unique genome. A unique genome thus could not be the ground of this long-standing belief in the unique human identity of each person.

Would human cloning violate instead what Hans Jonas called a right to ignorance, or what Joel Feinberg called a right to an open future (26,27)? Jonas argued that human cloning in which there is a substantial time gap between the beginning of the lives of the earlier and later twin is fundamentally different from the simultaneous beginning of the lives of homozygous twins that occur in nature. Although contemporaneous twins begin their lives with the same genetic inheritance, they also begin their lives or biographies at the same time, and so in ignorance of what the other who shares the same genome will by his or her choices make of his or her life. To whatever extent one's genome determines one's future, each begins ignorant of what that determination will be and so remains as free to choose a future, to construct a particular future from among open alternatives, as are individuals who do not have a twin. Ignorance of the effect of one's genome on one's future is necessary for the spontaneous, free, and authentic construction of a life and self.

A later twin created by human cloning, Jonas argues, knows, or at least believes he or she knows, too much about him or herself. For there is already in the world another person, one's earlier twin, who from the same genetic starting point has made the life choices that are still in the later twin's future. It will seem that one's life has already been lived and played out by another, that one's fate is already determined, so the later twin will lose the spontaneity of authentically creating and becoming his or her own self. One will lose the sense of human possibility in freely creating one's own future. It is tyrannical, Jonas claims, for the earlier twin to try to determine another's fate in this way. And even if it is a mistake to believe the crude genetic determinism according to which one's genes determine one's fate, what is important for one's experience of freedom and ability to create a life for oneself is whether one thinks one's future is open and undetermined, and so still to be determined by one's own choices.

One might try to interpret Jonas' objection so as not to assume either genetic determinism or a belief in it. A later twin might grant that he is not determined to follow in his earlier twin's footsteps, but claim nevertheless that the earlier twin's life would always haunt him, standing as an undue influence in shaping his life in ways to which others' lives are not vulnerable. But the force of the objection still seems to rest on a false assumption that having the same genome as his earlier twin unduly restricts his freedom to choose a different life than the earlier twin chose. A family environment also importantly shapes children's development, but there is no force to the claim of a younger sibling that the existence of an older sibling raised in that same family is an undue influence on his freedom to make a life for himself in that environment. Indeed, the younger twin or sibling might gain the benefit of being able to learn from the older twin's or sibling's mistakes.

In a different context, and without applying it to human cloning, Joel Feinberg has argued for a child's right to an open future. This requires that others raising a child not close off the future possibilities that the child would otherwise have so as to eliminate a reasonable range of opportunities for the child to choose autonomously and construct his or her own life. One way this right to an open future would be violated is to deny even a basic education to the child, and another way might be to create the child as a later twin so that he will believe his future has already

been set for him by the choices made and the life lived by his earlier twin.

A central difficulty in evaluating the implications for human cloning of a right either to ignorance or to an open future is whether the right is violated merely because the later twin may be likely to *believe* that its future is already determined, even if that belief is clearly false and supported only by the crudest genetic determinism. If the twin's future in reality remains open and his to freely choose, then someone's acting in a way that unintentionally leads him to believe that his future is closed and determined seems not to have violated his right to ignorance or to an open future. Consider an analogous case of causing a false belief that one's right has been violated. Suppose that you drive down the twin's street in your new car that is just like his, knowing that when he sees you he is likely to believe that you have stolen his car and therefore to abandon his driving plans for the day. You have not violated his property right to his car even though he may feel the same loss of opportunity to drive that day as if you had in fact stolen his car. In each case he is mistaken that his open future or car has been taken from him, and so no right of his to them has been violated. If we know that the twin will believe that his open future has been taken from him as a result of being cloned, even though in reality it has not, then we know that cloning will cause him psychological distress but not that it will violate his right. Thus Jonas's right to ignorance, and our employment of Feinberg's analogous right of a child to an open future, turn out not to be violated by human cloning, though they do point to psychological harms that a later twin may be likely to experience and that we will take up below.

The upshot of our consideration of a moral or human right either to a unique identity or to ignorance and an open future is that neither would be violated by human cloning. Perhaps there are other possible rights that would make good the charge that human cloning is a violation of moral or human rights, but it is not clear what they might be. We turn now to consideration of the harms that human cloning might produce.

What Individual or Social Harms Might Human Cloning Produce?

There are many possible individual or social harms that have been posited by one or another commentator, and we will consider only the more plausible and significant of them.

Largely Individual Harms

Human Cloning Would Produce Psychological Distress and Harm in the Later Twin. This is perhaps the most serious individual harm that opponents of human cloning foresee, and we have just seen that even if human cloning is no violation of rights, it may nevertheless cause psychological distress or harm. No doubt knowing the path in life taken by one's earlier twin might in many cases have several bad psychological effects (13,24,28–32). The later twin may feel, even if mistakenly, that his or her fate has already been substantially laid out, and so have difficulty freely and spontaneously taking responsibility for and making his or her own fate and life. The later twin's experience or sense of autonomy and freedom may be substantially diminished, even if in actual fact they are diminished much less than it seems to him or her. Together with this might be a diminished sense of one's own uniqueness and individuality, even if once again these are in fact diminished little or not at all by having an earlier twin with the same genome. If the later twin is the clone of a particularly exemplary individual, perhaps with some special capabilities and accomplishments, he or she may experience excessive pressure to reach the very high standards of ability and accomplishment of the earlier twin (31). All of these psychological effects may take a heavy toll on the later twin and be serious burdens under which he or she would live. One commentator has also cited special psychological harms to the first, or first few, human clones from the great publicity that would attend their creation (13). While public interest in the first clones would no doubt be enormous, medical confidentiality should protect their identity. Even if their identity became public knowledge, this would be a temporary effect only on the first few clones and the experience of Louise Brown, the first child conceived by IVF, suggests this publicity could be managed to limit its harmful effects.

While psychological harms of these kinds from human cloning are certainly possible, some would argue even likely, they remain at this point only speculative, since we have no experience with human cloning and the creation of earlier and later twins. With naturally occurring identical twins, while they sometimes struggle to achieve their own identity, a struggle shared by many people without a twin, there is typically a very strong emotional bond between the twins. Such twins are, if anything, generally psychologically stronger and better adjusted than nontwins (8). It is even possible that being a later twin would confer a psychological benefit on the twin; for example, having been deliberately cloned to have his or her specific genes might make the later twin feel especially wanted for the kind of person he or she is. Nevertheless, if experience with human cloning confirmed that serious and unavoidable psychological harms typically occurred to the later twin, that would be a serious moral reason to avoid the practice.

In the discussion above of potential psychological harms to a later twin, it was assumed that one later twin is cloned from an already existing adult individual. Cloning by means of embryo splitting, as carried out and reported by Hall and colleagues at George Washington University in 1993, has limits on the number of genetically identical twins that can be cloned (33). Cloning by nuclear transfer, however, has no technological limit to the number of genetically identical individuals who might be cloned. Intuitively many of the psychological burdens and harms noted above seem more likely and serious for a clone who is only one of many identical later twins cloned from one original source, whereby the clone might run into an identical twin around every street corner. This prospect could be a good reason to place sharp limits on the number of twins that can be cloned from any one source.

There is one argument that has been used by several commentators to undermine the apparent significance of

potential psychological harms to a later twin (8,24,25). The point derives from a general problem, called the nonidentity problem posed by the philosopher Derek Parfit and not originally directed to human cloning (34). Here is the argument. Even if all the psychological burdens and pressures from human cloning discussed above can not be avoided for any later twin, they are not harms to the twin, and so not reasons not to clone the twin. That is because the only way for the twin to avoid the harms is never to be cloned and so never to exist at all. But no one claims that these burdens and stresses, hard though they might be, are so bad as to make the twin's life, all things considered, not worth living — that is, to be worse than no life at all. So the later twin is not harmed by being given a life with these burdens and stresses, since the alternative of never existing at all is arguably worse — he or she loses a worthwhile life — but certainly not better for the twin. And if the later twin is not harmed by having been created with these unavoidable burdens and stresses, then how could he or she be wronged by having been created with them? And if the later twin is not wronged, then why is any wrong being done by human cloning? This argument has considerable potential import, for if it is sound it will undermine the apparent moral importance of any bad consequence of human cloning to the later twin that is not so serious as to make the twin's life all things considered not worth living.

Parfit originally posed the nonidentity problem, but he does not accept the above argument as sound. Instead, he believes that if one could have a *different* child without these psychological burdens (e.g., by using a different method of reproduction that does not result in a later twin), there is as strong a moral reason to do so as there would be not to cause similar burdens to an already existing child; this position has been defended in the general case of genetically transmitted handicaps or disabilities (35). The theoretical philosophical problem is to formulate the moral principle that implies this conclusion and that also has acceptable implications in other cases involving bringing people into existence, such as issues about population policy. The issues are too detailed and complex to pursue here and the nonidentity problem remains controversial and not fully resolved, but suffice it to say, what is necessary is a principle that permits comparison of the later twin with these psychological burdens and a different person who could have been created instead, for example, by a different means of reproduction, without such burdens. Choosing to create the later twin with serious psychological burdens instead of a different person who would be free of them, without a weighty overriding reason for choosing the former, would be morally irresponsible or wrong, even if doing so does not harm or wrong the later twin who could only exist with the burdens. At the least, the argument for disregarding the psychological burdens to the later twin because he or she could not exist without them is controversial, and it does not justify ignoring unavoidable psychological burdens to later twins among reasons against human cloning. Such psychological harms, as we will continue to call them, do remain speculative, but they should not be disregarded because of the nonidentity problem.

Human Cloning Procedures Would Carry Unacceptable Risks to the Clone. One version of this objection to human cloning concerns the research necessary to perfect the procedure, the other version concerns the later risks from its use. Wilmut's group had 276 failures before their success with Dolly, indicating that the procedure is far from perfected even with sheep. Further research on the procedure with animals is clearly necessary before it would be ethical to use the procedure on humans. But even assuming that cloning's safety and effectiveness was established with animals, research would need to be done to establish its safety and effectiveness for humans. Could this research be ethically done (36)? There would be little or no physical risk to the donor of the cell nucleus to be transferred, and his or her informed consent could and must always be obtained. There might be greater risks for the woman to whom a cloned embryo is transferred, but these should be comparable to those associated with IVF procedures and the woman's informed consent too could and must be obtained.

What of the risks to the cloned embryo itself? Judging by the experience of Wilmut's group in their work on cloning a sheep, the principal risk to the embryos cloned was their failure successfully to implant, grow, and develop. Comparable risks to cloned human embryos would apparently be their death or destruction long before most people or the law consider them to be persons with moral or legal protections of their life. Moreover artificial reproductive technologies now in use, such as IVF, have a known risk that some embryos will be destroyed or will not successfully implant and will die. It is premature to make a confident assessment of what the risks to human subjects would be of establishing the safety and effectiveness of human cloning procedures, but there are no unavoidable risks apparent at this time that would make the necessary research clearly ethically impermissible.

Could human cloning procedures meet ethical standards of safety and efficacy? Risks to an ovum donor (if any), a nucleus donor, and a woman who receives the embryo for implantation would likely be ethically acceptable with the informed consent of the involved parties. But what of the risks to the human clone if the procedure in some way goes wrong, or unanticipated harms come to the clone; for example, Harold Varmus, director of the National Institutes of Health, has raised the concern that a cell many years old from which a person is cloned could have accumulated genetic mutations during its years in another adult that could give the resulting clone a predisposition to cancer or other diseases of aging (37). Moreover it is impossible to obtain the informed consent of the clone to his or her own creation, but of course, no one else is able to give informed consent for their creation either.

It is too soon to say with any confidence whether unavoidable risks to the clone would make human cloning unethical. At a minimum, further research on cloning animals, as well as research to better define the potential risks to humans, is needed. For the reasons given above, we should not set aside risks to the clone on the ground that the clone would not be harmed by them, since its only alternative is not to exist at all; that is a bad argument. But we should not insist on a standard that requires risks

to be lower than those we accept in sexual reproduction, or in other forms of assisted reproduction. It is not possible now to know when, if ever, human cloning will satisfy an appropriate standard limiting risks to the clone.

Largely Social Harms

Human Cloning Would Lessen the Worth of Individuals and Diminish Respect For Human Life. Unelaborated claims to this effect in the media were common after the announcement of the cloning of Dolly. Ruth Macklin has explored and criticized the claim that human cloning would diminish the value we place on, and our respect for, human life because it would lead to persons being viewed as replaceable (24). Just as with a supposed right to a unique identity, only on a confused and indefensible notion of human identity is a person's identity determined solely by his or her genes. Instead, an individual's identity is determined by the interaction of his or her genes over time with his or her environment, including the choices the individual makes and the important relations he or she forms with other persons. This means in turn that no individual could be fully replaced by a later clone possessing the same genes. Ordinary people recognize this clearly. For example, parents of a 12-year-old child dying of a fatal disease would consider it insensitive and ludicrous if someone told them they should not grieve for their coming loss because it is possible to replace their dying child by cloning him; it is *their child who is dying* whom they love and value, and that child and his importance to them could never be replaced by a cloned later twin. Even if they would also come to love and value a later twin as much as their child who is dying, that would be to love and value that *different child* who could never replace the child they lost. Ordinary people are typically quite clear about the importance of the relations they have to distinct, historically situated individuals with whom over time they have shared experiences and their lives, and whose loss to them would therefore be irreplaceable.

A different version of this worry is that human cloning would result in persons' worth or value seeming diminished because we would now see humans as able to be manufactured or "hand-made." This demystification of the creation of human life might reduce our appreciation and awe of it and of its natural creation. It would be a mistake, however, to conclude that a human being created by human cloning is of less value or is less worthy of respect than one created by sexual reproduction. At least outside of some religious contexts, it is the nature of a being, not how she is created, that is the source of her value and makes her worthy of respect. Moreover, for many people gaining a scientific understanding of the extraordinary complexity of human reproduction and development increases, instead of decreases, their awe of the process and its product.

A more subtle route by which the value we place on each individual human life might be diminished could come from the use of human cloning with the aim of creating a child with a particular genome, either the genome of another individual especially meaningful to those doing the cloning or an individual with exceptional talents, abilities, and accomplishments. The child might then be valued only for his or her genome, or at least for his or her genome's expected phenotypic expression, and no longer be recognized as having the intrinsic equal moral value of all persons, simply as persons. For the moral value and respect due all persons to come to be seen as resting only on the instrumental value of individuals, or of individuals' particular qualities, to others would be to fundamentally change the moral status accorded to persons. Everyone would lose their moral standing as full and equal members of the moral community, replaced by the different instrumental value each of us has to others.

Such a change in the equal moral value and worth accorded to persons should be avoided at all costs, but it is far from clear that such a change would be the unavoidable result of permitting human cloning. Parents, for example, are quite capable of distinguishing their children's intrinsic value, just as individual persons, from their instrumental value based on their particular qualities or properties. The equal moral value and respect due all persons just as persons is not incompatible with the different instrumental value of people's particular qualities or properties; Einstein and an untalented physics graduate student have vastly different value as scientists, but share and are entitled to equal moral value and respect as persons. It would be a confusion and a mistake to conflate the two kinds of value and respect. Making a large number of clones from one original person might be more likely to foster this confusion and mistake in the public, and if so that would be a further reason to limit the number of clones that could be made from one individual.

Human Cloning Would Divert Resources from Other More Important Social and Medical Needs (13,28). As we saw in considering the reasons for, and potential benefits from, human cloning, in only a limited number of uses would it uniquely meet important human needs. There is little doubt that in the United States, and certainly elsewhere, there are more pressing unmet human needs, both medical or health needs and other social or individual needs. This is a reason for not using public funds to support human cloning, at least not if the funds will in fact be redirected to more important ends and needs. It is not a reason, however, either to prohibit other private individuals or institutions from using their own resources for research on human cloning or for human cloning itself, or to prohibit human cloning or research on human cloning.

The other important point about resource use is that it is not now clear how expensive human cloning would ultimately be, for example in comparison with other means of relieving infertility. The procedure itself is not extremely complex scientifically or technologically and might prove not to require a significant commitment of resources.

Human Cloning Might be Used by Commercial Interests for Financial Gain. Both opponents and proponents of human cloning agree that cloned embryos should not be able to be bought and sold. In a science fiction frame of mind, one can imagine commercial interests offering genetically certified and guaranteed embryos for sale, perhaps offering a catalog of different embryos cloned from individuals with a variety of talents, capacities, and other desirable properties. This would be a fundamental violation of the equal moral respect and dignity owed to all persons, treating them instead as objects to be differentially valued,

bought, and sold in the marketplace. Even if embryos are not yet persons at the time they would be purchased or sold, they would be being valued, bought, and sold for the persons they will become. The moral consensus against any commercial market in embryos, cloned or otherwise, should be enforced by law whatever public policy ultimately is on human cloning. It has been argued that the law may already forbid markets in embryos on grounds that they would violate the thirteenth amendment prohibiting slavery and involuntary servitude (38).

Human Cloning Might be Used by Governments or Other Groups for Immoral and Exploitative Purposes. In *Brave New World*, Aldous Huxley imagined cloning individuals who have been engineered with limited abilities and conditioned to do, and to be happy doing, the menial work that society needed done (39). Selection and control in the creation of people was exercised not in the interests of the persons created, but in the interests of the society and at the expense of the persons created. Any use of human cloning for such purposes would exploit the clones solely as means for the benefit of others, and would violate the equal moral respect and dignity they are owed as full moral persons. If human cloning is permitted to go forward, it should be with regulations that would clearly prohibit such immoral exploitation.

Fiction contains even more disturbing and bizarre uses of human cloning, such as the Nazi, Joseph Mengele's creation of many clones of Hitler in Ira Levin's *The Boys from Brazil*, Woody Allen's science fiction cinematic spoof *Sleeper* in which a dictator's only remaining part, his nose, must be destroyed to keep it from being cloned, and the contemporary science fiction film *Blade Runner*. Nightmare scenarios like Huxley's or Levin's may be quite improbable, but their impact should not be underestimated on public concern with technologies like human cloning. Regulation of human cloning must assure the public that even such farfetched abuses will not take place.

Human Cloning Used on a Very Widespread Basis Would Have a Disastrous Effect on the Human Gene Pool by Reducing Genetic Diversity and Our Capacity to Adapt to New Conditions (11). This is not a realistic concern, since there is little if any reason to believe that human cloning would be used on a wide enough scale to have the feared effect on the gene pool. The vast majority of humans seem quite satisfied with traditional sexual means of reproduction; if anything, from the standpoint of worldwide population, we could do with a bit less enthusiasm for it. Programs of eugenicists like Herman Mueller earlier in the century to impregnate thousands of women with the sperm of exceptional men, as well as the more recent establishment of sperm banks of Nobel laureates, have met with little or no public interest or success (40). People prefer sexual means of reproduction and they prefer to keep their own biological ties to their offspring.

CONCLUSION

Human cloning has until now received little serious and careful ethical attention because it was typically dismissed as science fiction, and it stirs deep, but difficult to articulate, uneasiness and even revulsion in many people. Any ethical assessment of human cloning at this point must be tentative and provisional. Fortunately the science and technology of human cloning are not yet in hand, so a public and professional debate is possible without the need for a hasty, precipitate policy response.

The ethical pros and cons of human cloning seem at this time to be sufficiently closely balanced and uncertain that there is not an ethically decisive case either for or against permitting it or doing it. Access to human cloning can plausibly be brought within a moral right to reproductive freedom, but it is not a central component of a moral right to reproductive freedom. The circumstances in which its use would have significant benefits appear at this time to be few and infrequent, and it is not the solution to any major or pressing individual or social needs. On the other hand, contrary to the pronouncements of many of its opponents, human cloning seems not to be a violation of any moral or human rights. While it does risk some significant individual or social harms, most are based on common confusions about genetic determinism, human identity, and the effects of human cloning. Because most moral reasons against doing human cloning remain speculative, they seem insufficient to warrant a complete legal prohibition of either research on or later use of human cloning. Legitimate moral concerns about the use and effects of human cloning, however, underline the need for careful public oversight of research on its development, together with a continued and wider public debate and review before cloning is used on human beings.

BIBLIOGRAPHY

1. I. Wilmut et al., *Nature* **385**, 810–813 (1997).
2. WHO (World Health Organization) Press Office, Geneva, Switzerland, March 11, 1997.
3. Council of Europe, *Recommendation 1046 on the Use of Human Embryos and Fetuses for Diagnostic, Therapeutic, Scientific, Industrial and Commercial Purposes*, 1988.
4. J. Cohen and G. Tomkin, *Kennedy Inst. Ethics J.* **4**, 193–204 (1994).
5. R. Dworkin, *Taking Rights Seriously*, Duckworth, London, 1978.
6. J.S. Mill, *On Liberty*, Bobbs-Merrill Publishing, Indianapolis, IN, 1859.
7. R. Rhodes, *Cambridge Q. Healthcare Ethics* **4**, 285–290 (1995).
8. J.A. Robertson, *Children of Choice: Freedom and the New Reproductive Technologies*, Princeton University Press, Princeton, NJ, 1994.
9. D.W. Brock, in D. Thomasma and J. Monagle, eds., *Health Care Ethics: Critical Issues for Health Professionals*, Aspen Publishers, Gaithersburg, MD, 1994.
10. G.J. Annas, *Kennedy Inst. Ethics J.* **4**(3), 235–49 (1994).
11. L. Eisenberg, *J. Med. Philos.* **1**, 318–331 (1976).
12. J.A. Robertson, *Hastings Center Rep.* **24**, 6–14 (1994).
13. M. LaBar, *Thought* **57**, 318–333 (1984).
14. NABER (National Advisory Board on Ethics in Reproduction), *Kennedy Inst. Ethics J.* **4**, 251–282 (1994).
15. C. Kahn, *Free Inquiry* **9**, 14–18 (1989).

16. J. Harris, *Wonderwoman and Superman: The Ethics of Biotechnology*, Oxford University Press, Oxford, UK, 1992.
17. L. Thomas, *N. Engl. J. Med.* **291**, 1296–1297 (1974).
18. J. Lederberg, *Am. Naturalist* **100**, 519–531 (1966).
19. R. McKinnell, *Cloning: A Biologist Reports*, University of Minnesota Press, Minneapolis, MN, 1979.
20. W.A.W. Walters, W.A.W. Walters, and P. Singer, eds., *Test-Tube Babies*, Oxford University Press, Oxford, UK, 1982.
21. G.P. Smith, *University of New South Wales Law J.* **6**, 119–132 (1983).
22. P. Ramsey, *Fabricated Man: The Ethics of Genetic Control*, Yale University Press, New Haven, CT, 1970.
23. L. Kass, *Toward a More Natural Science*, Free Press, New York, 1985.
24. R. Macklin, *Kennedy Inst. Ethics J.* **4**, 209–226 (1994).
25. R.F. Chadwick, *Philos.* **57**, 201–209 (1982).
26. H. Jonas, *Philosophical Essays: From Ancient Creed to Technological Man*, Prentice-Hall, Englewood Cliffs, NJ, 1974.
27. J. Feinberg, in *Parental Authority, and State Power, Whose Child? Children's Rights*, W. Aiken and H. LaFollette, eds., Rowman and Littlefield, Totowa, NJ, 1980.
28. D. Callahan, *Los Angeles Times*, November 12, p. B7 (1993).
29. R. McCormick, *Notes on Moral Theology: 1965 through 1980*, University Press of America, Washington, DC, 1981.
30. A. Studdard, *Man Med. J. Values Ethics Health Care* **3**, 109–114 (1978).
31. J.D. Rainer, "Commentary." *Man Med. J. Values Ethics Health Care* **3**, 115–117 (1978).
32. A.D. Verhey, *Kennedy Inst. Ethics J.* **4**, 227–234 (1994).
33. J.L. Hall et al., *1993 Abstracts of the Scientific Oral and Poster Sessions*, Program Suppl. Abstract 0-001, S1, American Fertility Society/Canadian Fertility and Andrology Society, 1993.
34. D. Parfit, *Reasons and Persons*, Oxford University Press, Oxford, UK, 1984.
35. D.W. Brock, *Bioethics* **9**, 269–275 (1995).
36. R. Pollack, *N.Y. Times*, November 17, 1993.
37. R. Weiss, *Int. Herald Trib*, March 7, 1997.
38. P.O. Turner, *University of Detroit J. Urban Law* **58**, 459–487 (1981).
39. A. Huxley, *Brave New World*, Chalto & Winders, London, 1932.
40. M. Adams, ed., *The Well-Born Science*, Oxford University Press, Oxford, UK, 1990.

See other CLONING entries; FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, NATIONAL BIOETHICS ADVISORY COMMISSION.

CLONING, OVERVIEW OF ANIMAL CLONING

GARY VARNER
Texas A&M University
College Station, Texas

OUTLINE

INTRODUCTION

Animal cloning is an area where the ethics and the science are both just coming into focus at the start of the twenty-first century. Although the cloning of embryos had been used for years in production agriculture, prior to the late 1990s, many or most scientists doubted that it would soon be possible to clone an adult animal of a complex species. Dolly's appearance in 1997 sparked renewed interest in cloning, especially the implications of applying the technique to human beings. However, many questions about the cloning of animals are just beginning to be addressed. On the scientific front, it is still unclear how to clone adults reliably and efficiently, and it is still unclear what the effects of the process are on the DNA of the resulting clones. On the philosophical front, although much has been published on the ethics of human cloning since Dolly's appearance, almost nothing has been published on the implications of adult cloning for animal welfare and animal rights (on the science of cloning, see K. Hanna, this volume).

ANIMAL CLONING: SCIENTIFIC QUESTIONS

In February 1997 a team led by Scottish scientist Ian Wilmut announced that it had successfully cloned a six-year-old ewe. Dolly, as they named the sheep clone, became a media celebrity overnight. Never before had a clone been produced from an adult animal. Cattle and sheep embryos had been cloned for years, but most scientists had doubted that it would be possible to clone adults in the foreseeable future.

The first sheep clone was produced in 1979 and the first bovine clone the next year, but these were produced by separating the cells of blastomeres. Blastomeres are very early embryos in which no differentiation of cells has yet occurred. Although still an expensive technique, by the 1990s blastomere separation had come to play an important role in production agriculture. For instance, an embryonic bull or cow of a promising breed line can be cloned, and the clones can then be frozen and shipped to farmers or stock breeders for implantation in surrogate mothers.

The technique used to produce Dolly was radically different. Although almost every cell of a mature animal contains its entire genetic code, in all somatic cells (i.e., body cells, as opposed to "germ" cells — i.e., sperm and

eggs) the vast majority of these genes have been somehow "turned off." The nucleus of a bone cell, for instance, contains all of the genetic "instructions" necessary to build the entire animal, but only those needed to produce a bone cell are "active." To clone an adult animal from a somatic cell it is necessary to somehow "reactivate" all this unused DNA, but since no one yet understood how genes are turned on and off, most scientists were skeptical that a clone of an adult animal would be produced within the foreseeable future. Indeed, when Dolly's existence was announced, some scientists publicly expressed skepticism that she was really a clone.

The Wilmut team used a process called nuclear transplantation, in which the nucleus of an egg is replaced by the nucleus of another cell. This technique had first been used in the 1980s, but until 1997 had only been known to work when the transferred nucleus came from an embryo cell. Attempts to clone adult frogs from skin cells, for instance, resulted in some tadpole births but no further development (1). The Wilmut team tried transferring nuclei from the mammary glands of an adult ewe and then manipulating them so that the cells first quit growing and then, after a quiescent phase, began dividing again. They reasoned that something about this manipulation would "facilitate reprogramming of gene expression" (2, p. 811). Apparently they were right. In the summer of 1998, DNA analysis confirmed that Dolly was indeed a clone of the six-year-old ewe the Wilmut team had claimed was the nucleus donor (3,4), and over the next 18 months a number of other teams announced successes using various adult cells as nucleus donors and using variations on the Wilmut team's process for manipulating the resulting blastocysts through a quiescent phase and back into growth. A Japanese team of scientists announced that it had successfully cloned large numbers of mice using cumulus cells (which surround eggs inside the ovaries), and had even made clones of adult clones (5). Then labs in France and Texas announced that they had cloned cattle from skin cells (6–8).

Although cloning of adult animals using the somatic cell nuclear transfer technique — which, following (9), we can simply call "somacloning" — is an acknowledged reality at the beginning of the twenty-first century, the process is still unreliable and questions remain about the health of somaclones. Media coverage of the early work described above emphasized the high ratio of nuclear transfers to live births and a range of concerns about the health of somaclones.

It took 277 nuclear transfers to produce Dolly, and the Wakayama team did almost 1400 transfers on the way to producing their first 31 live births. In some teams' work there was a high rate of lost pregnancies and/or perinatal deaths (6,7,9). And since mitochondria outside the nucleus of the egg into which the donor nucleus is inserted carry some of the egg donor's DNA, somaclones might face problems of compatibility between the mitochondrial DNA of the egg donor and the DNA of the nucleus donor.

One of the most discussed potential problems was premature aging due to telomere shortening. Telomeres are sequences of DNA at the ends of all chromosomes. Although they do not code for any genes, they function to protect the gene-coding portion of the chromosome from being lost during cell division, which strips several base pairs of DNA from each end of the chromosome every time the cell divides. Telomere shortening is widely suspected of being tied to senescence, so a clone produced from a 21-year-old steer might be expected to age suddenly and prematurely. In May 1999 Dolly's telomeres were found to be significantly shorter than age-matched ewes produced from embryonic donor nuclei using the nuclear transfer method (10), but Dolly had by then delivered two healthy lambs herself and the Wakayama team's second generation mouse clones had themselves reared healthy offspring. Of particular interest in this respect was Texas A&M University's cloning of a 21-year-old steer named Chance, whose clone (Second Chance) was born after Chance had died of old age. However, that bull is still quite young at the time of this writing, and at present, then, the jury is still out on the effects of telomere shortening in somaclones. However, even if it turns out that current somaclones suffer from telomere shortening, it may be possible in the future to treat the problem using telomerase, the enzyme which "resets" telomere length in sperm and eggs, allowing naturally produced embryos to begin life with full-length telomeres.

ANIMAL CLONING: ETHICAL AND PHILOSOPHICAL QUESTIONS

The scientific questions discussed in the preceding section are directly relevant to some of the ethical concerns commonly raised in early media coverage of somacloning. A spate of philosophical work on cloning followed Dolly's appearance, but almost none of the articles focused on animal cloning as anything other than a harbinger of, or a slippery slope toward, cloning humans. Accordingly this section addresses the ethical issues most commonly raised in media treatments of somacloning and considers the significance of these concerns from the perspective of animal rights and animal welfare philosophies as those were developed in the last quarter of the twentieth century by philosophers concerned primarily with more traditional uses of animals.

The most commonly raised concern about animal cloning, that it would lead to human cloning, will not be considered here. (For a summary of ethical concerns about human cloning, see D. Brock, this volume.) There is only a limited overlap between the concerns commonly expressed about human cloning and those expressed about animal cloning. Since even the most cognitively sophisticated nonhuman animals presumably lack the ability to understand that they are clones, worries commonly expressed about human clones resenting the fact that they are clones, or suffering from knowing that their earlier-born twins have made certain decisions or suffered certain inherited diseases are irrelevant to nonhuman animals. And because animal breeding has been used for millennia to "improve breeding stock," concerns about cloning of "gifted" individuals and about cloning reducing the worth of individuals seem at least less intense than the analogous concerns about human clones. There are legitimate economic and environmental

(and thus social) concerns about increasing genetic homogeneity in farm animals, and people who are worried about human cloning are concerned that the practice of animal cloning puts us on a slippery slope to that, but considered in and of itself, the cloning of animals has been objected to primarily on two grounds. The first is that it is unnatural or amounts to "playing God." The second is that it violates animals' rights and/or may have dramatic adverse impacts on their welfare. This section discusses these objections in turn, along with a related philosophical question about the moral significance of unactualized potentials.

The "It's Unnatural" and "Playing God" Objections

As Rolston (11) notes, there are two "absolute" and antithetically opposed senses of the term "natural." On one extreme, anything that happens in accordance with natural law (the laws of physics, biology, gravity, etc.) is natural, but in this sense nothing that human beings could do would be unnatural. If cloning works, it is consistent with the laws of biological nature, whether or not anyone presently understands how "reprogramming of gene expression" works. On the other extreme is an "artifactual" sense of the unnatural, according to which a process is natural to the extent that it is not guided by human intention or deliberation. In this sense, somacloning is extremely unnatural, but so are highways, computers, and good sanitation. And, Rolston notes, even the attempt to act naturally is in this sense unnatural, insofar as it is intentional or deliberate. So it is difficult to see how either of these "absolute" senses of "unnatural" helps us articulate a clear ethical concern about somacloning.

Varner (12) identifies two other senses of the "natural." First, the term is sometimes used to pick out what is distinctive or unique about a species. Aristotle employed this sense of "natural" in a famous argument in his *Nicomachean Ethics*. Aristotle used the Greek term *ergon* (translated variously as "work" or "function") to refer to whatever capacity the members of a taxon share with no "lower" organisms (13). He concluded that nutrition and growth are *ergon* for plants, sense perception and motion are *ergon* for animals, and reason is *ergon* for humans (14). However, it is unclear that this in any way helps us sort out ethical concerns about cloning, because cloning research and applications are paradigm cases of reasoned inquiry, and thus they are, in Aristotle's sense, perfectly natural things for humans to do. Also, in light of growing scientific evidence that animals engage in reasoning of various kinds, a more plausible candidate for the Aristotelean *ergon* of humans would be some kind of moral reasoning, so Aristotle's *ergon* argument at most raises the question of whether or not cloning is morally acceptable, without providing any part of the answer.

Varner notes that in yet another sense of the term, a process or activity is "natural" for a given species to the extent that it was characteristic of the species in its evolutionary past and/or it played an important role in natural selection's producing the various traits that are now characteristic of the species. Hunting is often claimed to be natural for humans in this sense, insofar as distinctively human traits like language and the very sophisticated group planning and coordination, which language makes possible, appear to have evolved because they were adaptive for our distant ancestors. Is cloning unnatural in this sense of the term? Many adult plants and animals reproduce by having clones "bud off" from them, and even in mammals, clonal reproduction occurs whenever identical twins are produced. So it is at most *somacloning* that is unnatural for humans and for the other mammals involved in research to date. However, it is unclear why this should make somacloning a bad thing, for plenty of things are unnatural in this sense without being bad, for instance, vegetarian diets and living without engaging in hunting.

Thus claims about what is natural and unnatural do not seem likely to be helpful in articulating the substance of ethical concerns about cloning, and similar things can be said about the objection that by engaging in cloning we are "playing God." Ruth Chadwick observes that claims about "playing God" are sometimes used to express fear of unforeseen consequences. The Greek concept of *hubris* involves people overestimating their power and suffering severe, unforeseen consequences, sometimes via the wrath of the gods. But, Chadwick argues, if the fear is of unforeseen consequences, it would be more clear to say so and "discard the concept of 'playing God'" (15, p. 203).

More generally, basing morality on God's will is philosophically problematic and basing public policy on religion is politically problematic in a pluralist society. Philosophically, the problem can be posed in terms of what is sometimes called "the Euthyphro Question" after a similar question posed about piety in a Platonic dialogue (16): *Is what is right right because God wills it, or does God will what He wills because it is right*.

In the former case, anything would be right if God did will it, including putatively unjust practices like slavery, sexism, wars of aggression, and so on. Also, if doing the right thing means nothing more than acting in accordance with God's will, then God's goodness is trivialized. For on such a view, "John is a good person" seems to mean roughly that "John does what God wills." But by parity of reasoning, on such a view, "God is good" would mean nothing more than "God does what God wills."

The second answer to the Euthyphro question, that "God wills what He wills because it is right," preserves the significance of claims about God's goodness. On such a view, it is at least possible, in principle, to investigate moral claims independently of God's will (since morality is logically prior to God's will on this view). In practice, doing so would be necessary for two reasons: first, to resolve ambiguities in the evidence about what God wills, and second, to determine what God would will about issues (like somacloning, presumably) that are not addressed in the available evidence regarding God's will. For on this view, God is like a wise and benevolent elder who can be depended upon for reliable advice about what is right and what is wrong, but when evidence about what God wills is lacking or ambiguous, the only way to determine what God wills would be to determine, independently, what would be the right thing for Him to will.

So just like claims that cloning is unnatural, claims that cloning is "playing God" seem unlikely to help

elucidate and evaluate ethical concerns about cloning. If the concerns cannot be expressed clearly and persuasively without relying on religious assumptions based only in faith, then most people would conclude that the concerns cannot justify sweeping restrictions on cloning in a pluralist society (17).

Animal Welfare and Animal Rights

Since the publication of Peter Singer's book *Animal Liberation* in 1975, philosophers have shown a keen interest in the areas of "animal rights" and "animal welfare." In her contribution to this volume, "Research on Animals, Ethics, and the Moral Status of Animals," Lilly-Marlene Russow questions the usefulness of this distinction. Some of the reasons she gives mirror points raised below, but the received interpretation of the distinction still provides a much better template for analyzing ethical concerns about animal cloning than does either talk of its unnaturalness or its being "playing God." The philosophical characterizations of animal welfare and animal rights sketched briefly below are fairly standard. More detail is available in Russow's contribution to this encyclopedia, and in (18–21) as well as the many works referenced there.

The Distinction as Popularly Conceived

Any careful consideration of animal welfare and animal rights concerns must begin with a clear characterization of what counts as one or the other type of view. Much confusion exists about this, because popular accounts of the distinction typically cut the pie very differently than philosophers writing on the topic.

In particular, in media treatments of animal issues, "animal rightists" are typically distinguished from "animal welfarists" in terms of the political ends and means the two endorse, rather than in terms of their underlying philosophical commitments. Accordingly animal rightists are typically portrayed as seeking the abolition of various practices, whereas animal welfarists are said to seek reform of problematic practices rather than elimination of animal use. Animal rightists are portrayed as willing to employ illegal and even violent tactics, whereas animal welfarists are portrayed as "working within the system." And, perhaps because the term seems to have been coined by agriculturalists, medical researchers, hunters, and so forth, whose colleagues felt threatened by self-professed animal rights activists, animal welfarists typically are portrayed as reasonable, well-informed people, in contrast to animal rightists who are often characterized as irrational and poorly informed people acting in the grip of their emotions.

Self-professed animal welfarists are seen to defend some uses of animals, while emphasizing that they take seriously a moral imperative to take animals' interests into account — that they are not neo-Cartesians who deny all significance to animals' lives and suffering. So it is not surprising that they wish to self-consciously distance themselves — in both their colleagues' and in the public's eyes — from self-professed animal rights activists, whose actions have sometimes richly earned them the above stereotype (21). However, this stark political characterization of the animal rights/animal welfare distinction in the media hides an underlying overlap in philosophical commitments that is crucial to assessing ethical objections to animal cloning.

Animal Welfare, Philosophically Conceived

Among philosophers writing on related issues, the animal welfare stance generally is taken to assume a specific ethical principle, utilitarianism, and a specific understanding of how animals' happiness or well-being, and the opposites of these, are to be assessed.

Utilitarianism is the view that right actions and institutions maximize aggregate happiness. The "classical" utilitarians of the nineteenth century — Jeremy Bentham (22) and John Stuart Mill (23) — championed various social causes on utilitarian grounds, arguing that various penal code, educational, and welfare reforms would increase the happiness of underprivileged peoples more than the marginal cost of such programs to more privileged persons, thus maximizing human happiness in the aggregate.

Bentham and Mill also both endorsed a hedonistic conception of happiness. Both claimed that the only thing that is intrinsically good is pleasure and the only thing that is intrinsically bad is pain. To say that something is intrinsically good is to say that its existence is good in and of itself, independently of its relationship to other things, and independently of its instrumental value in the pursuit of other ends. And since at least some nonhuman animals are capable of feeling pain, both Mill and Bentham concluded that these animals' happiness must be taken into consideration in the utilitarian calculus.

A widely cited passage in Bentham's *An Introduction to the Principles of Morals and Legislation* conveys the logic of this extension of the moral community and, taken in context, underlines an important point about animal welfare views as philosophically construed. Observing that animals have traditionally been mere "things" in the eyes of the law, Bentham adds:

> The day *may* come, when the rest the animal creation may acquire those rights which never could have been withholden from them but by the hand of tyranny. The French have already discovered that the blackness of the skin is no reason why a human being should be abandoned without redress to the caprice of a tormentor. It may come one day to be recognized, that the number of the legs, the villosity of the skin, or the termination of the *os sacrum*, are reasons equally insufficient for abandoning a sensitive being to the same fate. What else is it that should trace the insuperable line? Is it the faculty of reason, or, perhaps, the faculty of discourse? But a full-grown horse or dog is beyond comparison a more rational, as well as a more conversable animal, than an infant of a day, or a week, or even a month, old. But suppose the case were otherwise, what would it avail? The question is not, Can they *reason*? nor, Can they *talk*? but, Can they *suffer*? (22, p. 412).

In this passage Bentham clearly holds that sentience, rather than some more sophisticated cognitive capacity or species membership, is the criterion of moral standing, and

a criterion which many animals clearly meet. (Although "sentience" means, generally, consciousness or awareness of something, in discussions of animal welfare and animal rights the term is usually taken to mean consciousness of pain [and/or pleasure] specifically.) However, immediately preceding this widely quoted passage, Bentham states that

> If the being eaten were all, there is very good reason why we should be suffered to eat such of them as we like to eat; we are the better for it, and they are never the worse. They have none of those long-protracted anticipations of future misery which we have. The death they suffer in our hands commonly is, and always may be, a speedier and by that means a less painful one, than that which would await them in the inevitable course of nature (22, pp. 411–412).

That is, in hedonistic utilitarian terms, a sufficiently humane form of slaughter is perfectly permissible. (Which probably explains why only the portion of the passage quoted first above appears in animal rights tracts.)

Self-styled animal welfarists usually share Bentham's and Mill's basic philosophical commitments regarding animals. That is, they think in utilitarian terms (at least when it comes to animals — see the next subsection regarding their attribution of rights to humans), and they think of individual happiness in hedonistic terms (at least when it comes to animals). So it is common to hear animal welfarists defend humane agricultural systems in roughly Bentham's own terms, medical research on animals on the ground that the harms to animals are greatly outweighed by the suffering humans (and other animals) are spared when cures are found for debilitating diseases, and hunting in terms of reducing animal suffering on the whole. And in doing their utilitarian calculus, self-professed animal welfarists commonly speak as if animal happiness consists entirely in feeling pleasure while avoiding pain so that, for instance, if all pain and suffering is removed from the protocol of an experiment with important clinical implications, all ethical concerns about the experiment have been addressed.

Philosophically speaking, adopting the animal welfare perspective means using hedonistic utilitarian thinking to assess various human uses of animals, and so conceived, both defenders of animal research, agriculture, and hunting, and some prominent critics of these practices, are animal welfarists. Significantly Singer (24), whom the popular, political characterization of the distinction places squarely in the animal rights camp, is himself a utilitarian and, with some important qualifications (see below), a hedonistic utilitarian. How, then, does Singer reach such different conclusions regarding the very practices which self-professed animal welfarists defend? The answer is that while sharing certain deep philosophical commitments about animals, the two differ in their understanding of what those commitments imply in practice. Singer (19,24) argues that among other things, defenders of these practices exaggerate the magnitude and the certainty of benefits to humans and other animals, and underestimate the costs to animals involved in agriculture, research, and hunting.

Animal Rights, Philosophically Conceived

Philosophically construed, a true animal *rights* view must be nonutilitarian. This subsection briefly sketches a paradigm example: Tom Regan's view in *The Case for Animal Rights* (18).

Regan argues that whether or not the defenders of animal agriculture, research, and hunting are right that the aggregate benefits outweigh the aggregate costs, these practices are still wrong from a perspective that extends respect for individuals from humans to animals. In day-to-day talk about ethics, appeals to moral "rights" are often used to claim such respect for individual humans. To say that someone "has a moral right" to something usually means that we would not be justified in infringing that right on purely utilitarian grounds. For instance, an aggressive fundamentalist preacher's right to free speech may not be overridden just because the aggregate suffering of listeners and passers-by (in terms of anger, frustration, and hurt feelings) outweighs the joy the preacher and his followers get from him speaking. Self-styled animal welfarists commonly invoke such rights claims regarding humans, while denying that rights are relevant where animals are concerned. Thus one might claim that while it would never be permissible to experiment on humans without their consent, the utilitarian calculus suffices to justify this in the case of nonhuman animals.

In *The Case for Animal Rights* Regan examined the consequences of extending such respect for individuals from humans to animals. Specifically, Regan's view is that respect for individuals involves treating them as more than mere "utility receptacles," that if an individual "has moral rights," then we cannot justify harming him or her (or it) on the sole basis that doing so maximizes aggregate happiness. Respect for individuals requires nonutilitarian reasons for involuntarily imposing harm, or significant risk of harm.

Which animals deserve this kind of respect? Regan argues that given the range of humans who are usually thought to deserve it, consistency requires us to extend it to all animals who are "subjects of a life," by which he means, roughly, that they have memories, a sense of their own future, and desires or preferences about their future (18, p. 243). He argues that available empirical evidence shows convincingly that at least all normal adult mammals have these capacities, and that birds probably do too (18).

One of the obvious questions for anyone who, like Regan, construes rights claims as "trump cards" against utilitarian arguments is what to do where rights conflict. Regan defends two principles for use in such cases (18, pp. 301–312). For situations involving comparable harms to various individuals, Regan defends what he calls "the miniride" principle. This requires that where rights violations are inevitable, we minimize the violations of rights (hence the name of the principle). Although Regan admits that where it applies, this principle would have the same implications as utilitarianism, he denies that it is utilitarian in spirit or justification. Regan's other principle applies to cases involving dramatically unequal degrees of harm. This principle requires that we avoid harming "the worse-off individual," which means whatever individual under one option would be harmed significantly more

than any individual who would be harmed under any other possible option.

Obviously the concept of harm is central to Regan's view — it plays a prominent role both in his general characterization of "having moral rights" and in his two principles regarding conflicting rights. We saw in the preceding section that the classical utilitarians were hedonists, assessing harm in terms of felt pain and lost opportunities for pleasure. Regan endorses a very different, nonhedonistic view of harm, one that many interpreters of Mill have thought was implicitly at work in his thinking, and one that Peter Singer endorses with regard to at least humans and other mammals. On this alternative view, an individual is harmed to the extent that he or she fails to achieve an integrated satisfaction of his or her preferences. A lot is built into the notion of "an integrated satisfaction" of preferences, including, usually, a qualification about weeding out irrational or imprudent preferences. Without going into further detail, however, we can see how such a nonhedonistic conception of happiness would be attractive, at least when it comes to humans. For the death of a normal adult human in the prime of his or her life seems tragic in a way that is difficult to capture from the hedonistic perspective, which would assess the harm purely in terms of any pain felt while dying plus lost opportunities for pleasure in the future. But when a normal adult human dies prematurely, a whole network of preferences for the future — all of one's plans, projects, hopes, dreams, wishes, and so on — go unsatisfied whether one dies painlessly or not, and it is something about this impact on one's preferences that makes the human's death tragic.

Both Regan and Singer hold that this nonhedonistic conception of harm is applicable to many nonhuman animals, including at least all normal mammals. This is one of the important qualifications on Singer's view, alluded to above. Although Singer is a hedonistic utilitarian when it comes to animals who do not have a robust sense of their future, he is a *preference* utilitarian with regard to those who do, including but not limited to human beings. Singer's own view thus differs in two important respects from self-professed animal welfarists. First, he endorses a nonhedonistic conception of happiness for some animals (mammals) in addition to humans. But second, he is a thorough-going utilitarian, unlike those animal welfarists who would attribute rights to at least human beings Singer (24) denies that even humans hold "trump cards" against utilitarian arguments. So it is only the hedonistic utilitarian portion of Singer's thinking about animals which strictly corresponds to the standard philosophical conception of animal welfare.

Assessing Somacloning: Health Risks to Somaclones

Media reports on the early results of somacloning emphasized the low success rate and possible health risks to the resulting clones. Assessing the low success rates raises a complex philosophical question that will be discussed in the next subsection. This subsection describes the differing ways animal welfare and animal rights views would treat the potential health risks to somaclones after birth. First, however, it is necessary to recall the deep scientific uncertainties which still remain on the topic at the time of this writing.

Early work with genetic engineering or gene splicing produced some dramatically deformed animals. The so-called Beltsville pigs, who had a human growth hormone gene inserted into their DNA, suffered from numerous debilitating health problems (25). Because it is still not understood how DNA is re-programmed during the somacloning process, it is natural to worry that similar problems might result. For just as gene splicing alters a complex and well-evolved genetic code, so too might somacloning, if the DNA re-programming process was somehow incomplete or haphazard. Also, because mitochondrial DNA may play a role in gene activation and the enucleated egg cell into which the donor nucleus is transferred contains many mitochondria in its cytoplasm, somacloning by nuclear transfer might cause abnormalities if the two sources of DNA are somehow incompatible.

The number of somaclones produced to date is so small that it is hard to evaluate these risks. By January 2000 only a few dozen somaclones, most of them mice, are reported in the literature. A few health problems in somaclones are reported (6), but Dolly is reported to be healthy and to have delivered two healthy lambs (10) and second generation somacloned mice (somaclones of somaclones) are reported to have reared healthy offspring (5).

Probably the most widely discussed health risk to somaclones involves premature aging due to telomere shortening. At something over one year of age, Dolly's telomeres were found to be significantly shorter than age-matched ewes who had also been produced by nuclear transfer, but using embryonic rather than adult cells. However, "considering the large size distribution of sheep TRFs [terminal restriction fragments or telomeres], it remains to be seen whether a critical length will be reached during the animal's lifetime" (10, p. 317). So it may be that telomere shortening will only affect clones of clones, or perhaps only clones made from already elderly animals. In these respects the Wikayama team's multigeneration mice and Second Chance (the clone produced from an elderly steer at Texas A&M) are of particular interest, but it is just too soon to tell how important telomere shortening will be from published reports available at the time of this writing (5,7). Moreover, even if telomere shortening turns out to be a significant problem, it may be treatable, since in adult mammals and birds, sperm and eggs are produced with fully "repaired" telomeres, and the reagent involved, telemorase, has already been identified.

All in all, then, the magnitude and nature of health risks to somaclones are unclear. There is reason to be cautiously optimistic that somaclones will be of average health, but it is still quite possible that significant health problems will be found. In light of that uncertainty, what can be said about animal welfare and animal rights perspectives on somacloning research?

From the animal welfare perspective, any suffering that results from health problems counts as a negative in the moral ledger. In utilitarian terms, these costs must be weighed against the likely benefits in assessing

the research in question. The complex problem of assessing probabilities cannot be addressed here; in particular, whether or not it is even possible to rate experimental protocols in terms of their contributing to the development of valuable clinical applications is a very controversial issue. However, as a general line of research, somacloning seems likely to have a large range of valuable applications which a utilitarian would have to balance against whatever suffering somaclones endure, and if these benefits are very significant, then they could easily outweigh the still unsubstantiated risks somaclones face, at least from a utilitarian perspective.

How many and what kind of applications somacloning will have depends largely on what technical hurdles scientists are able to clear and on some related policy issues. It is all but certain that adult humans can be cloned by the nuclear transfer technique, paving the way for at least certain new reproductive technologies. With somacloning succeeding in mice, sheep, and cattle, whether and when these benefits of somacloning are realized is a matter of political rather than scientific uncertainty.

Other benefits of somacloning will only be realized if some significant remaining technical problems are solved. For instance, medical researchers envision cloning tissue from adults to produce rejection-free grafts of homogenous tissues like skin, bone, and liver. Such tissues can already be grown in culture, but since no one yet understands how DNA reprogramming occurs when embryos are manipulated after nuclear transfer, no one can yet say how difficult it will be to control the process so that instead of producing a whole new individual, an embryonic somaclone can be made to produce just one specific type of tissue cells. More ambitious treatments in the same vein would involve producing whole organs in vitro for rejection-free transplant into the DNA donor, but this vision of "organs in vats" still seems far-fetched.

Still other uses of somacloning are proposed besides such clinical applications in human medicine. The Missyplicity Project (26) is funded by a wealthy couple seeking to clone their pet dog, Missy, but the project has collateral goals with far-reaching implications (9). The project aims to establish commercial dog-cloning services which could be used to reproduce the best "service" dogs — guide, drug sniffing, and search and rescue dogs — whose largely genetically determined abilities can only be discovered through expensive training processes. The project is also intended to make the technology available free of charge to programs benefiting endangered canid species. In the latter case, somacloning could be used to preserve individual genotypes that would otherwise be lost to old age or when unsuccessfully released into the wild, and to introduce genotypes from the wild into captive breeding populations without removing the donor individuals from the wild.

With somacloning promising this diverse array of significant benefits, the health risks to somaclones would have to be very significant in order to outweigh the benefits in a utilitarian calculus. And when we recall that the standard animal welfare perspective assumes a hedonistic conception of happiness (at least where nonhuman animals are concerned), it seems even less likely that somacloning could be condemned from that perspective. Suppose, for instance, that telomere shortening necessarily reduces a somaclone's life span by a significant amount, but without causing it any kind of suffering apart from that involved in the normal aging process. Because the animal welfare perspective evaluates premature death purely in terms of lost opportunities for pleasure, it is hard to see this as tragic, especially since the somaclone itself would never have existed at all without this shortened life span. When the question is whether or not to bring into the world an individual with deformities or a disease, a utilitarian well might oppose doing so because there is the possibility of instead bringing into existence a similar individual without the deformity or disease. However, when the question is whether or not to go forward with research on somacloning, if some health problems for somaclones are inevitable, then there is no way to achieve the result in question without the attendant costs, and in light of the above consideration of potential uses of the technology, the result in question appears very important from a utilitarian perspective. So unless health problems facing somaclones cause fairly significant amounts of suffering, it is hard to see how this research could be opposed from an animal welfare perspective.

From an animal rights perspective, however, the situation is very different, for two reasons. First, recall that Regan's paradigm animal rights view assesses harm to individuals in terms of lost opportunities to form and satisfy desires. This means that for Regan, even a painless death can be a significant harm to an individual. Probably it is true, as Regan himself says, that mice, sheep, cattle, and even dogs have a significantly less robust sense of their own futures than do humans, but still the same medical problems are likely to translate into a different kind and degree of harm done to an animal when viewed from a perspective like Regan's rather than from the purely hedonistic perspective of animal welfare.

The second reason animal rights philosophies are likely to condemn somacloning research is that rights views are essentially antiutilitarian. Remember that Regan's work explores the implications of extending the kind of respect commonly attributed to human individuals — rights as "trump cards" against utilitarian arguments for harming them — to animals, and this involves denying that harm to the individual can be justified by appealing to the aggregate benefits of doing so. The array of benefits likely to emerge from cloning research thus appears *irrelevant* from an animal rights perspective. If the individuals involved are harmed and they have not consented (or cannot consent), then the research violates their rights, whether the individuals in question are humans or animals, and whatever the magnitude of the harms in question.

One obvious objection to Regan's blanket opposition to animal experimentation is suggested by what was said above while summarizing his view. Regan endorses two nonutilitarian principles for deciding whose rights to violate when rights violations are inevitable, and one of these, the worse-off principle, arguably implies that some animal experimentation would be not only

permissible but morally mandatory. For instance, suppose (not implausibly) that it would be possible to save more human burn victims in the future if somacloned skin grafts are developed. On Regan's view, the harm that death is to an individual is a function of how much capacity for desire formation and satisfaction that individual has, and Regan admits that this makes death to a human being (at least in the prime of his or her life) a significantly greater harm than death to any nonhuman animal, whose sense of its own future is much less robust. But then wouldn't the worse-off principle imply that the research ought to be done, since the future burn victims who would die without it would be harmed more than any nonhuman animal involved in the cloning research?

Regan responds to this sort of argument in *The Case for Animal Rights* (18, pp. 363–394) by claiming, in effect, that the worse-off principle only applies where background conditions of justice or fairness have been maintained, and that because the research involuntarily imposes risks that the experimental subjects would not otherwise face, that is not the case. That is, if either a human or an animal were going to die no matter what anyone did, then, Regan claims, the appropriate way to decide who should die would be using the worse-off principle — indeed, to decide on any other basis would be to fail to respect the two individuals equally. However, in the case of biomedical research, the risks of disease and death faced by the experimental animals have been imposed upon them without their consent. So, Regan argues, it seriously misrepresents the situation to paint it as a choice between saving one or the other life one of those lives is in jeopardy only because researchers have unfairly chosen to put it there, without the animal's consent.

Regan's response may encounter a problem when applied to somacloning research specifically (and some kinds of genetic engineering), however. Most animals involved in biomedical research only face special health risks if and when researchers choose to expose them to disease, injury, or drugs. Yet if somacloned animals necessarily face special risks, such as shortened life spans, then it is simply impossible for the individual somaclones to exist without facing that risk. And, if it is impossible for the individual in question to have existed without facing the risk in question, does it make sense to say that the individual is harmed by being brought into existence facing that risk? This raises complex philosophical issues in personal identity (some of which are touched on in Brock's contribution to this volume).

Assessing Somacloning: The Question of Potentiality

In addition to potential health risks to somaclones, media reports on early work in somacloning stressed that the success rate was abysmally low. For instance, Wilmut et al. (2) did 277 nuclear transfers to get the one pregnancy that resulted in Dolly. However, in assessing animal rights and animal welfare views on somacloning, it is important to note how many embryos failed to result in live births and for what reasons and the related question of these entities' moral status.

Most of the eggs into which donor nuclei were transferred by early researchers were disposed of without any attempt to implant them in surrogate mothers. For instance, Wilmut et al. only attempted 29 implantations, and from these attempts only one pregnancy resulted. However, at the same time they were trying to produce Dolly, the Wilmut team was using the nuclear transfer technique with embryo cells as the source of donor nuclei, and in this portion of their work, they apparently had several miscarriages and one neonatal death. They report attempting 135 implantations resulting in only 21 pregnancies, based on ultrasonic scans performed about 8 weeks after attempted implantation (around the end of the first trimester in a sheep's pregnancy). Subsequent scans revealed fewer pregnancies remaining, "suggesting either misdiagnosis or foetal loss," and in the end, only 8 live births resulted. Of these, one died within minutes of birth, although "post-mortem analysis failed to find any abnormality or infection" (2, p. 811). So apart from whether the source of the DNA is a somatic cell or a germ cell, the nuclear transfer technique may pose significant risks to the resulting embryos, fetuses, and neonates.

Although some researchers soon began to achieve higher success rates, what is the moral status of the many, many blastocysts disposed of prior to implantation, the spontaneously aborted embryos, and the neonates lost in research like the Wilmut team's? That depends crucially on a philosophical point about the moral significance of merely potential (as yet unactualized) traits or capacities: *If an individual has the potential to develop a trait or capacity that would make it deserving of some kind of respect, or moral rights, does that individual deserve the same kind of respect or rights now?* This we may call the question of potentiality.

Many people say yes to this question, at least when it comes to human beings. This is a secular, philosophical basis for a pro-life position on abortion. Just as many say no, however, and this seems to be a necessary plank in any pro-choice platform. For if we answer yes to the question of potentiality, how could we deny that every human embryo deserves the same respect or rights as any adult, and how could we fail to reach the conclusion that abortion is murder? Indeed, since even a newly fertilized egg has the potential to become a fully functioning adult, understanding the implications of saying yes to the question of potentiality helps us understand why many pro-life advocates could equate several forms of birth control with abortion (and, in turn, murder). For IUDs and "morning after" treatments (like RU 486 and "emergency contraception" with high doses of birth control pills) both function by preventing implantation in the uterus rather than by preventing conception, and even standard birth control pills have this as a side effect and second line of defense.

People often wonder how an animal rights advocate could consistently be pro-choice on the abortion issue, but if one answers no to the question of potentiality, then the moral status of the developing human being changes dramatically from conception through birth. From the animal welfare perspective, sentience (the ability to feel pain) is the criterion for moral standing, and on Regan's rights view, it is being a "subject of a life," by which he means having memories, desires, and a sense of one's own future. (As was noted earlier, Singer thinks that sentience

is sufficient for moral standing but also recognizes that having a robust sense of one's future gives one's life some added moral significance.) If sentience is the criterion for moral standing, then developing embryos probably do not qualify at all, since the first brain waves are detectable only at around six weeks and consciousness of pain may not develop until much later. And if having a sense of one's future is the criterion for having moral rights, humans may not qualify until sometime after birth. It is for this very reason that Regan restricts his claim about rights possession to "mentally normal mammals of a year or more" (18, p. 78). He does not deny that neonates have a sense of their future, but he admits this is more controversial than the claim that one-year-olds do.

Similar things can of course be said about developing animals. Although newborn animals are usually more precocious than are newborn humans, presumably the numerous animal embryos discarded before implantation during somacloning research, as well as at least early-term fetuses, are incapable of either feeling pain or thinking about their futures. When in an individual's development these capacities arise is a complex issue involving both scientific and philosophical claims that cannot adequately be addressed here (for more on the issue, see Refs. 12, 27, and 29). However, it is clear how, if we answer no to the question of potentiality, neither an animal welfare nor an animal rights position need find any risks faced by embryos and early fetuses in somacloning research morally problematic, leaving the emphasis on the kind of health risks discussed in the preceding subsection.

Apart from relying on one's antecedent intuitions about the moral status of embryos and newly fertilized eggs, how might one argue for a negative answer to the question of potentiality? A classic argument in favor of a negative answer was offered by S.I. Benn and Joel Feinberg in a pair of classic essays on the general concept of moral rights (29,30). They compare moral rights to legal rights, noting that a potential president does not have the right to command the armed forces until he or she actually becomes president. If moral rights and the more general concept of moral standing (which a utilitarian or animal welfarist would use) are treated analogously, then the potential to develop sentience or become a "subject of a life" does not give the individual the same moral status as actually developing these capacities.

A second classic argument offered for a negative answer is based on the phenomena of twinning and chimeras. At conception, a unique genotype is created, but there is more to personal identity than one's genotype. In nature, genetically identical twins result from fission of early embryos, yet in humans and other conscious animals, we consider identical twins distinct individuals in virtue of their differing psychological states. (Although this may be false when it comes to clonal reproduction in plants and even in animals which are not conscious; see Ref. 12.) The reverse phenomenon, when two early embryos with distinct genetic codes fuse, is also known to happen, albeit very infrequently, in nature. The resulting individual is called a chimera. Although every cell in a chimera's body has a genetic code identical to that of one of the two embryos from which it was produced, in any region of the chimera's body, 50 percent of the cells have the code of one embryo and 50 percent the other embryo's. Humans have been born cross-sexual chimeras, and in the lab, chimeras have been produced from closely related species (e.g., by fusing the embryos of a goat and a sheep). In an essay defending experimentation on early human embryos, Helga Kuhse and Peter Singer (31) claim that the above phenomena show that very early embryos are not properly regarded as specific individuals at all: Identical twins cannot both be identical with the single early embryo from which they sprang, nor can a single chimera be identical with the two early embryos from which it sprang.

The foregoing argument appears flawed, however. David Oderberg counters it with the following thought experiment about a "split brain operation." In the philosophical literature on personal identity, this refers to a hypothetical process by which one brain is split into two, which carry (to begin with) identical memories, personalities, and so on. Oderberg writes:

> Suppose it is certain that in five minutes Jones will undergo a split-brain operation. If Jones, being a person, has moral rights, then he is no less a person, and no less a possessor of moral rights, because of this certainty. . . .Similarly, the possibility, even the *certainty*, of division of a human embryo does not *of itself* show that the embryo is fair experimental game (32, pp. 277–278) .

So even if the correct inference to draw regarding twins and chimeras is that the early embryos involved in producing them are not identical with the resulting adults, that would not, by itself, prove that early embryos have no moral standing or moral rights. At most, it would prove that those early embryos are not the same individuals as either the twins or the chimera.

CONCLUSION

Media coverage of early work on somacloning emphasized health risks to the resulting clones both before and after birth. From an animal welfare stance, the benefits promised by somacloning seem likely to outweigh any presently foreseeable postpartum harms to the clones. From an animal rights perspective, the magnitude of these promised benefits is irrelevant — any significant harm inflicted upon the clones without their consent violates their rights, although a philosophical problem about identity may short-circuit that objection. From both perspectives, the question of potentiality determines the (in)significance of risks to all embryos and at least early-term fetuses.

BIBLIOGRAPHY

1. J.B. Gurdon, R.A. Laskey, and O.R. Reeves, *J. Embryol. Exp. Morphol.* **34**, 93–112 (1975).
2. I. Wilmut et al., *Nature* **385**, 810–813 (1997).
3. D. Ashworth et al., *Nature* **394**, 329 (1998).
4. E. Singer, Y. Dubrova, and A. Jeffreys, *Nature* **394**, 229–230 (1998).

5. T. Wakayama et al., *Nature* **394**, 369–374 (1998).
6. J.-P. Renard et al., *Lancet* **353**, 1489–1491 (1999).
7. J. Hill et al., *Biol. Reprod.* **62**, 1135–1140 (2000).
8. N. Wade, *N.Y. Times*, January 5, A14 (2000).
9. G. Varner, *Animal Welfare* **8**, 407–420 (1999).
10. P.G. Shiels et al., *Nature* **399**, 316–317 (1999).
11. H. Rolston, III, *Environ. Ethics* **1**(1), 7–30 (1979).
12. G. Varner, *In Nature's Interests? Interests, Animal Rights and Environmental Ethics*, Oxford University Press, Oxford, UK, 1998.
13. R. Kraut, *Can. J. Philos.* **9**(3), 467–478 (1979).
14. Aristotle, *Nicomachean Ethics*, T. Irwin, trans., Hackett Publishing Company, Indianapolis, IN, 1985.
15. R. Chadwick, *Philosophy* **57**, 201–209 (1982).
16. Plato, in E. Hamilton and H. Cairns, eds., *Plato: The Collected Dialogues*, L. Cooper, trans., Princeton University Press, Princeton, NJ, 1961, pp. 169–185.
17. L. Silver, *Cambridge Q. Healthcare Ethics* **7**, 168–172 (1998).
18. T. Regan, *The Case for Animal Rights*, University of California Press, Berkeley, CA, 1983.
19. P. Singer, *Animal Liberation*, 2nd ed., Avon Books, New York, 1990.
20. G. Varner, *Hastings Cent. Rep.* **24**(1), 24–28 (1994).
21. J. Jasper and D. Nelkin, *The Animal Rights Crusade: The Growth of a Moral Protest*, Free Press, New York, 1992.
22. J. Bentham, An Introduction to the Principles of Morals and Legislation, Basil Blackwell, Oxford, UK, 1948. Originally published in 1789.
23. J.S. Mill, *Utilitarianism*, Bobbs-Merril, Indianapolis, 1957. Originally published in 1861.
24. P. Singer, *Practical Ethics*, 2nd ed., Cambridge University Press, Cambridge, UK, 1993.
25. D. Bolt et al., in *Veterinary Perspectives on Genetically Engineered Animals*, 58-61, American Veterinary Medical Association, Schaumburg, IL, 1990.
26. L. Hawthorne, Missyplicity Project, Available at: *www.missyplicity.com*
27. D. DeGrazia, *Taking Animals Seriously*, Cambridge University Press, Cambridge, UK, 1996.
28. M. Tooley, *Abortion and Infanticide*, Clarendon Press, Oxford, UK, 1983.
29. S.I. Benn, in J. Feinberg, ed., *The Problem of Abortion*, Wadsworth, Belmont, CA, 1984, pp. 145–150.
30. J. Feinberg, in J. Feinberg, ed., *The Problem of Abortion*, Wadsworth, Belmont, CA, 1984, pp. 145–150.
31. H. Kuhse and P. Singer, in P. Singer et al., eds., *Embryo Experimentation*, Cambridge University Press, Cambridge, UK, 1990, pp. 65–75.
32. D.S. Oderberg, *Philos. Public Affairs* **26**, 259–298 (1997).

See other CLONING entries.

CLONING, OVERVIEW OF HUMAN CLONING

KATHI E. HANNA
Science and Health Policy Consultant
Prince Frederick, Maryland

OUTLINE

INTRODUCTION

The term "cloning" is used by scientists to describe many different processes that involve making duplicates of biological material. In most cases isolated genes or cells are duplicated for scientific study, and no new animal results. The experiment that led to the cloning of Dolly the sheep in 1997 was different: It used a cloning technique called *somatic cell nuclear transfer* and resulted in an animal that was a genetic twin — although delayed in time — of an adult sheep. This technique of transferring a nucleus from a somatic cell into an egg that produced Dolly was an extension of experiments that had been ongoing for over 40 years. It built on previous work in which embryos were cloned by a technique called *blastomere separation*, or twinning, and by a method of extracting the nucleus of an embryo and transferring it into another enucleated egg. Although the birth of Dolly was lauded as a success, in fact the procedure is not perfected. It is not yet clear whether Dolly will remain healthy or whether she could have subtle problems that might lead to serious diseases. Therefore the prospect of proceeding to application of this technique in humans is troubling for scientific and safety reasons as well as for additional ethical reasons having to do with family and the order of generations.

DEFINITION OF CLONING

The word "clone" is used in many different contexts in biological research but in its most simple and strict sense it refers to a precise genetic copy of a molecule, cell, plant, nonhuman animal, or human being. The feasibility of

cloning varies according to the complexity of the genetic material being used and its status in the plant or animal kingdom.

Genetically identical copies of whole organisms are commonplace in the plant-breeding world (referred to as *varieties* rather than clones) because it is relatively easy to regenerate a complete plant from a small cutting. However, the developmental process in animals does not usually permit cloning as easily as in plants, except in the simplest of invertebrate species. Although a single adult vertebrate cannot naturally generate another whole organism, cloning of vertebrates does occur in nature primarily with the formation of identical twins. Twins occur by chance in humans and other mammals when a single embryo splits into halves at an early stage of development.

At the molecular and cellular level, scientists have been cloning human and animal cells and genes for several decades. Such cloning provides greater quantities of identical cells or genes for study, as each cell or molecule is identical to the others. At the most basic level, biologists routinely make clones of deoxyribonucleic acid (DNA), the molecular basis of genes. DNA fragments containing genes are copied and amplified in a host cell, usually a bacterium. The availability of large quantities of identical DNA makes possible many scientific experiments. This process, often called *molecular cloning*, is the mainstay of recombinant DNA technology and has led to the production of important therapies, such as insulin to treat diabetes, tissue plasminogen activator (tPA) to dissolve clots after a heart attack, and erythropoietin (EPO) to treat anemia associated with dialysis for kidney disease.

In *cellular cloning* copies are made of cells derived from the soma, or body, by growing these cells in culture in a laboratory. The genetic makeup of the resulting cloned cells, called a *cell line*, is identical to that of the original cell. Since molecular and cellular cloning of this sort does not involve germ cells (eggs or sperm), the cloned cells are not capable of developing into a baby.

The third level of cloning aims to reproduce genetically identical animals. Cloning of animals can typically be divided into three distinct techniques, embryo cloning by *blastomere separation* (twinning), embryo cloning by *nuclear transplantation*, and cloning via *somatic cell nuclear transfer*. It is this last type of cloning that has raised so many legal, social, and ethical concerns, particularly after the February 1997 announcement that Ian Wilmut and colleagues had successfully cloned a sheep using somatic cell nuclear transfer (1).

CLONING BY BLASTOMERE SEPARATION (TWINNING)

Blastomere separation is an increasingly common form of cloning used in animal biotechnology. In this technique, the developing embryo is split very soon after fertilization when it is composed of two to eight cells (see Fig. 1) (2). To split the embryo, technicians dissolve the protective covering, or zone pellucida, of an early cleaving embryo and place the embryos in a medium in which they separate into individual cells. The cells, called *blastomeres*, are then placed in another solution that forms an artificial zone and are moved to a culturing medium, where they will begin cleaving (dividing). Each blastomere is able to produce an entirely new, individual organism. This is because blastomeres are considered to be totipotent; that is, they possess the total potential to make an entire new organism. This totipotency allows scientists to split animal embryos into several cells to produce multiple organisms that are genetically identical. This capability has tremendous relevance to breeding cattle and other livestock although it is an expensive technique and it is not yet perfected, having been associated with higher rates of congenital malformations.

Researchers first used blastomere separation to twin sheep embryos in 1979 and cattle embryos in 1980 (3). It is the sole cloning technique to have been attempted experimentally with human embryos. In 1993 investigators at the George Washington University (GWU) Medical Center separated the cells of 17 human embryos and generated 48 embryos (an average of three embryos for each original), demonstrating that one human embryo could be split to create two or more genetically identical embryos (4). In the GWU study the investigators used 22 eggs that were fertilized by more than one spermatozoan during in vitro fertilization (IVF) and were thereby considered ineligible for implantation; that is, they would have been

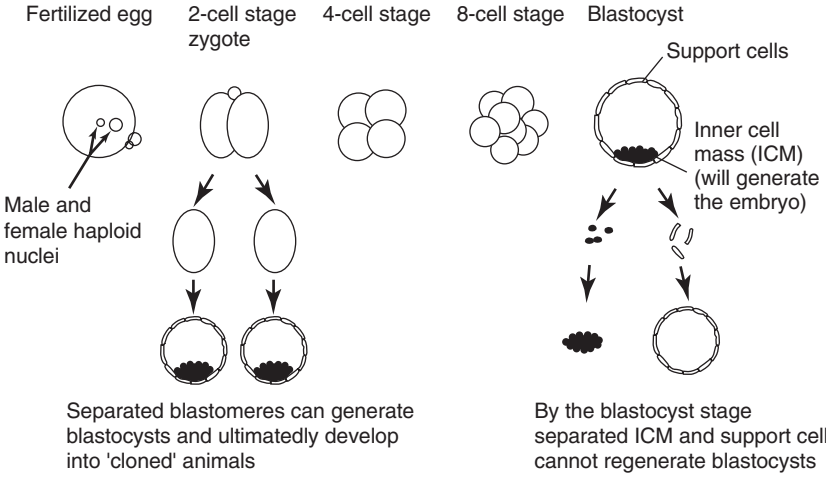

Figure 1. Preimplantation embryo development in mammals (*Source*: Ref. 21).

discarded (5). The researchers allowed the polyspermic eggs to divide into two, four, or eight cells. Cells from 17 of the embryos were placed in another solution that formed an artificial zone and were moved to a culturing medium. An average of three embryos developed from each original embryo, leading to a total of 48 cleaving cells. The GWU experiment showed that twinning is technically feasible, but it created an ethical firestorm and reopened debate about the appropriateness of research conducted on the preimplantation (ex utero) human embryo (6).

Proponents of blastomere separation justify its use as an aid in IVF. If too few embryos are available for transfer to the uterus, IVF success rates are low. Embryo twinning could increase the number of embryos and do so by manipulating (i.e., splitting) the embryos, rather than administering powerful and risky fertility drugs to women in order to stimulate the release of more eggs for fertilization. In addition, blastomere separation could facilitate preimplantation diagnosis, a technique in which a cell is removed from each embryo of a couple at risk for passing a deleterious gene to their offspring. The DNA from the extracted cell is amplified and examined for the presence of the disease-related gene. Only embryos lacking the deleterious gene are transferred. Preimplantation diagnosis remains problematic because in 30 percent of attempts DNA amplification fails due to insufficient quantities of DNA. With blastomere separation more than one cell can be used, which could improve the success rate of amplification and improve the accuracy of diagnoses (7).

Blastomere separation is also proposed as an aid to scientific inquiry. In research studies using human embryos this technique could provide a control embryo for each experimental embryo in order to hold constant as many variables as possible and state with more confidence that differences in the experimental and control groups were due to manipulation rather than to genetic differences in the two groups (8).

Ethical Considerations

One of the most cited objections to the use of twinning in humans is the questionable effect it might have on children. In one application of this technique multiple, identical embryos could be produced and all transferred at the same time to the prospective mother's uterus. If all implanted, the woman would give birth to identical twins, triplets, quadruplets, or more. Aside from the possible hardship this might cause the family, it is unlikely that the children would suffer psychologically as a result. However, if not all embryos were transferred at the same time, unique concerns arise. If the first transfer failed, a second frozen embryo could be thawed and transferred. If the transfer was successful, few would express concern, because only one live born child was produced as a result of multiple, although spaced, transfers.

An issue that has followed IVF since its first use concerns the disposition of frozen embryos. Whether embryos are identical, via blastomere separation or not couples can transfer them at a later date. If the embryos are identical, however, the result would be spaced twins. No one knows what it would be like to be the older or younger of spaced twins but because of the influence of social rearing it is highly unlikely that they would be identical in any way other than genotype. Rather, they would share many of the same environmental influences experienced by children reared in the same household. However, the temptation might exist for parents to unfairly compare the children. Thus concerns about the spacing of identical twins raises questions about the potential adverse psychosocial effects that might occur for the second twin because of unrealistic and unfounded expectations on the part of the parents.

Perhaps of more concern is the possibility that parents would store embryos in the event of the death of a child: The newly transferred identical would provide them with a genetic replica of the deceased child. Also a concern is the possibility that the stored embryos might later be used to generate cell lines for therapeutic transplantation for the living child who has become ill. Even more far-fetched, but possible, would be transgenerational transfer. That is, frozen embryos could be bequeathed to future generations, allowing a woman to give birth to her identical twin or a great grandchild to give birth to the identical twin of her great grandfather.

EMBRYO CLONING BY NUCLEAR TRANSPLANTATION

A more complex cloning technique is the nuclear transplantation of embryo cells to enucleated eggs. This technique was first used in animals in the early 1980s and the first birth in higher animals (lambs) occurred in 1986 when scientists transplanted the nuclei of cells obtained from four- to eight-cell embryos to enucleated eggs (9). The first experiments of this type were successful only when the donor cell was derived from an early embryo. Later, in 1996 and 1997, lambs were produced in this manner, this time using the nuclei of cells from a late stage embryo (1). In 1997 the technique was successfully used in Rhesus monkeys (10). In theory, large numbers of genetically identical animals could be produced through such nuclear transplantation cloning. In practice, the nuclei from embryos which have developed beyond a certain number of cells seem to lose their totipotency, limiting the number of animals that can be produced in a given period of time from a single, originating embryo.

The safety of this technique is questionable. It has not been attempted in humans. Although apparently healthy offspring have been produced in mammals, there is a high rate of fetal and infant loss. In the 1996 experiment with sheep, of the eight fetuses two were spontaneously aborted and two died at birth. Another died at 10 days (11). Other investigators have reported a high rate of congenital abnormalities (20 to 30 percent and larger than normal offspring) (12). The effects of the transfer of genetic material from one organism to another might include a compromised immune system.

Finally, more has to be learned about the role of mitochondrial DNA in development. A cellular component called the *mitochondrion* is the energy-producing component of the cell. Although most of the genes associated with the mitochondrion reside in the nucleus, the mitochondrion itself houses some genes. Thus, in nuclear transfer, mitochondrial genes are not transferred to the

enucleated egg along with the nuclear genes. Much has to be learned about the compatibility of nuclear DNA and mitochondrial DNA when such a transfer occurs (13). Preliminary studies of the effects of genetic imprinting suggest that mitochondrial DNA can play an important role in the early activation and function of some genes; 37 genes responsible for energy metabolism in cells are known to be coded by mitochondrial DNA (14). Thus mitochondrial DNA likely plays an important role in some disease expression, a role that must be considered when designing somatic cell nuclear transfer.

Ethical Considerations

Many of the issues raised by blastomere separation also apply to embryo cloning by nuclear transplantation, with some subtle differences. First, embryo cloning by nuclear transplantation would permit the creation of more embryos because it can take place at a later stage of embryonic development (8- to 16-cell stage). Thus many of the concerns posed by twinning are magnified by the potential to create more embryos through nuclear transplantation. Second, the embryos would not be as genetically identical in nuclear transplantation because the nuclei would be fused with the cytoplasm (and therefore the mitochondrial DNA) of the egg donor(s). This might diffuse some of the concern about loss of individuality. However, the safety issues appear to be the most pressing of concerns and perhaps the greatest ethical obstacle to proceeding in this technique.

SOMATIC CELL NUCLEAR TRANSFER CLONING

In the two previous examples of cloning, the genetic material being used in cloning was genetically unique, that is it was the product of a fusion between an egg and a sperm. The manipulation used to create clones occurred after the fusion. Thus a new individual was being replicated several times. What made Dolly different is that she was not genetically unique, that is, she was genetically identical to an existing six-year-old ewe. She was not created by the union of egg and sperm (see Fig. 2). To appreciate how her creation was possible, it is helpful to understand the science that led up to her birth. Much of it centers on an evolving understanding of how cells develop and differentiate, and how genes are expressed.

Cell Differentiation

Nearly every cell contains a spheroid organelle called the *nucleus*, which houses nearly all the genes of the organism. Genes are composed of DNA, which serve as a set of instructions to the cell to produce particular proteins. Although all somatic cells of an individual contain the same genes in the nucleus, the particular genes that are activated vary by the type of cell. For example, a differentiated somatic cell, such as a neuron, must keep a set of neural-specific genes active and silence those genes specific to the development and functioning of other types of cells such as muscle or liver cells. In contrast, gametes (eggs and sperm) do not differentiate but retain activity necessary to create new life after fusion with egg or sperm.

Investigations that began over 40 years ago sought to determine whether a differentiated somatic cell still contained all genes, even those it did not express. Early experiments in frogs and toads (15,16) provided evidence that the expression potential of the genes in differentiated cells is essentially unchanged from that of the early embryo. Nuclei from donor-differentiated cells were injected into recipient eggs in which the nucleus had been inactivated (Fig. 3). Another carefully controlled

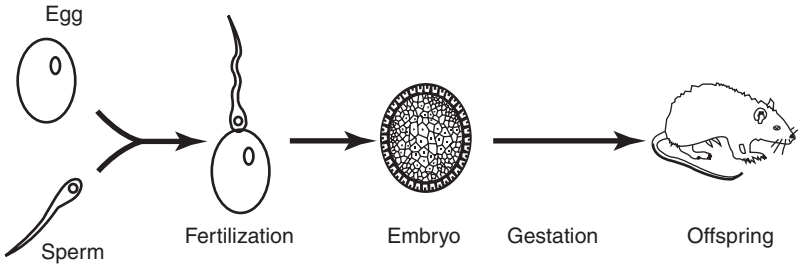

Figure 2. Sexual reproduction (*Source*: Ref. 21).

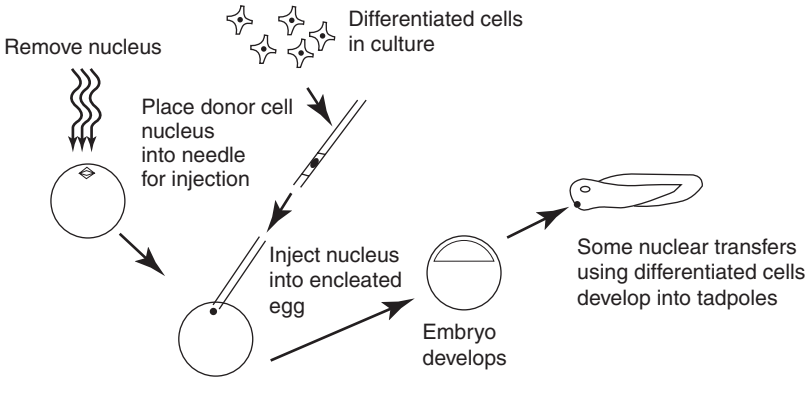

Figure 3. Nuclear transfer carried out in frogs (*Source*: Ref. 21).

series of experiments used nuclei from adult frog skin cells for transfer to an enucleated egg (17). Four percent of the nuclei transferred eventually gave rise to fully developed tadpoles. These experiments provided evidence that the genes contained in the nuclei of differentiated cells could be reactivated by the cytoplasm of the egg and thus direct normal development, but only up to a certain stage. No viable adult frog ever developed from these tadpoles and there was a decrease in the number of tadpoles born as the age of the transferred nucleus increased. This left open the possibility that complete reactivation of the adult nucleus was prevented by some irreversible change in the genetic material, and that there was a progressive decline in nuclear potential with age. Later analysis, however, suggested that the major reason for developmental failure of the transplanted embryos appeared to be chromosomal abnormalities that occurred during the process of nuclear transplantation itself.

Experiments in mammals also suggested that it is possible to reprogram adult somatic cells. Experiments in mice followed those in frogs and toads. Early development occurs at a slower rate in mammals than amphibians; thus it was believed that reprogramming of the donor nucleus would occur more efficiently in mice than in frogs. It was shown in mice in the 1980s that nuclei could be successfully exchanged between fertilized eggs, with 90 percent reaching the blastocyst stage of embryonic development and beyond, but only if the nuclei were recovered from an embryo at the two-cell stage (18). Many experiments since then have shown that blastomeres up to the early blastocyst stage are still totipotent when combined with other embryonic cells (19). This means that the failure of nuclear reprogramming has to be the result of something other than irreversible changes to the genetic material of the cells.

In 1986, experiments in sheep had results that differed from those in mice. Enucleated eggs from sheep could be fused with blastomeres taken from embryos at the eight-cell stage to provide donor nuclei and viable offspring were produced (20). Apparently the use of early-stage eggs prolongs the period of reprogramming before the donor nucleus has to undergo the first division. The advent in the last few years of electrofusion for both fusion of cells and activation of the egg has been another major advance, because activation and fusion occur simultaneously.

Enucleated eggs used for fusion only proceed to division after activation by some artificial signal, such as the electrical current used in the electrofusion technique. Because these experiments used fusion of two cells and not simple injection of an isolated nucleus, all of the cellular components are transferred. Thus the mitochondria, which contain some genes of their own, are transferred along with the nucleus. Because an enucleated egg also contains mitochondria, the result of a fusion experiment is a cell with a mixture of mitochondria from both the donor and the recipient. Since the mitochondrial genes represent an extremely small proportion of the total number of mammalian genes, mixing of mitochondria per se is not expected to have any major effects on the cell (2,21). However, if one of the donors suffers from a mitochondrial disease, then mixture of the mitochondria may significantly alleviate the disease.

Over the past 10 years or so, there have been numerous reports of successful nuclear transfer experiments in mammals, nearly all of them using cells taken directly from early embryos. The oldest embryonic nucleus that can successfully support development differs among species (22–26).

Reprogramming Gene Expression

More recently, studies suggested that it might be possible to reprogram the gene expression of somatic cells so that they perform a different task. Thus, it should be possible to activate or inactivate almost any gene in a cell, given the right cellular environment containing the appropriate regulatory molecules. To reprogram the gene expression of a somatic cell it is not essential to fuse it with an egg; in some cases re-programming can occur through fusion of two adult cells. Cell fusion experiments have demonstrated that extensive reprogramming of differentiated nuclei can occur (27). The knowledge that regulatory molecules can reprogram an adult nucleus led to the speculation that cloning via somatic cells nuclear transfer was a real possibility.

Synchronization of the Cell Division Cycle

All of the work just described struggled with understanding the relationships between the normal cell division cycle, the age of the embryo, and the ability of the nuclei to be reprogrammed (28,29). Work by Wilmut showed that the phase of cell cycle division at which transfer is attempted is critically important. Thus the need to transfer nuclei in a specific phase of division (called the G1 phase) before replication is initiated is important to avoid the chromosome damage that can occur and prevent development of the embryo into a viable offspring (1).

Cloning of Dolly

In the research that preceded the birth of Dolly, Wilmut and colleagues established cell lines from sheep blastocysts and used these cells as nuclear donors (11). In an attempt to avoid the problems of nuclear transfer of non-G1 nuclei into activated eggs, they starved the donor cell line by removing all nutrients from the medium prior to nuclear transfer. Under these starvation conditions, the cells exit the cell cycle and enter another phase (Gap phase 0), which is similar to the G1 phase in which chromosomes have not replicated. For Wilmut and colleagues, approximately 14 percent of fusions resulted in development of blastocysts, and 4 of the 34 (12 percent) embryos transferred developed into live lambs. Two died shortly after birth. The success rate in sheep and cow experiments was almost identical, and it suggests that division of cells in culture for many days does not inhibit the ability of their nuclei to be reprogrammed by the egg environment. These findings led Wilmut and colleagues to ask if the same would be true of nuclei from fully differentiated somatic cells.

In the most simple of terms, the technique used to produce Dolly the sheep, somatic cell nuclear transplantation cloning, involves removing the nucleus of an egg and replacing it with the diploid nucleus of a somatic cell.

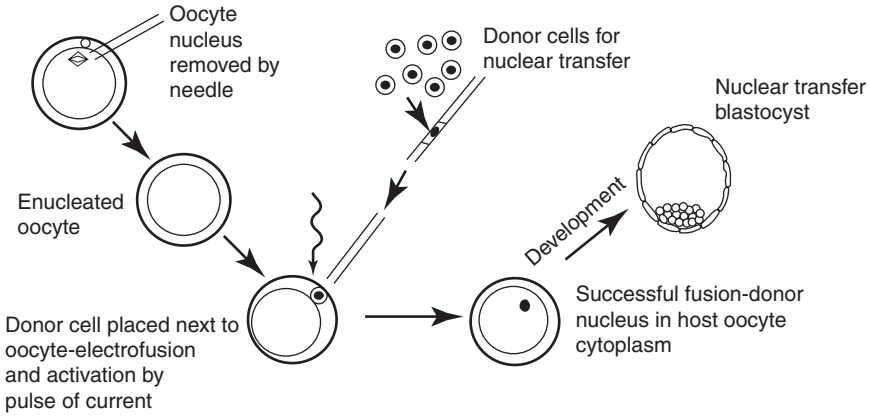

Figure 4. Nuclear transfer via electrofusion in mammals (*Source*: Ref. 21).

In such nuclear transplantation cloning there is a single genetic "parent," unlike sexual reproduction where a new organism is formed when the genetic material of the egg and sperm fuse (see Fig. 4). In addition, this technique differs from blastomere separation and embryo cloning via nuclear transplantation because it does not involve an existing embryo.

In the ground-breaking research reported in February 1997 (1), Wilmut and colleagues took late embryo, fetal cell cultures, and cell cultures derived from the mammary gland of an adult sheep and applied the same approach of synchronizing the cell in the G0 stage prior to nuclear transfer. They reported successful production of live offspring from all three cell types, although only 29 of 277 (11 percent) of successful fusions between adult mammary gland nuclei and enucleated oocytes developed to the blastocyst stage, and only 1 of 29 (3 percent) blastocysts transferred developed into a live lamb. This experiment was, in fact, the first time any fully developed animal had been born following transfer of a somatic cell nucleus.

It should be noted, however, that no attempt was made to document that the donor cells were fully differentiated cells, the genes of which expressed specialized mammary gland proteins. In the earlier experiments with frogs, the fact that the donor cells were fully differentiated was documented in such a manner. In the case of Dolly, it is possible that the cells could have been derived from a less-differentiated cell in the population, such as a mammary stem cell.

Scientific Uncertainties

Several important concerns remain about the science and safety of nuclear transfer cloning using adult cells as the source of nuclei. The first concerns the Dolly experiment itself. Only one animal was produced thus it is not clear that this technique is reproducible even in sheep.

Second, there might be true species differences in the ability to achieve successful nuclear transfer (21). It has been shown that nuclear transfer in mice is much less successful than in larger domestic animals. Part of this difference may reflect the intensity of research in this area in the last 10 years; agricultural interests have meant that more nuclear transfer work has been performed in domestic animals than in mice. But part of the species differences may be real and not simply reflect the greater recent effort in livestock. For example, in order for a differentiated nucleus to redirect development in the environment of the egg, its constellation of regulatory proteins must be replaced by those of the egg in time for the embryo to use the donor nucleus to direct normal development of the embryo. The species difference may be the result of the different times of embryonic gene activation. In mammals, unlike many other species, the early embryo rapidly activates its genes and cannot survive on the components stored in the egg. The time at which embryonic gene activation occurs varies between species (30,31). The later onset of embryonic gene activation and transcription in sheep provides an additional round or two of cell divisions during which nuclear reprogramming can occur, unlike the rapid genome activation in the mouse. Cross-species comparisons are needed to assess the importance of this difference in the time of genome activation for the success of nuclear transfer experiments (21). In humans, for example, the time period before gene activation is very short, which might not permit the proper reprogramming of genes after nuclear transfer to allow for subsequent normal development.

Third, the phenomenon of genetic imprinting may affect the ability of nuclei from later stages to reprogram development. Genetic imprinting refers to the relative effect on embryonic development of genes inherited from the father (paternal genes) versus those from the mother (maternal genes) (32). Some heritable imprint is established on the chromosomes during the development of the egg and the sperm such that certain genes are expressed only when inherited from the father or mother. Nuclei transferred from diploid cells, whether embryonic or adult, should contain maternal *and* paternal copies of the genome, and thus not have an imbalance between the maternally and paternally derived genes. However, an adult nucleus, if it is to be successfully reprogrammed, must retain intact the chromosomal imprints that under normal conditions would determine whether maternal or paternal gene copies will be active. There is some speculation that some instability of the imprint, particularly in cells in culture, could limit the efficiency of nuclear transfer from somatic cells (21). In addition, it is known that disturbances in imprinting lead

to growth abnormalities in mice and are associated with cancer and rare genetic conditions in children.

A fourth concern is whether cellular aging will affect the ability of somatic cell nuclei to program normal development. As somatic cells divide they progressively age and there is normally a defined number of cell divisions that they can undergo before senescence. Part of this aging process involves the progressive shortening of the ends of the chromosomes, the telomeres, and other genetic changes. Germ cells (eggs and sperm) evade telomere shortening by expressing an enzyme, telomerase, that can keep telomeres full length. It seems likely that returning an adult mammalian nucleus to the egg environment will expose it to sufficient telomerase activity to reset telomere length, since oocytes have been found to be potent sources of telomerase activity (33). In 1999 a team of researchers, including those that cloned Dolly, measured the telomeres in her cells as an indication of her actual age and found that her telomeres were 20 percent shorter than would be expected (34). However, it is not yet known whether the shorter telomeres actually make a difference in the physiological age of the cloned sheep. In addition, the 20 percent difference may be within the normal variation for sheep.

The health effects for the resulting liveborn, having been created with an "aged' nucleus, are unknown. Therefore, a fifth concern is raised by the possibility that the mutations that accumulate in somatic cells might affect nuclear transfer efficiency and lead to cancer and other diseases in the offspring. As cells divide and organisms age, mutations in the DNA will inevitably occur and will accumulate with time. If these mistakes occur in the sperm or the egg, the mutation will be inherited in the offspring. Normally mutations that occur in somatic cells affect only that cell and its descendants, which are ultimately dispensable. Nevertheless, such mutations are not necessarily harmless. Transfer of a nucleus from a somatic cell carrying such a mutation to an egg would transform a sporadic somatic mutation into a germline mutation (i.e., transmitted to all of the cells of the body).

POTENTIAL THERAPEUTIC APPLICATIONS OF NUCLEAR TRANSFER CLONING

The demonstration that in mammals as in frogs, the nucleus of a somatic cell can be reprogrammed by the egg environment provides further impetus to studies on how to reactivate embryonic programs of development in adult cells. These studies have exciting prospects for regeneration and repair of diseased or damaged human tissues and organs, and they may provide clues as to how to reprogram adult differentiated cells directly without the need for oocyte fusion. In addition, the use of nuclear transfer has potential application in the field of assisted reproduction.

Potential Applications in Organ and Tissue Transplantation

Some human diseases can be treated effectively by organ or tissue transplantation, including some leukemias, liver failure, and heart and kidney disease. In some instances, the organ required is nonvital, that is, it can be taken from the donor without great risk (e.g., bone marrow, blood, kidney). In other cases, the organ is obviously vital and required for the survival of the individual, such as the heart. All transplantation is imperfect, with the exception of that which occurs between identical twins, because transplantation of organs between individuals requires genetic compatibility.

The application of nuclear transfer cloning to humans could provide a potential source of organs or tissues of a predetermined genetic background. The notion of using human cloning to produce individuals for use solely as organ donors is not only morally repugnant but also illegal as it is unconstitutional and may violate the prohibition against slavery. A morally more acceptable and potentially feasible approach is to direct differentiation along a specific path to produce specific tissues (e.g., muscle or nerve) for therapeutic transplantation rather than to produce an entire individual. Given current uncertainties about the feasibility of this, however, much research would be needed in animal systems before it would be scientifically sound, and therefore potentially morally acceptable, to go forward with this approach.

Potential Applications in Cell-Based Therapies

Another therapeutic possibility raised by cloning is transplantation of cells or tissues not from an individual donor, but from an early embryo or embryonic stem cells. Blastomeres and stem cells are primitive, undifferentiated (or totipotent) cells. This potential application would not require the generation and birth of a cloned individual. Embryonic stem cells provide an interesting model for such studies, since they represent the precursors of all cell lineages in the body. Mouse embryonic stem cells can be stimulated to differentiate in vitro into precursors of the blood, neuronal and muscle cell lineages, among others (35), and they thus provide a potential source of stem cells for regeneration of all tissues of the body.

It might be possible to take a cell from an early blastomere and treat it in such a manner as to direct its differentiation along a specific path. By this procedure it might be possible to generate in the laboratory sufficient numbers of specialized cells, for example, bone marrow stem cells, liver cells, or pancreatic beta-cells (which produce insulin) for transplantation. If even a single tissue type could be generated from early embryonic cells by these methods and used clinically, it would constitute a major advance in transplantation medicine by providing cells that are genetically identical to the recipient.

One could imagine the prospect of nuclear transfer from a somatic cell to generate an early embryo and from it an embryonic stem cell line for each individual human, which would be ideally tissue-matched for later transplant purposes. This might be a rather expensive and far-fetched scenario. An alternative scenario would involve the generation of a few, widely used and well-characterized human embryonic stem cell lines, genetically altered to prevent graft rejection in all possible recipients.

The preceding scenarios depend on using cells of early human embryos, generated either by IVF or nuclear transfer into an egg. Because of ethical and moral concerns raised by the use of embryos for research purposes, it

would be far more desirable to explore the direct use of human cells of adult origin to produce specialized cells or tissues for transplantation into patients (2). It may not be necessary to reprogram terminally differentiated cells but rather to stimulate proliferation and differentiation of the quiescent stem cells, which are known to exist in many adult tissues, including even the nervous system (36). Experiments in this area are likely to focus more on the conditions required for direct stimulation of the stem cells in specific tissues, than on actual use of nuclear transfer to activate novel developmental programs. These approaches to cellular repair using adult stem cells will be greatly aided by an understanding of how stem cells are established during embryogenesis.

Another strategy for cell-based therapies would be to identify methods by which somatic cells could be "de-differentiated" and then re-differentiated along a particular path. This would eliminate the need to use cells obtained from embryos. Such an approach would permit the growth of specialized cells compatible with a specific individual person for transplantation. Although at the current time this strategy is highly speculative, ongoing research in animal systems may identify new approaches or new molecular targets that might make this approach feasible.

It will be of great importance to understand through experiments in animals how the environment of the egg reprograms a somatic cell nucleus. What cellular mechanisms can be elucidated? What components are involved in these processes? Can we direct cells along particular developmental pathways in the laboratory and use these cells for therapy? The capacity to grow human cells of different lineages in culture would also dramatically improve prospects for effective somatic gene therapy.

Assisted Reproduction

Another area of medicine where the knowledge gained from animal work has potential application is in the area of assisted reproduction. Assisted reproduction technologies are already widely used and encompass a variety of parental and biological situations, that is, donor and recipient relationships. In most cases an infertile couple seeks remedy through either artificial insemination or IVF using sperm from either the male or an anonymous donor, an egg from the woman or a donor, and in some cases surrogacy. In those instances where both individuals of a couple are infertile or the prospective father has nonfunctional sperm, one might envision using cloning of one member of the couple's nuclei to produce a child.

Although this constitutes an extension of current clinical practice, aside from the serious, moral, and ethical issues surrounding this approach, there are significant technical and medical causes for caution, some of which were described in the research questions enumerated above. In most situations of assisted reproduction, other than the intentional union of the gametes by IVF techniques, the fertilized egg and initial cells of the early embryo are not otherwise manipulated. In some rare cases, such as preimplantation genetic diagnosis, the embryo is manipulated by the removal of one of the identical cells of the blastomere to test its genetic status. In contrast, if nuclear transfer were to be used as a reproductive option, it would entail substantially more invasive manipulation. Thus far, the animal cloning of Dolly is a singular success, one seemingly normal animal produced from 277 nuclear transfers. Until the experiment is replicated the efficiency, and even the validity, of the procedure cannot be fully determined. It is likely that the mere act of manipulating a nucleus and transferring it into an egg could decrease the percentage of eggs that go on to develop and implant normally, as well as increase the rate of birth defects.

CLONING AND GENETIC DETERMINISM

The announcement of Dolly sparked widespread speculation about a human child being created using somatic cell nuclear transfer. Much of the perceived fear that greeted this announcement centered on the misperception that a child or many children could be produced who would be identical to an already existing person. This fear reflects an erroneous belief that a person's genes bear a simple relationship to the physical and psychological traits that compose that individual. This belief, that genes alone determine all aspects of an individual, is called *genetic determinism* (37). Although genes play an essential role in the formation of physical and behavioral characteristics, each individual is in fact the result of a complex interaction between his or her genes and the environment within which they develop. As social and biological beings we are creatures of our biological, physical, social, political, historical, and psychological environments. Indeed, the great lesson of modern molecular genetics is the profound complexity of both gene–gene interactions and gene–environment interactions in the determination of whether a specific trait or characteristic is expressed.

While the concept of complete genetic determinism is wrong and overly simplistic, genes do play a major role in determining biological characteristics including a predisposition to certain diseases. Moreover, the existence of families in which many members are affected by these diseases suggest that there is a single gene that is passed down with each generation that causes the disease. When such a disease gene is identified, scientists often say they have "cloned the gene for" breast cancer, for instance, implying a direct cause and effect of gene and disease. Indeed, the recent efforts of the Human Genome Project (HGP) have led to the isolation of a large number of genes that are mutated in specific diseases, such as Duchenne muscular dystrophy, and certain types of breast and colon cancer.

However, recent scientific findings have revealed that a "one-gene, one-disease" approach is far too simplistic. Even in the relatively small list of genes currently associated with a specific disease, knowing the complete DNA sequence of the gene does not allow a scientist to predict if a given person will get the disease. For example, in breast cancer there can be many different changes in the DNA, and for some specific mutations there is a calculated risk of developing the disease, while for other changes the risk is unknown. Even when a specific genetic change is identified that "causes" the disease in some

people, others may be found who have the same change but do not get the disease. This is because other factors, either genetic or environmental, are altered that mask or compensate for "the" disease gene. Thus even with the most sophisticated understanding of genes, one cannot determine with certainty what will happen to a given person with a single change in a single gene.

Once again, the reason rigid genetic determinism is false is that genes interact with each other and with the environment in extremely complex ways. For example, the likelihood of developing colon cancer, a disease with a strong hereditary component and for which researchers have identified a single "causative" gene, is also strongly influenced by diet. When one considers a human trait that is determined by multiple genes, the situation becomes even more complex. The number of interactions between genes and environment increases dramatically. In fact the ability to predict what a person will be like knowing only their genes becomes virtually impossible because it is not possible to know how the environment and chance factors will influence the outcome.

Thus the idea that one could make through somatic cell nuclear transfer a team of super athletes or a superior military force is simply false. Knowing the complete genetic makeup of an individual would not tell you what kind of person that individual would become. Even identical twins that grow up together and thus share the same genes and a similar home environment have different likes and dislikes, and can have very different talents. The increasingly sophisticated studies coming out of human genetics research are showing that the better we understand gene function, the less likely it is we will ever be able to produce at will a person with any given complex trait.

CONCLUSIONS

Dolly's birth demonstrates the feasibility of somatic cell nuclear transplant cloning. There are many applications that nuclear transfer cloning might have for biotechnology, livestock production, and new medical approaches. Work with embryonic stem cells and genetic manipulation of early embryos in animal species (including nuclear transfer) is already providing unparalleled insights into fundamental biological processes and promises to provide great practical benefit in terms of improved livestock, improved means of producing pharmaceutical proteins, and prospects for regeneration and repair of human tissues.

However, the possibility of using human cloning for the purposes of creating a new individual entails significant scientific uncertainty and medical risk at this time (2). Potential risks include those known to be associated with the manipulation of nuclei and eggs and those yet unknown, such as the effects of aging, somatic mutation, and improper imprinting. These effects could result in high rates of failed attempts at pregnancy as well as the increased likelihood of developmentally and genetically abnormal embryos.

In addition, cloning in this manner would change the nature of procreation so that an individual would not need a partner of the opposite gender to have offspring (although a woman would be needed to carry the pregnancy). What makes this technique distinctive is that for the first time women and men would not need to share genes to create a child. There are various arguments against this form of procreation based on evolutionary biology and the need for a genetic "lottery."

More troubling are the issues of how cloning might feed the egos of the very wealthy, the very powerful, or the less than honorable. It would also necessitate a new way of thinking about relationships because the child would have only one biological parent. This elimination of one biological parent would pose a unique challenge to our existing concepts of parenthood and the order of generations. Finally, it is not clear whether being created in this manner would create adverse psychological effects, and there is reason to believe that there may be significant adverse physical impacts.

The three types of cloning described raise similar ethical concerns about the emotional well-being of the resulting child, effects on human relationships, the safety of the procedures, and the use of artificial means to produce a child. The motivation for the use of each technique, however, may differ dramatically. For example, blastomere separation for immediate transfer as a treatment for infertility raises fewer concerns than the use of somatic cell nuclear transfer to create a child in one's own image.

ACKNOWLEDGMENT

Some of this article is based on "The Science and Application of Cloning," a chapter in *Cloning Human Beings* (National Bioethics Advisory Commission, 1997) for which the author was a writer and consultant. In turn, much of the Commission's scientific analysis relied on material contained in two commissioned papers provided by Janet Rossant, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, and by Stuart Orkin, Children's Hospital and Dana Farber Cancer Institute, Boston.

BIBLIOGRAPHY

1. I. Wilmut et al., *Nature (London)* **385**, 810–813 (1997).
2. National Bioethics Advisory Commission, *Cloning Human Beings: Report and Recommendations of the National Bioethics Advisory Commission*, NBAC, Rockville, MD, 1997.
3. H.W. Jones, R.G. Edwards, and G.E. Seidel, *Fertil. Steril.* **61**, 423–426 (1994).
4. J.L. Hall et al., Paper presented at the 1993 Annual Meeting of the American Fertility Society (1993).
5. R.R. Angell, A.A. Templeton, and I.E. Messinis, *Cytogenet. Cell Genet.* **42**, 1–7 (1986).
6. National Institutes of Health (NIH), *Report of the Human Embryo Research Panel*, NIH, Bethesda, MD, 1994.
7. R. Kolberg, *Science* **262**, 652–653 (1993).
8. R. Voelker, *J. Am. Med. Assoc.* **271**, 331–332 (1994).
9. D. Solter, *Nature (London)* **380**, 24–25 (1996).
10. L. Meng and D.P. Wolf, *Abstr. 51st Annu. Meet. Am. Soc. Reprod. Med.* Program Suppl. S236 (1996).
11. K.H.S. Campbell et al., *Nature (London)* **380**, 64–66 (1996).

12. M. Sims and N.L. First, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6143–6147 (1994).
13. D.S. Rubenstein et al., *Cambridge Q. Healthcare Ethics* **4**(3), 316–339 (1995).
14. J.H. Tanne, *Br. Med. J.* **319**, 593 (1999).
15. J.B. Gurdon, *J. Embryol. Exp. Morphol.* **10**, 622–640 (1962).
16. R. Briggs and T.J. King, *Proc. Natl. Acad. Sci. U.S.A.* **38**, 455–463 (1952).
17. J.B. Gurdon, R.A. Laskey, and O.R. Reeves, *J. Embryol. Exp. Morphol.* **34**, 93–112 (1975).
18. J. McGrath and D. Solter, *Science* **226**, 1317–1319 (1984).
19. J. Rossant and R.A. Pederson, *Experimental Approaches to Mammalian Embryonic Development*, Cambridge University Press, Cambridge, UK, 1986.
20. S.M. Willadsen, *Nature (London)* **320**, 63–65 (1986).
21. J. Rossant, *The Science of Animal Cloning*, commissioned paper prepared for the National Bioethics Advisory Commission, Rockville, MD, 1997.
22. H.T. Cheong, Y. Takahashi, and H. Kanagawa, *Biol. Reprod.* **48**, 958–963 (1993).
23. P. Collas and F.L. Barnes, *Mol. Reprod. Dev.* **38**, 264–267 (1994).
24. L.C. Smith and I. Wilmut, *Biol. Reprod.* **40**, 1027–1035 (1989).
25. R.S. Prather, M.M. Sims, and N.L. First, *Biol. Reprod.* **41**, 414–418 (1989).
26. X. Yang et al., *Biol. Reprod.* **47**, 636–643 (1992).
27. H.M. Blau et al., *Science* **230**, 758–766 (1985).
28. K.H.S. Campbell et al., *Biol. Reprod.* **50**, 385–1393 (1994).
29. S.L. Stice, C.L. Keefer, and L. Mathews, *Mol. Reprod. Dev.* **38**, 61–68 (1994).
30. P. Braude, V. Bolton, and S. Moore, *Nature (London)* **332**, 459–461 (1988).
31. R.M. Schultz, *BioEssays* **15**, 531–538 (1993).
32. D. Solter, *Annu. Rev. Genet.* **22**, 127–146 (1988).
33. L.L. Mantell and C.W. Greider, *EMBO J.* **13**, 3211–3217 (1994).
34. P.G. Shiels et al., *Nature (London)* **399**, 316–317 (1999).
35. M.J. Weiss and S.H. Orkin, *J. Clin. Invest.* **97**, 591–595 (1995).
36. F.H. Gage, J. Ray, and L.J. Fisher, *Annu. Rev. Neurosci.* **18**, 159–192 (1995).
37. P. Ramsey, *Fabricated Man: The Ethics of Genetic Control*, Yale University Press, New Haven, CT, 1970.

See other CLONING entries; FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, NATIONAL BIOETHICS ADVISORY COMMISSION.

CLONING, POLICY ISSUES

ELISA EISEMAN
RAND Science and Technology Policy Institute
Washington, District of Columbia

OUTLINE

INTRODUCTION

The birth of Dolly, the first mammal to be cloned from an adult, has brought into sharp focus the future possibility of cloning human beings. Because of the inherent moral, ethical, and legal implications associated with cloning of human beings, policy responses around the world have been intense. In the United States there has been legislative action taken at the state and federal levels. The U.S. Food and Drug Administration (FDA) has also asserted its authority to regulate the cloning of human beings. Similarly several other nations have enacted laws prohibiting the cloning of human beings, and international organizations have issued policy statements calling for a worldwide ban on the cloning of human beings. While the cloning of Dolly has revolutionized science by proving that it is possible to clone an adult mammal and cloning technology may one day transform medicine by providing improved treatments for diseases, there appears to be broad international agreement that cloning of human beings for reproductive purposes should be prohibited.

BACKGROUND

Cloning, which literally means to make a copy, is the asexual reproduction of a precise genetic copy of a molecule, cell, tissue, plant, or animal. However, scientists use the word “cloning” in many different ways. Molecular cloning refers to the copying of DNA fragments. For example, the human gene for insulin has been cloned into bacteria to produce insulin for the treatment of diabetes. In

addition, human cells are routinely cloned to study cancer or genetic diseases. These types of cloning are integral tools in biotechnology, and they have been used to produce breakthrough medicines, diagnostics and vaccines to treat heart attacks, cancer, diabetes, hepatitis, cystic fibrosis, and many other diseases.

The cloning of animals was originally conceived of as a way to understand the genetic processes that regulate development and differentiation. The first attempts to clone animals occurred in 1952 with amphibians and in the 1980s with mammals (1). In these experiments, animals were successfully cloned only when cells from embryos were used, and none were cloned from cells of adult animals. It was not until the birth of Dolly on July 5, 1996, that scientists were able to show that it was possible to use genetic material from a single adult mammalian cell to develop a new individual (2). The ability to clone adult animals has moved the prospects for cloning into areas well beyond basic developmental biology, and has paved the way for major advances in biotechnology, reproductive medicine, and cell-based therapies.

In the last four years, since the cloning of Dolly, there have been several scientific advances. The ability to clone mammals other than sheep from adult cells has been reported for cows and mice (3,4). In addition techniques have been developed to produce cloned animals carrying specific genes, providing an efficient means for producing genetically engineered animals that can make proteins in their milk that could then be used for pharmaceutical or clinical purposes (5,6). The last cow of a rare breed has been cloned in an effort to save the breed from extinction (7), and scientists are preparing to clone other rare and endangered animals (8). Nuclear transfer technology has also been used for human applications, including in preimplantation diagnosis during in vitro fertilization (IVF) (9), to try to treat infertility (10), and in an attempt to produce human embryonic stem cells (11). Many of these advances hold the promise of improved treatments for diseases for which there are currently no good alternatives.

While cloning techniques may one day provide improved treatments for diseases, revolutionize the production of biopharmaceuticals, and save endangered species, mammalian cloning does have its risks. In addition to high rates of spontaneous abortion late in pregnancy and death soon after birth, mammalian cloning has been linked to a developmental defect of the immune system (12) and may be associated with premature aging (13). Thus the question of safety remains, and this casts doubt on the future uses of mammalian cloning, especially for human reproductive purposes. Beyond the safety concerns, the prospect of cloning human beings raises several other ethical, social, and legal concerns. These concerns have been addressed at the state, national, and international level.

POLICY AND LEGISLATION IN THE UNITED STATES

Immediately after the announcement of Dolly's birth, President Clinton asked the National Bioethics Advisory Commission (NBAC) for their recommendations on the use of cloning technology. Soon after, legislation was introduced in the 105th Congress and approximately a dozen states, aimed at prohibiting the cloning of human beings (14). President Clinton also transmitted legislation to Congress that would make it illegal for anyone to create a human being through cloning. More recently two bills have been introduced in the 106th Congress. While there was no law in the United States that directly prohibited creating a child through somatic cell nuclear transfer, there were already a variety of state and federal laws and some existing policies that did apply. Summarized in the discussion below are the NBAC's recommendations as well as enacted and pending legislation at the federal and state levels that both directly and indirectly prohibit the cloning of human beings (Tables 1–4).

National Bioethics Advisory Commission

On February 24, 1997, two days after the news about the birth of Dolly, President Clinton asked NBAC to deliver a report to him within 90 days on the legal and ethical issues involved in the cloning of human beings and "possible federal actions to prevent its abuse." On June 9, 1997, NBAC delivered its recommendations to the President. In their report *Cloning Human Beings*, NBAC agreed that the creation of a child by somatic cell nuclear transfer is scientifically and ethically objectionable at this time. The reasons cited were the following: (*1*) the efficiency of nuclear transfer is so low and the chance of abnormal offspring so high that experimentation of this sort in humans was premature, and (*2*) the cloning of an already existing human being may have a negative impact on issues of personal and social well-being such as family relationships, identity and individuality, religious beliefs, and expectations of sameness (14).

NBAC also recommended a continuation of both the moratorium on the use of federal funding in support of any attempt to create a child by somatic cell nuclear transfer, as well as the voluntary moratorium for the private and nonfederally funded sectors (14). NBAC further recommended that federal legislation be enacted to prohibit anyone from attempting, whether in a research or clinical setting, to create a child through somatic cell nuclear transfer cloning (14). Such legislation should include a sunset clause to ensure that Congress reviews the issue after a specified time period (three to five years) in order to decide whether the prohibition continues to be needed (14). In addition any regulatory or legislative actions undertaken to prohibit the creation of a child by somatic cell nuclear transfer should be carefully written so as not to interfere with other important areas of scientific research, such as the cloning of human DNA sequences and cell lines, neither of which raises the scientific and ethical issues that would arise from an attempt to create a child through somatic cell nuclear transfer (14).

Administration Policy

On March 4, 1997, President Clinton released a statement to the heads of executive departments and agencies prohibiting the use of federal funds for cloning of human beings. Even though restrictions already exist on the use of federal funds for the creation of human embryos

for research purposes (see Federal Legislation below), these restrictions do not explicitly cover the creation of human embryos for implantation and do not cover all federal agencies. Therefore President Clinton issued his statement "to make it absolutely clear that no federal funds shall be allocated for cloning of human beings." In addition to the ban on the use of federal funds, President Clinton also asked for a voluntary moratorium by researchers funded by private sources.

Acting on NBAC's key recommendation, President Clinton announced the "Cloning Prohibition Act" of 1997 on June 9, 1997. Consistent with NBAC's recommendations, the President's legislative proposals prohibited the use of somatic cell nuclear transfer to create a human being for five years and directed the NBAC to report to the President in four and a half years on whether to continue the ban. The proposal was carefully worded to ensure that it would not interfere with beneficial biomedical and agricultural activities. This legislation therefore would not prohibit the use of somatic cell nuclear transfer techniques to clone DNA in cells and it would not ban the cloning of animals. To date, this legislation has not been signed into law; however, the ban on federal funding the President declared in March remains in effect. In addition the President called upon the private sector to refrain voluntarily from using this technology to attempt to clone a human being.

The Office of Management and Budget released a Statement of Administrative Policy on February 9, 1998, in response to Senator Lott's Human Cloning Prohibition Act (S.1601, see below). The Statement detailed the Administration's position that it did not support the passage of S.1601 in its current from because it was "too far-reaching" and it would "prohibit important biomedical research aimed at preventing and treating serious and life-threatening diseases." Instead, the Administration offered several amendments to S.1601, including: (*1*) a five-year sunset on the prohibition on human somatic cell nuclear transfer technology to ensure that there is a continuing examination of the risks and benefits, (*2*) permitting somatic cell nuclear transfer using human cells for the purpose of developing stem cell technology, (*3*) striking the bill's criminal penalties and instead making any property derived from or used to commit violations of the Act subject to forfeiture to the United States, and (*4*) striking the provision establishing a new Commission to Promote a National Dialogue on Bioethics, since it would be duplicative of NBAC's mission. The President's proposal would "prohibit any attempt to create a human being using somatic cell nuclear transfer, provide for further review of the ethical and scientific issues associated with the use of somatic cell nuclear transfer, and protect important biomedical research."

Federal Legislation

In fiscal years 1996 and 1997, Congress passed prohibitions on the use of funds appropriated to the Departments of Labor, Health and Human Services, and Education, and Related Agencies for any research that involves exposing embryos to risk of destruction for nontherapeutic research (P.L. 104-91 and P.L. 104-208). The net effect of these appropriation decisions has been to eliminate virtually all federal funding for research to perfect methods for cloning human beings, including research aimed at initiating pregnancy, since it would probably involve the destruction of many embryos that failed to develop normally (14). This type of research could, however, proceed uninhibited in the private sector.

More recently, language that directly prohibits the use of federal funds for cloning of humans beings has been included in the appropriations legislation for the Departments of Labor, Health and Human Services, and Education, and Related Agencies in fiscal years 1998, 1999, and 2000 (Table 1). These appropriations continue the human embryo research ban in the public sector by prohibiting the use of federal funds for the creation of a human embryo for research purposes or for research in which a human embryo is destroyed, discarded, or knowingly subjected to risk of injury or death greater than that allowed for research on fetuses in utero. By expanding the definition of a human embryo to "include any organism, not protected as a human subject under 45 CFR 46 as of the date of the enactment of this Act, that is derived by fertilization, parthenogenesis, cloning, or any other means from one or more human gametes or human diploid cells" (P.L. 105-78 and P.L. 105-277), these appropriations also effectively prohibit the use of federal funds for the cloning of human beings.

There are two other federal laws and policies that do not directly prohibit cloning, but may have some applicability. The Fertility Clinic Success Rate and Certification Act of 1992 (42 U.S.C.A. Sec, 263a-1 et seq.) requires that clinics using assisted reproduction techniques, such as IVF be monitored. The Act covers all laboratories and treatments that involve manipulation of human eggs and embryos, and requires that pregnancy success rates be reported to the Department of Health and Human Services (DHHS) for publication in a consumer guide. DHHS is also directed to develop a model program to be implemented by the states for the inspection and certification of laboratories that use human embryos. Since any effort to use cloning to create a child would involve manipulation of human eggs and embryos, these requirements would probably also apply to efforts to clone human beings.

The Federal Policy for the Protection of Human Subjects (also called the "Common Rule") describes the requirements for conducting research on human subjects, such as ensuring that human subjects are not exposed to unreasonably risky experiments and are enrolled in research only after giving informed consent (45 CFR Part 46, Subpart A). The Common Rule, promulgated by 17 federal agencies that conduct, support or otherwise regulate human subjects research, governs research that is conducted with federal funds or is performed at institutions that have executed an assurance with the federal government. (These assurances typically promise that any researcher affiliated with the institution will abide by the federal regulations even if that particular researcher is not using federal funds.) Other human subjects regulations codified at Title 45 Part 46 of the Code of Federal Regulations include additional protections pertaining to research involving fetuses, pregnant women, and human IVF. Enforcement of these protections is

Table 1. Enacted Federal Legislation Prohibiting Cloning of Human Beings

Public Law	Title	Synopsis	Status
P.L. 106-113	Health and Human Services FY2000 Appropriations bill (part of the Omnibus Appropriations Bill FY2000)	Continues the ban on the use of federal research funds for human embryo research. This means that federal funds may not be used for the creation of a human embryo for research purposes or for research in which a human embryo is destroyed, discarded or knowingly subjected to risk of injury or death greater than that allowed for research on fetuses in utero. The definition of a human embryo includes "any organism, not protected as a human subject under 45 CFR 46 as of the date of the enactment of this Act, that is derived by fertilization, parthenogenesis, cloning, or any other means from one or more human gametes or human diploid cells."	Sponsor: Rep Istook Introduced as H.R. 3242 by Rep. Young: 11/17/99. Incorporated into H.R. 3194. Became Public Law 106-113: 11/29/99.
P.L. 105-277	Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, 1999 (part of the Omnibus Appropriations Bill FY99)	Continues the ban on the use of federal research funds for human embryo research. This means that federal funds may not be used for the creation of a human embryo for research purposes or for research in which a human embryo is destroyed, discarded or knowingly subjected to risk of injury or death greater than that allowed for research on fetuses in utero. The definition of a human embryo includes "any organism, not protected as a human subject under 45 CFR 46 as of the date of the enactment of this Act, that is derived by fertilization, parthenogenesis, cloning, or any other means from one or more human gametes or human diploid cells."	Sponsor: Rep Wolf Introduced as H.R. 4274 by Rep. Porter: 07/20/98. Incorporated into H.R. 4328. Became Public Law 105-277: 10/21/98.
P.L. 105-78	Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, 1998	Continues the ban on the use of federal research funds for human embryo research. This means that federal funds may not be used for the creation of a human embryo for research purposes or for research in which a human embryo is destroyed, discarded or knowingly subjected to risk of injury or death greater than that allowed for research on fetuses in utero. Congress also expanded the definition of a human embryo to "include any organism, not protected as a human subject under 45 CFR 46 as of the date of the enactment of this Act, that is derived by fertilization, parthenogenesis, cloning, or any other means from one or more human gametes or human diploid cells."	Sponsor: Rep. Porter Introduced as H.R. 2264: 07/25/97. Became Public Law: 105-78: 11/13/97.

primarily the responsibility of Institutional Review Boards (IRBs), which review experiments before people can be enrolled. Any effort to use federal funds to clone a human being would raise serious questions about the physical harms that might result, making it difficult for an IRB to approve such research.

Just 11 days after the announcement of Dolly, Representative Ehlers introduced two bills, H.R. 922 and H.R. 923, in the House. H.R. 922 would have prohibited the "expenditure of Federal funds to conduct or support any research on the cloning of humans." H.R. 923 would have made it unlawful for any person to use a human somatic cell for the process of producing a human clone, and set forth a civil money penalty. Several other bills were introduced almost a year later, following Richard Seed's announcement that he intended to clone human beings. Altogether, nine bills prohibiting the cloning of human beings were introduced in the 105th Congress, six in the Senate and three in the House.

While no action was taken on any of these bills before the 105th Congress adjourned, one bill, H.R. 3133, was reintroduced in the 106th Congress by Representative Stearns as H.R. 2326—the Human Cloning Research Prohibition Act. H.R. 2326 prohibits the use of federal funds to "conduct or support any project of research that includes the use of somatic cell nuclear transfer technology to produce an oocyte that is undergoing cell division toward development of a fetus" (Table 2). To date, one other bill prohibiting federal funding of human cloning has been introduced during the 106th Congress. The Human Cloning Prevention Act of 1999, H.R.571, sponsored by Representative Paul, prohibits "federal payments to any business, institution, or organization that engages in human cloning or human cloning techniques" (Table 2).

While both bills expressly prohibit the use of federal funds for research on the cloning of a human being, neither of these bills set forth penalties, such as fines or prison time, for these actions. It is also interesting to note that H.R. 2326 specifically prohibits the use of human somatic cell nuclear transfer technology, while H.R. 571 does not specify any particular technique for cloning. H.R. 2326 contains additional language that strives not to restrict areas of biomedical and agricultural research that use somatic cell nuclear transfer to clone molecules, DNA, cells, tissues, and nonhuman animals. H.R. 2326 also contains language that requires the Director of the National Science Foundation (NSF) to enter into an agreement with the National Research Council (NRC) to

Table 2. Pending Federal Cloning Legislation in the 106th Congress

Bill	Title	Synopsis	Status
H.R.2326.IH	Human Cloning Research Prohibition Act	Prohibits the expenditure of federal funds to conduct or support research on the cloning of humans, and to express the sense of the Congress that other countries should establish substantially equivalent restrictions.	Sponsor: Rep. Stearns Introduced in the House: 06/23/99
H.R.571.IH	Human Cloning Prevention Act of 1999	Prohibits federal payments to any business, institution, or organization that engages in human cloning or human cloning techniques.	Sponsor: Rep. Paul Introduced in the House: 02/04/99

review the implementation of any legislation prohibiting the cloning of human beings. Finally, H.R. 2326 states that foreign countries should establish similar restrictions set forth in the bill to prohibit the cloning of human beings. None of this additional language is included in H.R. 571.

State Legislation

Even though most states do not have legislation directly regulating assisted reproduction techniques, a number of state laws regarding the management of embryos could restrict even privately funded research aimed at cloning human beings (15). Ten states have laws regulating research and/or experimentation on conceptuses, embryos, fetuses, or unborn children that use broad enough language to include early stage conceptuses: Florida, Louisiana, Maine, Massachusetts, Michigan, Minnesota, North Dakota, New Hampshire, Pennsylvania, and Rhode Island (15).

Five states have enacted legislation that directly prohibits cloning of a human being, California, Louisiana, Missouri, Michigan, and Rhode Island (Table 3). California was the first state to enact legislation prohibiting the cloning of human beings. Within three weeks of the announcement of the cloning of Dolly, California introduced a bill into the Senate (SB1344). The bill was signed into law by the governor on October 4, 1997.

Michigan has enacted four separate bills all prohibiting cloning of human beings. Three bills were introduced in the House (HB5475, HB4846, and HB4962), and one bill was introduced in the Senate (SB864). All three of the House bills needed to be enacted into law for each act to take effect. All four bills passed the Senate by a vote of 37 to zero. In addition all four bills were presented to the governor for signature on May 20, 1998, and signed into law by the governor on June 3, 1998.

Rhode Island and Missouri both introduced bills in January 1998. In Rhode Island, the bill became law without the governor's signature on July 7, 1998 (HB7123). In Missouri, the bill was signed into law by the governor on July 10, 1998 (SB722). Louisiana is the latest state to enacted cloning legislation. Legislation banning the cloning of human beings was introduced into Louisiana's Senate on March 29, 1999, and signed into law by the governor on July 2, 1999 (SB825).

In addition, there are four states that have pending legislation to prohibit the cloning of human beings (Table 4). Massachusetts has three bills pending, New York has five bills pending, and New Jersey and Ohio each have one bill pending. Six states, Arkansas, Connecticut, Illinois, Oregon, South Carolina, and Virginia, introduced proposed legislation; however, bills in these states are now inactive because of the adjournment of their legislatures.

REGULATION IN THE UNITED STATES

Food and Drug Administration

In response to provocative statements by scientist Richard Seed, who announced January 7, 1998, that he plans to clone human beings, FDA announced that it has the authority to regulate human cloning. FDA asserts that human cloning by somatic cell nuclear transfer requires "more than minimal manipulation" of a cell, and therefore requires approval by the FDA under Section 351 of the *Public Health Service Act* (16). In addition cellular products resulting from "more than minimal manipulation" of cells would require approval for safety and efficacy under provisions in the *Public Health Service Act* that regulate products derived from human materials (16). Acting Commissioner Michael Friedman has affirmed that the FDA will take legal action against anyone who attempts to clone a human being without obtaining prior approval from the FDA (17).

In October 1998 Stuart L. Nightingale, the Associate Commissioner for Health Affairs at FDA, distributed a letter detailing FDA's position on the use of cloning technology to create a human being. The purpose of this letter was to confirm to IRBs that FDA has jurisdiction over clinical research involving cloning of human beings, and to inform IRBs of the FDA regulatory process that is required before any investigator can proceed with such a clinical investigation. The letter states:

> Clinical research using cloning technology to create a human being is subject to FDA regulation under the *Public Health Service Act* and the *Federal Food, Drug, and Cosmetic Act.* Under these statutes and FDA's implementing regulations, before such research may begin, the sponsor of the research is required to submit to FDA an Investigational New Drug Application (IND) describing the proposed research plan; to obtain authorization from a properly constituted and functioning IRB; and to obtain a commitment from the investigators to obtain informed consent from all human subjects of the research. Such research may proceed only when an IND is in effect. Since FDA believes that there are major unresolved safety questions pertaining to the use of cloning technology to create a human being, until those questions are appropriately addressed in the IND, FDA would not permit any such investigation to proceed (17a).

Table 3. Enacted State Cloning Legislation

State	Bill	Synopsis	Status
California	SB1344	Prohibits a person from cloning a human being and from purchasing or selling an ovum, zygote, embryo, or fetus for cloning purposes. Penalties for a corporation not to exceed $1,000,000, penalties for an individual $250,000. Those in violation may lose their license. Provisions will be repealed on 1/1/03 unless a later enacted statue deletes or extends that date.	Sponsor: Johnston Introduced: 3/11/97. Passed Assembly: 9/2/97. Passed Senate: 9/10/97. Signed by the governor: 10/4/97. Filed with Secretary of State: 10/6/97.
Louisiana	SB825	Prohibits human cloning.	Sponsor: Hines Introduced: 3/29/99. Passed Senate: 5/11/99. Passed House; to Senate for concurrence: 6/14/99. Signed by governor: 7/2/99.
Michigan	HB4846	Amends the Public Health Code relating to the practice of a health profession by a licensee, a registrant, or an applicant for licensure or regulation. Prohibits a licensee or registrant or other individual from cloning or attempting to clone a human being. A licensee or registrant or other individual who violates this subsection is subject to a civil penalty of $10,000,000.	Sponsor: Profit Introduced: 1/14/98. Passed House: 1/29/98. Passed Senate by a 37-0 vote: 4/28/98. Bill received House concurrence and with SB864, HB5475, HB4962 was presented to governor for signature: 5/20/98. Signed by the governor: 6/3/98.
Michigan	HB4962	Amends the Penal Code to prohibit an individual from cloning or attempting to clone a human being. Provides felony penalties of not more than 10 years imprisonment or a fine of not more than $5,000.00 or both. Definitions of "clone," "cloning," "human somatic cell nuclear transfer," and "somatic cell" are the same as in HB 4846.	Sponsor: McManus Introduced: 1/14/98. Passed House: 1/29/98. Passed Senate by a 37-0 vote: 4/28/98. Bill received House concurrence and with SB864, HB5475, HB4846 was presented to governor for signature: 5/20/98. Signed by the governor: 6/3/98.
Michigan	HB5475	Prohibits the expenditure of state funds to clone a human being or to conduct or to support research on the cloning of human beings. Definitions are the same as HB 4846 and HB 4962. All three bills must be enacted into law for each act to take effect.	Sponsor: Mans Introduced: 1/14/98. Passed the House: 1/29/98. Passed Senate by a 37-0 vote: 4/28/98. Bill received House concurrence and with SB 864, HB4962, HB4846 was presented to governor for signature: 5/20/98. Signed by the governor: 6/3/98.
Michigan	SB864	Prohibits human cloning for a period of 5 years and provides for civil and criminal penalties.	Sponsor: Bennett Introduced: 2/5/98. Passed Senate by a 37-0 vote: 4/28/98. SB864, HB4846, HB4962 & HB5475 were concurred and presented to governor for signature: 5/22/98. Signed by the governor: 6/3/98.
Missouri	SB722	The Senate bill incorporates HB 1316. Section 17 of the bill states that no state funds shall be used for research with respect to the cloning of a human person. For purposes of this section, the term "cloning" means the replication of a human person by taking a cell with genetic material and cultivating such cell through the egg, embryo, fetal, and newborn stages of development into a new human person.	Sponsor: Sims Introduced: 1/14/98. Passed Senate: 3/4/98. Passed House; to the Senate for concurrence: 5/6/98. Senate concurred in House amendments: 5/8/98. Signed by the governor: 7/10/98.
Rhode Island	HB7123	Prohibits cloning of a human being and purchasing or selling of an ovum, zygote, embryo, or fetus for the purpose of cloning a human being. Provides for civil penalties in the amount not to exceed $1,000,000 for a corporation, etc.; not to exceed $250,000 for an individual for violations of the act.	Sponsor: Cambio Introduced: 1/9/98. Passed House: 4/29/98. Passed on Senate Floor: 6/26/98. House concurred with amendment: 6/29/98. Became law without governor's signature: 7/7/98.

Source: From Pharmaceutical Research and Manufacturers of America (PHRMA), Cloning Legislation and Regulation, http://www.phrma.org/genomics/cloning/legislation.html

Table 4. Pending State Legislation

State	Bill	Synopsis	Status
Massachusetts	HB2455	Prohibits science of cloning.	Sponsor: Fallon Introduced 1/6/99.
Massachusetts	HB2462	Regulates the science of cloning.	Sponsor: Harkins Introduced 1/6/99.
Massachusetts	SB1394	Prohibits human cloning.	Sponsor: Magnani Introduced: 1/6/99. Referred to Senate Committee on Science and Technology: 1/6/99.
New Jersey	AB329	States that a person who knowingly engages or assists, directly or indirectly, in the cloning of a human being is guilty of a crime of the first degree. Defines "cloning of a human being" to mean the replication of a human individual by cultivating a cell with genetic material through the egg, embryo, fetal and newborn stages into a new human individual. Amends the Genetic Privacy Act of 1996 to provide that an individual's genetic information is the property of the individual, and deletes exceptions from current New Jersey law where procedures for obtaining informed written consent already are governed by national standards.	Sponsors: Doria and Gill Introduced: 1/13/98. Referred to Assembly Health Committee: 1/13/98. Carryover to 1999 legislative session.
New York	SB2123	Provides that appropriations and reappropriations to the New York State Advisory Commission on Cloning and Genetic Engineering shall be subject to the provisions which apply to all other legislative commissions; creates the temporary state commission on cloning and genetic engineering to examine, evaluate, and make recommendations to the legislature and governor on the scientific, technical, moral, and ethical issues raised by cloning.	Sponsor: Goodman Introduced: 2/3/99. Referred to Senate Committee on Finance: 2/3/99. Withdrawn from Senate Committee on Finance: 3/18/99. Referred to Senate Committee on Corporations, Authorities, and Commissions: 3/18/99.
New York	SB1954	Enacts Cloning Prohibition and Research Protection Act, prohibits cloning of human beings and provides a $250 civil fine for violation of prohibition.	Sponsor: Goodman Introduced: 2/1/99. Referred to Senate Committee on Health: 2/1/99.
New York	SB1179	Prohibits any person from cloning a human and from purchasing or selling an ovum, zygote, embryo, or fetus for the purpose of cloning a human and establishes civil penalties of up to a specified amount.	Sponsor: Marchi Introduced: 1/15/99. Referred to Senate Committee on Health: 1/15/99.
New York	AB6874	Prohibits human cloning and the use of public funds, resources, property, employees, or those of political subdivisions or public corporations in furtherance thereof; makes violation a felony and grounds for license revocation.	Sponsor: Labriola Introduced: 3/10/99. Referred to Assembly Committee on Health: 3/10/99.
New York	AB3026	Prohibits any person from cloning a human being and from purchasing or selling an ovum, zygote, embryo, or fetus for the purpose of cloning a human being; establishes civil penalties; requires the Commissioner of Health to submit a report to the governor and the legislature on the implications of human cloning.	Sponsor: Connelly Introduced: 1/28/99. Referred to Assembly Committee on Health: 1/28/99.
Ohio	SB102	Prohibits cloning a human being. Makes it unlawful to purchase or sell an ovum, zygote, embryo, or fetus for the purpose of cloning. Creates a "Cloning Enforcement Fund" in the state treasury which would consist of moneys from civil penalties. Civil penalty would not exceed $5,000.	Sponsor: Ray Introduced: 3/11/99. Referred to Senate Committee on Reference: 3/11/99. Senate Committee on Reference recommended referral: 3/16/99. Sent to Senate for second reading; read a second time: 3/16/99. Sent to Senate Committee on Judiciary: 3/16/99. Hearing in Senate Committee on Judiciary: 4/21/99.

Source: From Pharmaceutical Research and Manufacturers of America (PHRMA), Cloning Legislation and Regulation, http://www.phrma.org/genomics/cloning/legislation.html

However, the FDA has not specified which provision of current law grants it such authority (17). There are three possible bases for FDA's assertion of jurisdiction over cloning of human beings: (*1*) classification as a "drug" under Section 201(g) of the Federal Food, Drug, and Cosmetic Act (FDCA), (*2*) classification as a "medical device" under Section 201(h) of the FDCA, and (*3*) classification as a "biological product" under Section 351(a) of the Public Health Service Act (PHSA) (17). If human cloning is covered by any of these statutory provisions, the FDA would have authority to require premarket approval and/or licensing based on reasonable clinical assurance of safety and efficacy (17). However, FDA's authority to regulate cloning of human beings has been questioned, and the matter may require a statutory amendment to expand FDA's authority (17).

CLONING LEGISLATION IN OTHER NATIONS

Due to the transnational characteristics of science, there exists a need for international cooperation regarding the conduct of scientific and medical research. Some of this need may be met through legislation adopted on a country-by-country basis, but some international agreement is probably also needed. NBAC recognized the importance of international cooperation in the effort to prohibit the cloning of human beings, and concluded that "[t]he United States Government should cooperate with other nations and international organizations to enforce any common aspects of their respective policies on the cloning of human beings" (14).

The possibility of cloning human beings has prompted responses from several nations. Several countries already had existing legislation that prohibited the cloning of human beings, including Australia, Austria, Denmark, France, Germany, Norway, Slovakia, Spain, Sweden, Switzerland, and the United Kingdom. Three countries, Israel, Malaysia, and Peru, passed cloning legislation in response to the news of Dolly. In addition Argentina, Belgium, Canada, China, Japan, and South Korea have proposed legislation but have not yet passed laws to prohibit the cloning of human beings. Countries that already have laws or have announced plans to pass laws prohibiting the cloning of human beings are discussed below and in Table 5.

Countries With Existing Cloning Legislation Before Dolly Was Cloned

Even though the announcement of the first cloning of an adult mammal seemed to take everyone by surprise, several countries already had existing legislation that prohibited the cloning of human beings. South Australia and Spain have had laws prohibiting the cloning of human beings since 1988. In Victoria, Australia, and the United Kingdom legislation was drafted and implemented based on reports from their national ethics commissions. In addition the ethics commissions in Australia, France, and the United Kingdom provided their respective governments with further recommendations after the cloning of Dolly was announced. In contrast, Austria, Norway, Slovakia, and Sweden have laws that only implicitly prohibit cloning of human beings.

Australia. Three Australian states, Victoria, South Australia, and Western Australia, already have existing legislation preventing reproductive cloning. In addition, in October 1997, the New South Wales government issued a discussion paper entitled "Review of the Human Tissue Act 1983." In this paper the Minister for Health of New South Wales announced that the government had introduced a law to ban human cloning and trans-species fertilization involving human gametes or embryos. This ban was developed in response to community concern.

The Commonwealth Minister for Health and Aged Care requested the Australian Health Ethics Committee (18) of the National Health and Medical Research Council (NHMRC) to advise him on the need for possible legislation regarding cloning of human beings. In their report of December 18, 1998, entitled "Scientific, Ethical and Regulatory Considerations Relevant to Cloning of Human Beings," the AHEC advised that:

- A basic distinction should be drawn between the cloning of a *whole* human individual and the copying (also referred to as "cloning") of the component *parts* of a human (such as DNA and cells);
- The cloning of individual human beings is prohibited by State legislation in Victoria, South Australia and Western Australia and is prohibited by NHMRC guidelines; and
- Legislation should be introduced in the remaining States and Territories to regulate human embryo research and to prohibit research on human embryos except as it is permitted in the NHMRC's *Ethical guidelines on assisted reproductive technology* (18).

NHMRC's *Ethical guidelines on assisted reproductive technology* already prohibits experimentation with the intent to produce two or more genetically identical individuals, including development of human embryonic stem cell-lines with the aim of producing clones of individuals. Although infringement of these guidelines is not a legal offense, sanctions usually involve loss of access to NHMRC research funds. These guidelines are regarded as national standards of acceptable practice.

In May 1998, the Australian Academy of Science initiated a project on human cloning to contribute to the public debate in this area. In February 1999, the Academy released its position statement "On Human Cloning" (19). In its statement the Academy distinguishes between "reproductive cloning" to produce a human fetus and "therapeutic cloning" to produce human stem cells, tissues and organs, and bases its recommendations on this distinction (19). The Academy's first recommendation states "that reproductive cloning to produce human fetuses is unethical and unsafe and should be prohibited" (19). The statement goes on to say that "human cells, whether derived from cloning techniques, from embryonic stem cell lines, or from primordial germ cells should not be precluded from use in approved research activities in cellular and developmental biology" (19). Based on its recommendations, the Academy concludes that the

Table 5. Legislation in Other Countries Prohibiting Cloning of Human Beings

Country	Law	Date	Synopsis
Argentina		(proposed)	Intend to deter efforts to clone human beings using somatic cell nuclear transfer.
Victoria, Australia	Infertility Treatment Act	1995	Bans cloning of human beings.
South Australia	Reproductive Technology Act	1988	Bans cloning of human beings.
Western Australia	Human Reproductive Technology Act	1991	Bans cloning of human beings.
Austria	Federal Law on Medically Assisted Procreation	1992	Implicitly prohibits cloning of human beings.
Belgium		(proposed)	Legislation covering medical ethics including cloning is currently being considered by Parliament.
Canada	An Act to amend the Criminal Code (genetic manipulation) (Bill C-247)	(proposed)	Would criminalize human cloning and germ-line genetic alteration without prohibiting beneficial scientific research in genetics.
China	Human Reproductive Technology Bill	(proposed)	Intend to deter efforts to clone human beings using somatic cell nuclear transfer.
Denmark	Scientific Ethical Committee System and the Handling of Biomedical Research Projects (Act No. 503)	1992	Research on cloning (production of genetically identical individuals) is forbidden, as is nuclear substitution.
Denmark	Medically Assisted Procreation in Connection with Medical Treatment, Diagnosis and Research (Act No. 460)	1997	Confirms the Danish Parliament's position of January 25, 1995, that treatment can not be initiated in areas where a research ban already exists under the 1992 Act.
France	Federal Bioethics Legislation (Laws 94-653 and 94-654)	1994	Implicitly prohibits human cloning. Bioethics Committee recommended that the ban should be made more explicit when the bioethics legislation is revised in 1999.
Germany	Federal Embryo Protection Act	1990	The creation of an embryo genetically identical to another embryo, fetus, or any living or dead person is an offense.
Israel	Anti-genetic Intervention Law	1998	Places a five year moratorium on any attempt to clone human beings or create a human being through germ-line gene therapy. Does not prohibit research and development of cloning technologies.
Japan		(proposed)	A Committee of the Council for Science and Technology has proposed a ban on human cloning.
Malaysia		1997	Bans the cloning of human beings.
Norway	Medical Use of Biotechnology (Law No. 56)	1994	Implicitly prohibits embryo cloning.
Peru	General Health Law	1997	Prohibits human cloning.
Slovakia	1994 Health Care Law	1994	Implicitly prohibits embryo cloning.
South Korea		(proposed)	Legislators want to ban all human cloning experiments except those that relate to disease research. A proposal before the National Assembly creates a committee of representatives from government, religious groups, research, and industry.
Spain	Spanish Civil Code-Assisted Reproduction Procedures (Law No. 35/1988)	1988	Explicitly prohibits embryo and oocyte cloning with criminal sanctions.
Sweden	Measures for the Purposes of Research or Treatment in Connection with Fertilized Human Oocytes (Law No. 115)	1991	Implicitly prohibits embryo and oocyte cloning with criminal sanctions.

(*continued*)

Table 5. ***Continued***

Country	Law	Date	Synopsis
Switzerland	Law on Reproductive Medicine in Humans	1990	prohibits interventions on the genetic material of gametes, live embryos, and fetuses; prohibits measures aimed at influencing the sex or inherited characteristics of the unborn child; prohibits use of live embryos, fetuses, and parts thereof for research purposes; and prohibits cloning, creation of chimeras, interspecies hybridization, and extracorporeal procreation.
Switzerland	Amendment of Federal Constitution	1992	Legally binding, implicitly prohibits embryo cloning.
Switzerland	Federal Bill on Medically Assisted Procreation	1996	Proposes criminal sanctions for the artificial creation of genetically identical beings.
United Kingdom	Human Fertilisation and Embryology Act	1990	The nuclear substitution of an embryo, or any cell while it forms part of an embryo is expressly prohibited.

Sources: From Refs. 22 and 24.

1996 NHMRC *Ethical guidelines on assisted reproductive technology* and relevant State legislation should be revised to allow research on therapeutic cloning thereby allowing "Australia to participate fully and capture benefits from recent progress in cloning research" (19).

Denmark. Denmark passed *Act No. 503 on a Scientific Ethical Committee System on the Handling of Biomedical Research Projects* in 1992 (20). The 1992 Act forbids research on cloning and nuclear substitution. Cloning is defined as the production of genetically identical individuals. In 1997, *Act No. 460 on Medically Assisted Procreation in Connection with Medical Treatment, and Research* confirms the Danish Parliament's position of January 25, 1995, that treatment cannot be initiated in areas where a research ban already exists under the 1992 Act (20).

France. The Federal Bioethics Legislation, passed in 1994, basically bans embryo research, only allowing research on human embryos if it does not harm the integrity of the embryo. In May 1997, France's national bioethics committee recommended that the ban on human embryo research be loosened to allow for the use of excess embryos from IVF for the development of embryonic stem cells for fundamental and therapeutic research (21). The bioethics committee qualified its recommendation with multiple safeguards including requiring informed consent from the parents of the embryo, as well as bans on the creation of embryos for research, germline modifications, and cloning of human beings. The bioethics committee recommended that the legislature adopt their recommendations during the scheduled revision of France's bioethics legislation in 1999 (21).

Germany. Germany already has some of the world's most restrictive laws on genetic engineering, applying even to food plants such as tomatoes and soybeans. The *Federal Embryo Protection Act 1990* makes the creation of an embryo genetically identical to another embryo, fetus, or any living or dead person an offense punishable by up to five years imprisonment or by a fine (20). The Act also prohibits alteration of the genetic information of the human germline, and the creation of chimeras and hybrids. In March 1997, the German Parliament passed a resolution calling for a comprehensive international ban on human cloning.

Spain. Spain's law on *Assisted Reproduction Procedures* (Law No. 35/1988), passed in 1988, explicitly prohibits embryo and oocyte cloning with criminal sanctions (20). It also prohibits the fertilization of a human ovum for any other purpose than human procreation. This legislation is sufficiently broad enough to prohibit both embryo twinning and somatic cell nuclear transfer because it concentrates on the result rather than the technique used (22).

Switzerland. Switzerland's *Law on Reproductive Medicine in Humans* of October 18, 1990, prohibits interventions on the genetic material of gametes, live embryos, and fetuses (23). It likewise prohibits measures aimed at influencing the sex or inherited characteristics of the unborn child. Live embryos, fetuses, and parts thereof may not be used for research purposes. Furthermore the following are prohibited: cloning, the creation of chimeras, interspecies hybridization, and extracorporeal procreation. Switzerland's Federal Constitution is itself a legally binding document that implicitly prohibits embryo cloning (20). In 1996 Switzerland proposed the *Federal Bill on Medically Assisted Procreation* that would explicitly prohibit the artificial creation of genetically identical beings by imposing criminal sanctions (22).

United Kingdom. The 1984 Report of the Committee of Inquiry into Human Fertilisation and Embryology (Warnock Report) was commissioned by the government following the 1978 birth of the first baby conceived through IVF. The Warnock report was the basis for The *Human*

Fertilisation and Embryology Act of 1990. The *Human Fertilisation and Embryology Act* makes provisions to regulate and monitor treatment centers and to ensure that research using human embryos is carried out in a responsible way. This is done by means of a licensing system administered through the Human Fertilisation and Embryology Authority (HFEA). Three areas of activity are covered by the Act: (*1*) any fertility treatment that involves the use of donated eggs or sperm, or embryos created outside the body (IVF)—these are referred to as licensed treatments, (*2*) storage of eggs, sperm, and embryos, and (*3*) research on human embryos (20).

The *Human Fertilisation and Embryology Act* expressly prohibits one type of cloning technique, the nuclear substitution of any cell while it forms part of an embryo (20,24). However, it does not expressly prohibit embryo splitting or nuclear transplantation. Since both of these techniques involve the creation of embryos outside the body, a license is required from the HFEA. In 1997 the HFEA announced a policy not to issue licenses for any procedures involving embryo splitting or nuclear transfer to any IVF practice either in the private or public sector.

In response to the cloning of Dolly, the Human Genetics Advisory Commission (HGAC) and the HFEA decided to hold a consultation exercise on cloning. In their report issued December 1998, entitled "Cloning Issues in Reproduction, Science and Medicine," it was concluded that the *Human Fertilisation and Embryology Act* has been effective in dealing with new developments relating to human cloning, and should be extended to ban all human reproductive cloning regardless of the technique used (20). In addition it was recommended that somatic cell nuclear transfer to create embryonic stem cells should be allowed. However, under the *Human Fertilisation and Embryology Act*, laboratory research is allowed on human embryos less than 14 days old only if it is used for research into the treatment of infertility and congenital diseases, but research cannot be aimed at developing replacement tissue. Therefore, the scientists at HGAC and HFEA advised that the Secretary of State for Health should consider specifying in regulations two further purposes for which the HFEA might issue licenses for research, so that potential benefits can clearly be explored: (*1*) the development of methods of therapy for mitochondrial disease, and (*2*) the development of therapeutic treatments for diseased or damaged tissues or organs (20).

Austria, Norway, Slovakia, and Sweden. In contrast to the countries described above that have laws explicitly prohibiting cloning of human beings, the laws in Austria, Norway, Slovakia, and Sweden are implicit. Austria's Federal Law of 1992 regulating Medically Assisted Procreation implicitly prohibits cloning of human beings by stating that assisted reproductive techniques must use "viable cells" to achieve pregnancy (22). Sweden's Law No. 115, passed on March 14, 1991, implicitly prohibits embryo and oocyte cloning with criminal sanctions (20). Section 2 states that "the purpose of experimentation shall not be to develop methods aimed at causing heritable genetic effects." Norway's *Law No. 56 on the Medical Use of Biotechnology 1994* and Slovakia's 1994 *Health Care Law* also implicitly prohibit embryo cloning (20).

Countries That Passed Cloning Legislation in Response to the News of Dolly

Three countries, Israel, Malaysia, and Peru, passed cloning legislation in response to the news of Dolly. Legislation was passed in both Malaysia and Peru because cloning of human beings was viewed as unnatural. However in Malaysia, cloning of animals for scientific purposes is allowed. In Israel, there is a five-year moratorium on cloning of human beings; however, the law allows cloning for medical purposes if the Health Minister deems that it does not violate human dignity.

Israel. The Israeli Knesset unanimously passed an anti-genetic intervention law on December 29, 1998. The law places a five-year moratorium on any attempt to clone human beings. Germ-line gene therapy is also forbidden. The law does allow genetic intervention for medical purposes, such as cloning a healthy organ for donation. However, specific clinical research proposals would only be allowed to proceed if safety and efficacy could be established, and if the Health Minister deemed them not to violate human dignity. The Health Minister will be responsible for deciding how to supervise such intervention. Violation of the ban is punishable by two years in prison (25).

Interestingly, the law does not specifically say that human genetic intervention is opposed to human dignity. During the ban, the law states that the Supreme Helsinki Committee will act as an advisory committee "to follow developments in medicine, science, and biotechnology in the sphere of genetic experimentation on human beings, and to report annually, advise, and make recommendations to the Health Minister as to how to proceed, to continue as is, or to reformulate the law" (26).

Malaysia. The Malaysian Cabinet has banned the cloning of human beings because it is "against nature" (27). The cloning of human beings was seen as unethical and an interference with God's creation. However, cloning of animals is allowed for scientific purposes.

Peru. Peru was the first Latin American nation to ban human cloning in a new General Health Law passed by Congress on June 12, 1997 (28). The Congress' health committee found that cloning of human beings goes against people's individuality. The law's aim is to avoid "creating unnatural procreation."

Countries with Proposed Legislation to Prohibit Cloning

Argentina, Belgium, Canada, China, Japan, and South Korea have proposed legislation but have not yet passed laws to prohibit the cloning of human beings. Until Canada passes legislation, cloning of human beings is subject to a voluntary moratorium introduced by the Minister of Health in July 1995. Even though China and Japan have not yet passed legislation, the Chinese Minister of Health and the Japanese Education Ministry have stated that they will not provide funding for research on cloning human beings. South Korea's Ministry of Health and Social Welfare has proposed an expansion of existing rules

that ban the implantation of genetically engineered human embryos to include a prohibition on human cloning. In Argentina and Belgium, legislation to regulate the cloning of human beings is also being considered.

Canada. In its final report the Royal Commission on New Reproductive Technologies concluded that "certain activities conflict so sharply with the values espoused by Canadians and by this Commission, and are so potentially harmful to the interests of individuals and of society, that they must be prohibited by the federal government under threat of criminal sanction." These activities include human zygote or embryo research related to ectogenesis, cloning, animal or human hybrids, and the transfer of zygotes to another species.

Based on the recommendations of the Royal Commission, Canada proposed a comprehensive national policy on the management of human reproductive and genetic technologies in June 1996. The *Human Reproductive and Genetic Technologies Act* would have prohibited 13 unacceptable uses of new reproductive and genetic technologies, including cloning of human embryos, germ-line genetic alteration, and other practices that commercialize reproduction and are contrary to the principles of human dignity, respect for life, and protection of the vulnerable. However, the legislation died on the order paper in April 1997, leaving all research and experiments in Canada subject to a voluntary moratorium introduced by the Minister of Health in July 1995.

On October 9, 1997, Bill C-247, an Act to amend the Criminal Code by adding a section on genetic manipulation was introduced into the House of Commons as a Private Member's Bill. (Private Members' Public Bills, sponsored by a private Member who is not a Minister of the Crown, are public policy initiatives that affect the entire general public or a portion thereof.) Bill C-247 criminalizes human cloning and germ-line genetic alteration without prohibiting beneficial scientific research in genetics. The bill states that

> No person shall knowingly (*a*) manipulate an ovum, zygote or embryo for the purpose of producing a zygote or embryo that contains the same genetic information as a living or deceased human being or a zygote, embryo or foetus, or implant in a woman a zygote or embryo so produced; or (*b*) alter the genetic structure of an ovum, human sperm, zygote or embryo if the altered structure is capable of transmission to a subsequent generation.

Violation of the above prohibitions would be a criminal offense punishable by a fine of up to $500,000, imprisonment for up to ten years, or both. This bill was still being considered by the House of Commons as late as February 1999 (29).

In addition, a group composed of three of Canada's major funding bodies, the Medical Research Council, the Natural Sciences and Engineering Research Council, and the Social Sciences and Humanities Research Council, issued a policy statement entitled the "Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans" in August 1998 (30). This policy statement describes standards and procedures for governing research involving human subjects. Included in the policy statement is a section on "Research Involving Human Gametes, Embryos or Foetuses," Article 9.5 of which states:

> It is not ethically acceptable to undertake research that involves ectogenesis, cloning human beings by any means including somatic cell nuclear transfer, formation of animal/human hybrids, or the transfer of embryos between humans and other species.

While these policies only apply to research funded by these three Councils, application of these policies to privately funded research is being considered.

China. In May 1997 the Chinese Academy of Sciences, China's leading institute of scientific research, banned the cloning of human beings, and called for a committee to set standards for cloning animals (31). In response to strong objections to human cloning by both scientists and the Chinese government, legislation similar to that currently being implemented in Hong Kong will probably soon be passed (32). The new *Human Reproductive Technology Bill* will prohibit the cloning of any human embryo, and specifically outlaw cloning by nuclear transfer (32). In Hong Kong, a statutory monitoring committee has been set up together with an ethics committee to exercise tight control of reproductive technology, and a similar body comprising scientists, ethicists, and government agencies has been strongly advocated in mainland China (32). Meanwhile the Chinese Minister of Health has emphasized the "Policy of the Four Nos" toward research on human cloning, No support, No approval, No license, No acceptance (32).

Japan. In March 1997, the Japanese Ministry of Education, Science, Sports, and Culture announced that it would not provide funding for scientific research on cloning human beings. However, this is an administrative guideline that only applies to state run institutions and carries no penalties for violators (33). In August 1998, the Japanese Science Council, an advisory panel of the Ministry of Education, introduced strict controls on cloning research carried out at universities and national research institutes (34). Regulations restrict the application of techniques, such as somatic cell transfer to nonhuman cells, and all cloning projects have to undergo scrutiny by a committee of experts in ethics, medicine, and law. In addition the Council for Science and Technology, the country's principal science policy body, has proposed a legal ban on human cloning (34). According to media reports in Japan, the government is preparing to submit a bill to parliament based on the Council's recommendations, representing the first legal prohibition of life science research in Japan (33).

South Korea. In response to the December 1998 announcement by Korean scientists of the cloning of a human embryo, politicians are working to expand the 1997 rules adopted by the Ministry of Health and Social Welfare that cover genetic research that bans the implantation of genetically engineered human embryos, but not human cloning (35). Therefore South Korea's Parliament is now

considering legislation to ban cloning of human cells except for disease research (35). One proposal before the National Assembly gives the task of reviewing such experiments to a committee of representatives from government, religious groups, research, and industry (35).

Argentina and Belgium. Argentina has proposed legislation that is intended to deter efforts to clone human beings using somatic cell nuclear transfer. In Belgium, legislation covering medical ethics including cloning is currently being considered by parliament (20).

POLICY STATEMENTS AND ETHICAL GUIDELINES OF INTERNATIONAL ORGANIZATIONS

The possibility of cloning human beings has prompted responses from several international organizations. The Council of Europe, the World Health Organization (WHO), UNESCO's International Bioethics Committee (IBC), the Human Genome Organization (HUGO), the European Commission's bioethics advisory panel, and the Denver Summit of Eight have all called for a worldwide ban on the cloning of human beings (Table 6). (In addition to the original G7 leaders from the world's leading industrialized countries, Britain, Canada, France, Germany, Italy, Japan, and the United States, the 1997 Denver Summit of the Eight included Russia.) The policy statements of these international organizations are detailed below.

Council of Europe

On April 4, 1997, 21 countries associated with the Council of Europe signed an international convention, the *Convention for the Protection of Human Rights and Dignity with Regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine*, which calls for a ban on human cloning (36). (The countries that signed were Denmark, Estonia, Finland, France, Greece, Iceland, Italy, Latvia, Lithuania, Luxembourg, the Netherlands, Norway, Portugal, Romania, San-Marino, Slovakia, Slovenia, Spain, Sweden, Turkey, and Macedonia; http://www.coe.fr/oviedo/index.htm.) In addition, Article 13 of the *Convention on Human Rights and Biomedicine* prohibits interventions seeking to introduce any modification in the genome of any descendants and therefore, implicitly, forbids cloning of human beings including by use of somatic (nonreproductive) cells (36). The Convention is open for signature to the Council's 40 member countries as well as Australia, Canada, Japan, the United States, and the Holy See, which contributed to the drafting process. This text represents the first binding legal instrument ever drafted on an international scale with a view to safeguarding human dignity and fundamental rights against any improper applications of medicine and biology.

On January 12, 1998, representatives from 19 members of the Council of Europe signed an *Additional Protocol to the Convention on Human Rights and Biomedicine on the Prohibition of Cloning Human Beings* that committed their countries to prohibiting by law "any intervention seeking to create human beings genetically identical to another human being, whether living or dead" (37). The Protocol is limited to a ban on the cloning of human beings by embryo splitting or nuclear transfer. It does not prohibit the cloning of cells and it does not deal with the use of embryonic stem cells. Two European countries did not sign the Protocol. Germany claimed that the measure was weaker than a current German law that forbids all research on human embryos (38).

Table 6. Policy Statements and Ethical Guidelines of International Organizations

Organization	Policy/Guideline	Date	Synopsis
Council of Europe	Additional Protocol to the Convention on Human Rights and Biomedicine on the Prohibition of Cloning Human Beings	January 1998	Prohibits any intervention seeking to create human beings genetically identical to another human being, whether living or dead.
World Health Organization (WHO)	Resolution on Human Cloning (WHA50.37)	1997	Affirmed that the use of cloning for the replication of human individuals is ethically unacceptable and contrary to human integrity and morality.
UNESCO's International Bioethics Committee (IBC)	Universal Declaration on the Human Genome and Human Rights (29 C/Resolution 17)	November 1997	Prohibits practices which are contrary to human dignity, such as reproductive cloning of human beings.
Human Genome Organization (HUGO)	Statement on cloning	March 1999	States that there should be no attempt to produce a genetic "copy" of an existing human being by somatic cell nuclear transfer.
European Commission	Meeting at the Hague	June 1997	Called human cloning ethically unacceptable and should be prohibited by law.
Denver Summit of the Eight	Communique: The Denver Summit of the Eight	June 1997	The heads of state for the United States, Japan, Germany, England, France, Italy, and Canada, endorsed a worldwide ban on human cloning.

The United Kingdom did not sign because of its strong tradition of defending the freedoms of scientific research (38). Initially, the Netherlands also refused to sign the Protocol. However, following a debate in the Lower House of the Dutch Parliament, the Dutch government decided to sign the Protocol with the caveat that the term "human being" be defined as humans who are already born (39). The other countries that signed were Denmark, Estonia, Finland, France, Greece, Iceland, Italy, Latvia, Luxembourg, Moldova, Norway, Portugal, Romania, San Marino, Slovenia, Spain, Sweden, Macedonia, and Turkey.

World Health Organization

On March 11, 1997, Dr. Hiroshi Nakajima, the Director-General of WHO, issued a statement condemning human cloning:

> WHO considers the use of cloning for the replication of human individuals to be ethically unacceptable as it would violate some of the basic principles which govern medically assisted procreation. These include respect for the dignity of the human being and protection of the security of human genetic material (40).

However, other uses of cloning technology, such as animal cloning and the routine cloning of human DNA, genes, and cells, should not be banned (40). These uses of cloning technology hold the promise of advancing biomedical research on the diagnosis and treatment of diseases such as cancer, heart disease and diabetes.

In his statement the Director-General also referred to the guiding principles set forth in 1992 by the scientific group convened by the Special Programme of Research, Development and Research Training in Human Reproduction. The role of this group was to review the technical aspects of medically assisted procreation and related ethical issues. The group upheld "the right of everyone to enjoy the benefits of scientific progress and it applications" and the need "to respect the freedom indispensable for scientific research and creative activity" (40). They also stressed that "there is a universal consensus on the need to prohibit extreme forms of experimentation, such as human cloning, interspecies fertilization, the creation of chimeras and, at present, the alteration of germ-cell genome" (40).

In May 1997, at the meeting in Geneva, the Fiftieth World Health Assembly adopted a resolution affirming that "the use of cloning for the replication of human individuals is ethically unacceptable and contrary to human integrity and morality" (41). The Director-General was asked to clarify the potential applications of cloning procedures in human health and their ethical, scientific and social implications. This resolution was affirmed and upheld in 1998 at the Fifty-first World Health assembly (42).

In October 1998, a small working group of independent and government experts met at WHO headquarters to consider a report containing a first draft of guiding principles and recommendations to WHO and its Member States entitled *Cloning in Human Health* (43). The draft guiding principles were inspired by the basic principles of medical ethics, including beneficence, nonmaleficence, confidentiality, autonomy, equity and access to care for all, and were based on fundamental values such as dignity, human rights, and freedom (43). The draft guiding principles included subjects such as the need for public education on genetic research, the interaction of genes and the environment, the right to retain control over one's genetic material and the information derived from it, and gene therapy (43).

United Nations Economic, Scientific, and Cultural Organization

The *Universal Declaration on the Human Genome and Human Rights* was formulated in December 1996 by the United Nations Economic, Scientific and Cultural Organization (44) International Bioethics Committee (IBC). The Declaration received widespread support and was unanimously adopted on November 11, 1997, by UNESCO's 186 member States. On November 19, 1998, the 86 member countries of the United Nations Commission on Human Rights approved the Declaration, and on December 9, 1998, it was adopted by the United Nations General Assembly.

Article 11 of the Declaration addresses the issue of cloning of human beings. Article 11 states:

> Practices which are contrary to human dignity, such as reproductive cloning of human beings, shall not be permitted. States and competent international organisations are invited to co-operate in identifying such practices and in determining, nationally or internationally, appropriate measures to be taken to ensure that the principles set out in this Declaration are respected (44).

Human Genome Organization

In March 1996, about a year before Dolly was cloned, the International Ethics Committee of HUGO issued the *Statement on the Principled Conduct of Genetic Research* (45). The statement is concerned with research under the Human Genome Project (HGP) and Human Genome Diversity Project (HGDP). In its background principles, the statement refers to the "acceptance and upholding of human dignity and freedom." Cloning of human beings would violate these principles. In addition the cloning of a human being would violate a principle referred to in the statement's preamble that is concerned with the "reduction of human beings to their DNA sequences and attribution of social and other human problems to genetic causes."

In March 1999 the HUGO Ethics Committee issued its *Statement on Cloning* that makes specific recommendations on both animal and human cloning (46). The recommendations on human cloning are subdivided according to the purposes for which the cloning is carried out, reproductive cloning, basic research, and therapeutic cloning (46). The HUGO Ethics Committee makes the following recommendations:

- *Animal cloning.* Animal cloning should be subject to the same principles for animal welfare as other experimentation on animals, and possible consequences on biodiversity should be considered.

- *Reproductive cloning.* There should be no attempt to produce a genetic "copy" of an existing human being by somatic cell nuclear transfer. However, the use of somatic cell nuclear transfer may be supported if it is used to avoid a disease, such as an error in mitochondrial DNA.
- *Basic research.* In both humans and animals, cloning techniques should be supported to investigate a wide variety of scientific questions, including the study of gene expression and the study of aging.
- *Therapeutic cloning.* Research to produce cells and tissues for therapeutic transplants should be supported.

HUGO also states that the creation of human embryos should be considered for certain types of research that may be of widespread benefit to humanity, such as the development of embryonic stem cells (46).

European Commission

In June 1997 at a meeting at the Hague, the European Commission's bioethics advisory panel called human cloning ethically unacceptable and said it should be prohibited by law (47). The bioethics panel also specifically rejected the idea of embryo splitting in order to increase the success rate of IVF. However, the panel did recognize that cloning research might have important therapeutic implication such as in the study of aging and cancer, or the development of stem cells that could be used to repair or regenerate human organs. However, the European Commission must leave legislation against such experiments up to its individual member nations.

Denver Summit of the Eight

The Denver Summit of the Eight concluded their 23rd annual summit calling for specific actions on a host of economic, global, and political issues. The 18-page final communiqué, issued June 22, 1997, included a specific article related to cloning. Article 47 of the communiqué states that the G8 "agree on the need for appropriate domestic measures and close international cooperation to prohibit the use of somatic cell nuclear transfer to create a child" (48).

CONCLUSIONS

The cloning of Dolly has paved the way for major advances in biotechnology, reproductive medicine, and cell-based therapies. Before long, the preservation of genetically important strains and mutants of laboratory and farm animals, the preservation and propagation of rare and endangered species, and the unlimited multiplication of elite animals from selected matings will be routine. Combining cloning technology with transgenic techniques will provide an efficient way to produce animals that can make proteins in their milk that could then be used for pharmaceutical or clinical purposes. By genetically engineering cloned animals to express human proteins (e.g., histocompatibility antigens) on the surface of cells and organs, the risk of immune rejection in xenotransplantation may be significantly reduced. In addition, cloning technology may lead to the development of customized (e.g., autologous) human embryonic stem cells for use as cell and tissue-based therapies that would not be rejected by the patient's immune system. Many of these advances hold the promise of improved treatments for diseases for which there are currently no good alternatives.

While cloning techniques may one day provide improved treatments for diseases, revolutionize the production of biopharmaceuticals, and save endangered species, mammalian cloning does have its risks. In addition to high rates of spontaneous abortion late in pregnancy and death soon after birth, mammalian cloning has been linked to a developmental defect of the immune system and may be associated with premature aging. Thus, the question of safety remains and casts doubt on the future uses of mammalian cloning.

Beyond the safety concerns, the prospect of cloning human beings raises several other ethical concerns. These concerns have prompted calls for worldwide bans. Consequently language that directly prohibits the use of federal funds for cloning of human beings was included in appropriations legislation that prohibits the use of federal funds for human embryo research. In addition, five states have enacted legislation prohibiting cloning of human beings. FDA has also asserted its authority to regulate the cloning of human beings. Similarly, several other nations and international organizations have also enacted laws or issued policy statements prohibiting the cloning of human beings.

There appears to be broad international agreement that cloning of human beings for reproductive purposes should be prohibited. However, there is less agreement as to whether or not the use of cloning technology to develop novel therapeutic applications should be allowed. Some of the legislation and policies have specifically recognized the potential benefits of the use of cloning technology for therapeutic purposes. However, other policies are very broad and essentially prohibit any use of somatic cell nuclear transfer using human cells.

It is clear that the potential benefits that may be realized through the use of cloning technology are many. However, the potential for cloning a child is an issue that we will be grappling with for a long time to come. Therefore, responsible public policy will need to be crafted in such a way as to prevent the use of cloning technology for purposes for which it is found to be ethically unacceptable, while allowing for beneficial uses that hold so much promise for curing human diseases.

BIBLIOGRAPHY

1. N.L. First and R.S. Prather, *Differentiation (Berlin)* **48**, 1–8 (1991).
2. I. Wilmut et al., *Nature (London)* **385**, 810–813 (1997).
3. Y. Kato et al., *Science* **282**, 2095–2098 (1998).
4. T. Wakayama et al., *Nature (London)* **394**, 369–374 (1998).
5. A.E. Schnieke et al., *Science* **278**, 2130–2133 (1997).
6. A. Fitzgerald, *The Detroit News/Associated Press*, January 21, 1998.

7. R. Weiss, *The Washington Post*, August 20, 1998, p. A2.
8. M. Farley, *Los Angeles Times*, September 11, 1998.
9. P. Cohen, *New Sci.* **158**, 6 (1998).
10. R. Weiss, *The Washington Post*, October 9, 1998, p. A1.
11. N. Wade, *N.Y. Times*, November 12, 1998, p. A1.
12. J.P. Renard et al., *Lancet* **353**, 1489–1491 (1999).
13. P.G. Shiels et al., *Nature (London)* **399**, 316–317 (1999).
14. National Bioethics Advisory Commission (NBAC), *Cloning Human Beings: Report and Recommendations of the National Bioethics Advisory Commissions*, NBAC, Rockville, MD, 1997.
15. L. Andrews, *Cloning Human Beings: Report and Recommendations of the National Bioethics Advisory Commission*, vol. II, NBAC, Rockville, MD, 1997.
16. The Bureau of National Affairs (BNA), *Bur. Natl. Affairs* **3** (1998).
17. E.C. Price, *Harv. J. Law Technol.* **11**, 619–641 (1998).

17a. S.L. Nightingale, Letter from the Food and Drug Administration, Rockville, MD, October 26, 1998.

18. Australian Health Ethics Committee (AHEC) of the National Health and Medical Research Council, *Scientific, Ethical and Regulatory Considerations Relevant to Cloning Human Beings*, AHEC, Canberra, Australia, 1998.
19. Australian Academy of Sciences (AAS), *On Human Cloning: A Position Statement*, AAS, Canberra, Australia, 1999.
20. Human Genetics Advisory Commission (HGAC) and Human Fertilization and Embryology Authority (HFEA), *Cloning Issues in Reproduction, Science and Medicine*, HGAC, United Kingdom, 1998.
21. D. Butler, *Nature (London)* **387**, 218 (1997).
22. B. Knoppers, *Cloning Human Beings: Report and Recommendations of the National Bioethics Advisory Commission*, vol. II, NBAC, Rockville, MD, 1997.
23. International Digest of Health Legislation, *Int. Dig. Health Legislation* **44**, 256–257 (1993).
24. Human Genetics Advisory Commission (HGAC) and Human Fertilisation and Embryology Authority (HFEA), *Cloning Issues in Reproduction, Science and Medicine*, HGAC, United Kingdom, 1998.
25. S. Bashi, *The Associated Press*, December 30, 1998.
26. R.H.B. Fishman, *Lancet* **353**, 218 (1999).
27. Reuter, March 19, 1997.
28. Reuter, June 14, 1997.
29. Debates of the House of Commons of Canada, *Edited Hansard*, No. 179, February 11, 1999. Available at: *http://www.parl.gc.ca/36/1/parlbus/chambus/house/debates/179_1999-02-11/toc179-e.htm*
30. Tri-Council Policy Statement, *Ethical Conduct for Research Involving Humans*, MRC, NSERC, and SSHRC, 1998.
31. Associated Press, May 12, 1997.
32. G.K. Becker, *Eubios J. Asian Int. Bioethics* **7**, 175–178 (1997).
33. J. Lamar, *Br. Med. J.* **319**, 1390 (1999).
34. A. Saegusa, *Nat. Med.* **4**, 993 (1998).
35. M. Baker, *Science* **283**, 16–17 (1999).
36. Council of Europe, *Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine*, Council of Europe, Oviedo, Spain, 1997.
37. Council of Europe, *Additional Protocol to the Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine, on the Prohibition of Cloning of Human Beings*, Council of Europe, Paris, 1998.
38. J. Schuman, Associated Press, January 12, 1998.
39. B.I. Gordijn, *Professional Option: The Cloning of Human Beings. The Dutch Debate in an International Context*, Occasional Paper No. 5, Centre for Professional Ethics, University of Central Lancashire, Preston, Lancashire, UK, 1999.
40. World Health Organization (WHO), *WHO Director-General Condemns Human Cloning*, Press Release WHO/20; WHO, Geneva, 1997. Available at: *http://www.who.ch/press/1997who20.html*
41. World Health Organization (WHO), *World Health Assembly States its Position on Cloning in Human Reproduction*, Press Release WHA/9; WHO, Geneva, 1997. Available at: *http://www.who.ch/press/1997/wha9.html*
42. World Health Organization (WHO), *Ethical, Scientific and Social Implications of Cloning in Human Health*, WHA51.10, Agenda Item 20; Fifty-First World Health Assembly, WHO, Geneva, 1998.
43. World Health Organization (WHO), *Cloning in Human Health: Report by the Secretariat*, A52/12, Provisional agenda Item 13; Fifty-Second World Health Assembly, WHO, Geneva, 1999.
44. UNESCO, *Universal Declaration on the Human Genome and Human Rights*, UNESCO, Geneva, 1997.
45. Human Genome Organization (HUGO), *HUGO Statement on the Principled Conduct of Genetics Research: HUGO Ethical, Legal, and Social Issues Committee Report to HUGO Council*, HUGO, 1996, Available at: http://www.gene.ucl.ac.uk/hugo/conduct.htm
46. R. Chadwick, *Ebios J. Asian Int. Bioethics* **9**, 70 (1999).
47. R. Herman, *The Washington Post*, June 10, 1997, p. Z19.
48. Denver Summit of the Eight, *Communique: The Denver Summit of the Eight*, 1997.

See other CLONING entries.

D

DISABILITY AND BIOTECHNOLOGY

ANI B. SATZ
Princeton University
Princeton, New Jersey

University of Michigan Law School
Ann Arbor, Michigan

ANITA SILVERS
San Francisco State University
San Francisco, California

OUTLINE

INTRODUCTION

From the medical point of view, people are disabled when they are less functionally proficient than is commonplace for humans, and when their dysfunction is associated with a biological anomaly. Medicine traditionally has aimed at least to reduce, and preferably to cure, such dysfunction, and thus eventually to eliminate disability. There are four main ways in which biotechnology may be expected to help achieve this goal. First, biotechnology may prevent the inception of people who are biologically anomalous. For instance, technology derived from our increasingly accurate understanding of human biology identifies "at-risk" individuals who may be dissuaded from reproducing when apprised of their liability of having a disabled child. Second, biotechnology may prevent or protect people from being biologically anomalous and thereby becoming disabled. Technologies that immunize against disabling diseases, or that delay the disabling effects of aging, use this strategy. Third, biotechnology may repair anomalies by curing them, as do technologies which restore diseased persons to wellness. Fourth, biotechnology may mitigate a biological anomaly's impact by creating compensatory products, such as insulin for people with diabetes, or prosthetics, such as cochlear implants for people with nerve deafness.

Pervasive as the medical view of disability may be, people with disabilities often interpret their situation differently. They understand their limitations in terms of social rather than personal deficits. This social model of disability transforms the notion of "handicapping condition" from a biological state which disadvantages unfortunate individuals to a state of society which disadvantages an oppressed minority. The social model attributes the dysfunction experienced by people with disabilities primarily to hostile social arrangements. On this social view of disability, people who do not function in species-typical ways often are obstructed by socially constructed barriers. These range from discriminatory practice such as disability-based denial of employment to thoughtlessly inaccessible design such as the installation of steps rather than ramps. Sometimes the absence of adequate support services and health care benefits also is construed as a barrier to the effective functioning of people with certain kinds of impairments (1).

From this viewpoint, reforming social arrangements to achieve equitable opportunity and accessibility is the best route to reducing dysfunction in biologically anomalous people. From this perspective, the preeminent strategy of the medical model — namely altering biologically anomalous people to make them species-typical or normal — unfairly disparages personal traits central to the identity of people with disabilities. Further, by placing a premium on medically altering them so as to bring their modes or levels of functioning into better conformity with species-typical functioning, this medicalized approach to disability can be coercive and can expose the disabled to risky or ineffective medical interventions. To illustrate, Deaf Culture advocates believe that implanting cochlear devices in prelingually deaf children hazards their future by supplanting the proven effectiveness of communication in Sign language with a device whose success is unpredictable. They charge that such risky intervention is impelled by a mistaken idea, namely the unfounded assumption that living as a hearing person necessarily is better than living as deaf.

At least four fears converge to make disabled people suspicious about the promise of biotechnology. In the medical domain, the prospect that biotechnology can normalize disabled people may invite inadequately tested or coercive interventions and expose them to otherwise unacceptable levels of medical risk. In the social domain, this prospect may induce policy makers to promote medical strategies for altering individuals with disabilities over social strategies for accommodating them. In the political

domain, biotechnology may diminish the proportion of the population who are disabled and thereby may attenuate the political influence of disabled people and endanger the allocation of benefits and special services they need. The fourth fear, a philosophical one, sees biotechnology as an irresistible instrument for promoting homogenization and thereby reducing the diversity in capabilities that is now a feature of humankind.

Running through all four fears is anxiety about practices which assume that there is a biological mandate for functioning in the normal ways, that is, in the modes and at the levels of performance most common to the species. Indeed, this supposed biological mandate often is invoked to argue that restoring anomalous individuals to species-typical functioning is preeminently desirable, and thereby to procure public resources for doing so (2). In the experience of many people with disabilities, however, medical practice traditionally has been dominated by policies that distend the importance of species-typical functioning and consequently damage the disabled by devaluing their differences and discounting their alternative approaches of functioning. To illustrate, individuals in whom thalidomide caused congenital anomalies of the upper limbs report that, throughout their childhood, medical professionals forced them to wear dysfunctional prosthetic hands and disparaged the superior function they could achieve by using their feet to pick up and manipulate objects (3).

The problem is by no means new. Medicine has compiled a mixed record in treating the disabled. The progress medicine has made in reducing mortality rates from disabling disease and accident swells the numbers of people with disabilities. The historical record also reveals that the medical model of disability prompted the institutionalization of many biologically anomalous people merely because they did not seem normal (4). Medicine also imposed unnecessary suffering through worthless treatments aimed at making biologically anomalous people who functioned capably appear more normal. For example, walking, rather than wheeling, is the most common way humans gain mobility. Consequently it has not been unusual for individuals with walking limitations to be discouraged from atypical, but effective, compensatory modes of functioning and to be subjected to dangerous, ineffective surgery merely to try to make them mobile in the species-typical mode.

Biotechnology thus is threatening because it appears to make the pursuit of biological homogeneity an eminently practicable enterprise. In this regard disability advocates have urged that the obligation to protect the collective interests of the disabled, a minority group whose members definitively do not meet the standard of normality, takes precedence over the obligation to develop biotechnology that relieves some individuals of burdens arising from biological anomalies.

Because identifying as disabled quintessentially means experiencing one's self as an exception among normal people, people with disabilities may not agree that being normal is unquestionably valuable. Nevertheless, in the past, people with disabilities often acceded to medicine's traditional aim of normalizing them. From a disability perspective, molecular genetic medicine may make this goal more problematic because it has the capacity to effectuate a much more thorough program of normalizing than traditional drug and surgical interventions could accomplish. From a disability perspective, genetic technology's potential for thoroughly normalizing the population by reducing the natural variety in the functional capacities of humans increases the urgency of weighing its possible dangers.

DISABILITY AND GENETIC TECHNOLOGY

From its inception, the science of human genetics has aimed at enhancing people's lives. Reducing the incidence or impact of inherited disability, and thereby raising the aggregated level of human achievement, seemed an obvious policy for furthering this aim. While the benefits of such a strategy are apparent, many wrongs have been committed in its pursuit. Camouflaged by the claim that they were merely liberating unfortunate individuals from living out a destiny blighted by biologically predetermined disadvantage, eugenics programs targeted individuals with disabilities.

The excesses of eugenics programs committed to reducing disability make some people with anomalous inherited traits suspicious of any technology that facilitates genetic intervention. They are reluctant to permit some current, and many proposed, applications of genetic technology. Nevertheless, there are other people with disabilities who welcome new genetic technologies and invest in them. They believe that with recent developments in molecular genetic medicine, the science of human inheritance now is positioned to fulfill its promise in regard to alleviating the burdens of disability.

The Question About Disability

At the very least, history illustrates how often medical interventions intended to counter disability by elevating corporeal or cognitive capacity have harmed people who did not meet the desired standard. In this regard contemporary bioethical conversations about distinguishing beneficial from detrimental applications of genetic technology are wise to consider such deleterious interventions into the processes of human inheritance as the Nazi program for euthanizing "defective" Germans (5) or the U.S. practice of sterilizing "defective" Americans, notoriously endorsed by the Supreme Court in *Buck v. Bell* (6). Although many bioethicists see little resemblance between these historical incidents and contemporary applications of human genetics, disability activists believe that this history shifts the burden of proof to advocates of expanded usage of genetic technology.

When we move from the level of phenotype to the level of genotype, this debate translates into questions about the impact of molecular genetic medicine on people with disabilities. Specifically, do practices such as the termination of pregnancies because pre-natal testing is positive for genetic deficits in the fetus, or the alteration of fetal chromosomes by inserting genes needed to preclude genetic deficits, fall squarely within the benefits of therapeutic medicine? Or do they instead display the morally problematic aspects of the destructive eugenics programs of the past?

Relevant Genetic Technologies

Scientific literature refers to four categories of genetic technology with implications for disability: drug treatments and pharmacogenetics, cloning, genetic testing, and somatic-cell gene therapy and germ-cell genetic engineering. Advances in recombinant DNA technology have resulted in the capability for mass production of some gene products that may be used to treat disabling conditions. (Recombinant DNA technology allows the insertion of DNA into a bacterial or other host for duplication.) Insulin, for example, which now can be produced cheaply and with better control of the quality of the product, has clinical applications in the treatment of some forms of diabetes. Drugs are being developed to target specific genes for disabling conditions in order to suppress their expression, although such treatments are not yet readily available. Advances in our understanding of human genetics may also contribute to pharmacogenetics, or the understanding of how genes influence drug treatment for disabling and other conditions. Although cloning is commonly thought of as the production of one living organism from another, the technology also may be used to clone cells that will not develop into an organism. Thus, blood, tissue, and organ replacement in the treatment of disabling conditions may be possible through cloning.

Drug treatment enabled by genetic technology and pharmacogenetics, and the use of cloning to replace body parts both appear simply to be new ways to facilitate traditional therapies. Proteins and other substances, as well as organs or tissues, created in vitro can be used for the same treatment purposes as products obtained from human donors and other animals. The last two categories-genetic testing, and gene therapy and genetic engineering-may differ in kind from nongenetic therapies because they have the potential to eliminate disability altogether. These applications may evoke concerns propelled by recollections of the past eugenics programs that victimized people with disabilities.

REDUCING INHERITED DISABILITY: GENETIC TESTING

Genetic testing may be used to detect single gene, multigene (polygenic), or environmentally influenced (multifactorial) genetic conditions that are associated with disability. (Genetic screening refers to genetic testing programs involving either targeted populations or the general population, or to testing programs used to determine the need for further diagnostic testing.) Testing may be direct or through linkage analysis. The latter involves testing family members and identifying certain normal variations in genetic sequences called polymorphisms that serve as markers, indicating the potential presence of a genetic anomaly. The predictive value of information about disabling conditions generated by genetic testing is limited by variances in gene expression, false negatives and positives, and in the case of linkage analysis specifically, by genetic recombination causing the genetic marker or markers to separate from the disease gene. These technological limitations compound concerns about detecting and preventing disability because they indicate that information generated by genetic testing may be inaccurate or of limited predictive value. Despite these known uncertainties and imprecisions, our aversion to disability is so great that people who receive a positive result for a disabling genetic condition may be stigmatized.

Genetic testing is difficult to define because it is conceivable that many diseases have a genetic component. Further, analysis of some nongenetic material, such as metabolites, may furnish strong indication of a genetically anomalous condition. Tests that examine genes for mutations (DNA diagnostic tests), analyze gene products such as RNA, amino acids, proteins, and their associated enzymes, or identify the structure of chromosomes (cytogenetic diagnostic tests) are commonly viewed as genetic tests. Tests for metabolites such as blood sugar and cholesterol level may be viewed as genetic tests when they are highly indicative of mutations in single genes (7). In the United States, the genetic/nongenetic distinction is important for purposes of reimbursement and privacy protection. Currently reimbursement for commonly recognized genetic tests is limited under government and private insurance schemes. Privacy laws in some states afford special protection for information generated from genetic tests, possibly decreasing access to these tests as employers become more reluctant to provide coverage for them in their insurance benefit plans and to expose themselves to greater litigation risk.

There are several occasions on which genetic testing may arise. Each of these has implications for disability. Genetic testing may be used to diagnose (diagnostic testing) or predict a condition associated with disability in embryos (embryonic screening), fetuses (prenatal testing), or living persons (pre-symptomatic testing), or to predict carriers of such a condition (carrier testing). Disabling genetic conditions detected by these tests include Tay-Sachs, Duchenne's muscular dystrophy, cystic fibrosis, fragile-X syndrome, hemophilia A and Down syndrome. (Down syndrome is a congenital rather than a hereditary condition, though it is detected through chromosomal analysis. Risk may be identified prior to such analysis by maternal serum screening or chemical analysis of fetal gene product or biochemical analysis.) Genetic tests may also be used to identify genetic conditions such as deafness or anchondroplasia (the most common kind of dwarfism), which may or may not be viewed as disabilities.

Prenatal testing detects genetic conditions in fetuses. Studies indicate that prenatal testing that reveals certain genetic condition leads to abortion in most cases, although some expecting parents use the information to prepare to care for and support a disabled child. (In a compilation of international surveys, 73–100 percent of individuals chose to abort their fetuses when Down syndrome was detected, 100 percent made this choice for metabolic disorders such as Tay-Sachs, 95–100 percent for thalassemia, 38–63 percent for sex chromosome abnormalities, and 39–54 percent for sickle cell anemia.) (8). Abortion is considered therapeutic for some genetic anomalies such as Tay-Sachs disease, which causes children to suffer pain and severe deterioration of functioning and not to survive past the age of five. On the other hand, parents with conditions such as deafness or achondroplasia sometimes explicitly wish to select for fetuses with

the conditions they possess (9). They may believe, for instance, that they can most effectively parent children who are like themselves. Similar selection is possible through embryo screening, which involves testing embryos for conditions prior to implantation during assisted reproduction.

Diagnostic genetic testing confirms a suspected diagnosis or else eliminates the possibility of a genetic condition associated with disability. This contributes to more accurate identification of illnesses and avoids unnecessary or inappropriate drug treatment. By revealing the underlying genetic causes of some disabling conditions, developing diagnostic genetic tests facilitates research that may mitigate or cure such conditions. Diagnostic testing also identifies individuals in whom the effects of such conditions may be cured or mitigated. Newborn screening programs for PKU and congenital hypothyroidism are conducted throughout the United States for this reason. Similar motivations underlie diagnostic testing of adults. Once diagnosed, individuals with hemochromatosis may be stabilized by phlebotomy, individuals with cystic fibrosis may receive antibiotic and pulmonary therapies, and individuals suffering from the copper build-up caused by Wilson's disease may be placed in remission through treatment with chelating agents or zinc.

Presymptomatic genetic testing discovers individuals' predisposition for genetic conditions associated with disability. A negative result may provide comfort and reassurance, while a positive result could offer time for psychological, financial, and familial preparations for the onset of the condition and, when available, the opportunity to take prophylactic measures to delay or prevent the onset of dysfunction. Here again, testing seems to place disabling conditions within the context of medicine by focusing on prevention, preparation, and cure. Thus, where prophylactic or curative measures for a condition are available, testing for members of at-risk families is strongly recommended.

Carrier testing identifies individuals who do not have genetic impairments but who are carriers of certain genes for such impairments. This form of testing usually is conducted in order to allow for more informed family planning. Carrier screening programs — for example, within the Northern European-American population for cystic fibrosis, the Ashkenazy Jewish-American population for Tay-Sachs, the Mediterranean-American population for thalessemia, and the African-American population for sickle trait — have sometimes been implemented when there is an elevated risk of genetic anomaly. As presymptomatic individuals can be identified by carrier testing for certain late-onset conditions, such as Huntington's disease, the two forms of testing could raise similar ethical issues.

Are there Unique Benefits and Harms?

As a technology, genetic testing is sometimes thought to bestow unique benefits, or else to threaten unique harms, for people with disabilities. This is, in part, because genetic testing may occur within a medical setting, predict disability, generates shared information about disability, and have eugenic implications. Many of the genetic conditions associated with disability are the product of molecular anomalies we think of as diseases, but some are not. Acondroplasia and deafness are deemed disabilities, at least for purposes of protection under disability discrimination law, but they are neither diseases nor illnesses. Nevertheless, testing for all genetic anomalies occurs within a medical setting, either at a primary health center, such as a clinic or hospital, or at a clinical genetics laboratory, subsequent to physician referral. Considering genetic anomalies associated with disability to be diseases introduces possibilities of both benefits and burdens for individuals in whom the anomaly is detected, for their families, and for other people with disabilities. These benefits and burdens are not unique, however; they are imposed by genetic and nongenetic diagnostic technologies alike.

The general benefits associated with detecting present or future disabling conditions that are construed specifically as diseases include prevention, prophylaxis, and treatment. Detection within a medical setting may confer indirect benefits of clinical quality controls, genetic counseling, and physician fiduciary obligations. However, when genetic counseling is conducted in a climate of disability prevention, its neutrality may be so compromised that its benefits become questionable. (Consider the mission of genetic testing centers such as the Murdoch Institute in Victoria, Australia, which advertises, "[o]ur aim is to help every child to be born healthy and with normal abilities.") (10).

Three concerns arise in regard to understanding genetic disability as a medical condition. First, testing may promote the medicalizing of genetic characteristics that are not illnesses or impairments but are considered to be weaknesses or are otherwise thought of as undesirable. For instance, red hair might come to be considered a sign of being diseased because it is associated with an elevated risk of skin cancer. If so, having red hair might disable people from being hired for occupations that require them to work out of doors. Likewise, being irresponsible or aggressive or aloof might be counted as biological impairments requiring medical intervention, rather than as personal problems requiring character building, if genetic testing were to correlate them with genetic anomalies.

Second, individuals may feel obligated to prevent, cure, or treat inherited medical conditions because testing for them is available. They might be coerced into testing even if the conditions are valued, as achondroplasia and deafness are, within certain communities or families, and even if it is only in hostile social environments that individuals with these conditions are dysfunctional. They might, for instance, be thought irresponsible or be refused insurance coverage unless they submit to testing.

Third, opportunities to test may result in social pressure to eliminate mildly disabling genetic conditions that do not occasion significant dysfunction. Further compounding this concern is that testing for some conditions, such as Down syndrome and fragile-X syndrome, leaves the severity of the predicted dysfunction unclear. For example, Williams syndrome occasionally severely limits people who have the condition; however, in many cases people with Williams syndrome are better described as different rather than as dysfunctional.

The moral complexities of genetic testing are illustrated by the conflicting considerations weighed by people with achondroplasia. There is a fatal form of achondroplasia that occurs only when both parents are achondroplastic dwarfs; of the 10 to 15 percent of achondroplastic babies that are born to achondroplastic parents, one-fourth have the fatal, homozygous condition (inheriting the dominant achondroplasia gene from both parents), which is 2 to 4 percent of all achondroplasia births. Should genetic testing for achondroplastic fetuses be allowed? Should it be recommended? Because three-quarters of achondroplastic children are born to two average-size parents who are unprepared for them, fetal testing could result in dramatically fewer dwarf children being born. Nevertheless, such testing benefits dwarf couples who can avoid bringing to term high-risk pregnancies where the child does not survive. Some disability activists argue that the benefit genetic testing bestows on individual parents in a very small percentage of all dwarf births is not worth risking the collective future of the dwarf community. Although some people with short stature support this argument against permitting genetic testing, many reject it (11).

Reductionism and the Genetic Identification of Disability

Some commentators argue that testing for the genetic anomalies that cause physical and mental impairments ignores the fact that hostile social environments may contribute to the performance limitations associated with disability. This concern may be understood as a complaint either about reducing disability to biology or, more specifically, reducing disability to genetics (12). The criticism is leveled against procedures that suggest the source of the individual's limitation lies in herself rather than in the unfavorable way society treats people like her.

Claims of this kind fail to withstand further scrutiny. While genetic testing may identify an impairment as arising from a biological or genetic anomaly, it is a non-sequitur to suggest that doing so ignores how the social environment limits people with disabilities. Whether an individual uses a wheelchair because of accident or illness, genetic or nongenetic, is irrelevant to the concern that society limits the opportunities of such people by denying them access to education, employment, transportation, and both public and private places of commerce and accommodation.

Further, the charge that genetic testing promotes genetic reductionism by picturing people with disabilities as the victims of their genes is problematic. Most genetic tests, especially tests for polygenic or multifactorial conditions, do not predict genetic disease with certainty. They only identify predispositions for disease. Some tests for rare, highly penetrative, autosomal dominant conditions, such as Huntington's disease, are 100 percent predictive, though expression and age of onset, both very important to the definition of disease and the social impact of disability, vary. In this regard individuals with the Huntington's disease gene may have 35 to 50 years of existence free of disability. Our new ability to assess the probability of a currently disease-free individual's future disability raises questions about how to describe such individuals, as well as how to protect them against discrimination. While information that a person will develop Huntington's disease or has a high risk of familial Alzheimer's disease often suffices to stigmatize the individual, the U.S. Supreme Court appears to have narrowed protection against disability discrimination to the class of people who are presently rather than prospectively disabled (13).

Genetic reductionism generally is understood as the broad concept that genes determine who we are or what we will become (14). Even if a genetic test predicted expression and the onset of dysfunction with 100 percent accuracy and precision, it is wrong to imagine that a person's genetic condition necessarily undermines her self-conception, or that it is so inextricably entwined with her self-identity that it determines who she is and will become. While some individuals may identify strongly with their genetic makeup, the majority do not. Thus, even the ability to predict expression with precision does not entail genetic reductionism. It is true that some impairments may be so severe, or may be so socially stigmatizing, that they leave no room for conceptions of self that ignore the dysfunctional state. In such cases, a predictive test that identifies the genetic cause of the condition also reveals the genetic determinants of the disabled person's identity. However, it is the disabling condition itself, not the disclosure of its source or cause, that determines identity.

A remaining concern is that it is harmful to place disabling genetic conditions in a medical context because to do so invites coercive efforts for prevention or cure (15). Concerns of this nature have arisen with regard to deaf parents who prefer having deaf to hearing offspring. In the rich traditions and culture of the Deaf community, deafness, a characteristic deemed an impairment in the medical community, is viewed as a capability by deaf people (16). Deaf people who undergo genetic testing are much less likely to seek to prevent or treat their own deafness or deafness in their offspring than hearing people. Similarly, prospective parents with sickle cell disease are less likely than those who have not lived with the disease to abort fetuses affected with the same condition (17).

Thus genetic testing generates information that can be valuable to individuals, their offspring, and others in their care for many purposes other than prevention or repair. Genetic information may contribute to financial and psychological preparedness and may offer the comfort of taking control of one's current or future health state. It also often suggests the best means for reducing or avoiding the pain and suffering associated with disease.

Because of the high percentage rate of terminations of pregnancies following detection of genetic anomalies, prenatal testing constitutes the most troubling application of genetic testing. Even in this situation, however, the problem does not lie in the fact that the test identifies a genetic condition. Just as a genetic test may disclose that a fetus is likely to develop myotonic dystrophy, a genetic condition that eventually will limit use of the limbs, a nongenetic ultrasound may reveal a fetus with malformed arms and legs that also will limit use of the limbs. Termination of both pregnancies then arises from

the same belief, namely that the child will not have normal use of his appendages.

Predicting Disability

Predicting disabling genetic conditions raises the question of the moral relevance of predictive information about disability. Although one could loosely predict other causes for conditions that occasion impairment, for example, that those who do not wear seat belts or bike helmets may have a greater percentage of brain or spinal chord injury, these are at best correlative generalizations. Nongenetic diagnostics such as ultrasound may predict some physical deformities, genetic and nongenetic, but genetic testing generally is the most effective way of predicting disabling genetic conditions associated with individual genomes or familial gene pools. To illustrate, although achondroplasia can be detected by ultrasound, this technique cannot determine whether a fetus with two achondroplastic parents has inherited the gene from both parents, the fatal "double-dominant" condition occurring in a quarter of such pregnancies, or only inherited one parent's achondroplasic gene, a nonfatal condition.

Although predicting impairments may allow for their treatment or prevention, or preparation for the onset of the disabling condition, it has at least two morally troubling aspects. One is the effect of predictive information upon the autonomy of presymptomatic individuals. The second is the use of predictive information to prevent the birth of people with disabilities.

For presymptomatic individuals, predicting impairments may support autonomous behavior or hinder it. Some people find that predictive information preserves their autonomy by allowing them time to plan and prepare for the onset of a disorder to which they are susceptible. Others may find their autonomy is compromised, especially in a future or dispositional sense, because they must alter their life plans and restrict their choices, so as to prepare for life with a disability. The disparity between these alternative responses to the same news suggests that the impact of predictive testing upon autonomy is determined by the character of the recipient of the news, not by the character of the news itself.

Of course, nongenetic tests may also predict disability. Predictions about the effects of environmental carcinogens, exposure to contagious disease, and suffering from domestic violence may be made with respect to individuals or families. These predictions, like genetic predictions, invoke the same concerns over current and future autonomy of individuals of all ages. It is consequently difficult to establish that genetic testing presents a special threat to the autonomy of people with disabilities.

The most pronounced example of the impact of predictive information upon autonomy is demonstrated by testing children to learn whether they will become disabled. Testing the child and revealing the results to the child precludes the opportunity for that child to decide later, as an adult, whether she desires the predictive information. In this sense, testing children may violate both the child's current autonomy, if the child does not want to be tested, as well as the child's future or dispositional autonomy (18). Dispositional autonomy is reduced if the child is preempted from deciding, once she becomes a competent adult, whether she should have this information. Arguably, the best practice balances the detriments of paternalistic intervention against the benefits of prophylactics or available treatment. In instances where the child would reach the age of majority before the onset of disability, or where there are no measures to prevent or ameliorate impairment brought on by the condition, testing the child could be discouraged. The child could be tested if prophylactics or treatments are available.

Carrier and prenatal testing raise a different set of issues with regard to predicting disability. Although carrier testing might reveal genetic predispositions and raise concerns about autonomy as enumerated above, it is usually employed to detect individuals who carry deleterious genes but who do not exhibit the conditions therewith associated. Prenatal genetic testing reveals information about fetuses, which, regardless of debate about other aspects of their moral status, cannot be autonomous.

Disclosing the Results of Tests for Disability

Some of the concerns about carrier and prenatal testing grow out of fears that positive results will expose the subjects of such tests to various forms of social bias against disability. Genetic testing for disabling conditions may reveal shared information among biological relations. Although this is not a unique feature of genetic testing, as nongenetic information about contagious disease or exposure to carcinogens may be shared, it highlights some reasons for hesitating to share information about the results of tests for disabling conditions. Families may be discouraged from supporting their members who seek testing for fear that information about them could become public or that they could receive information about themselves that they do not wish to know.

The benefits and burdens associated with shared genetic information usually are discussed with respect to overlapping autonomy and privacy considerations. Autonomy in this context pertains to the right of an individual to know or not to know information about herself. This form of autonomy is also described as personal privacy. Informing an individual of her own predicted future disability may violate her privacy (19). It may drain her confidence in the feasibility of her life plans and in doing so may enervate her dispositional autonomy. Current and future autonomy may be constrained by the loss of her social identity as a fully-functioning individual (20). Or, if her identity has been formed by the possibility that she may have inherited a familial condition, learning that she has no great likelihood of doing so may demolish the basis on which she has planned to live her life. On the other hand, avoiding predictions of disability for fear that others will gain access to this news may conflict with one's need, as an autonomous individual, to learn enough about one's future to facilitate self-determination.

Certain features of contemporary techniques for predicting genetic disability bring the values of privacy and autonomy into conflict with the value associated with

preventing disability. If an individual believes that being tested herself might reveal the genetic condition of a family member who neither wants to know it nor wants anyone else to know it, she may be obligated to refuse to be tested, so as to respect the privacy and autonomy of another person. On the other hand, if the social value of preventing disability, or of acknowledging it, outweighs the individualistic values of privacy and autonomy, the person may be morally obligated to be tested in order to aid a relative who wishes to have linkage analysis for personal or family planning purposes, and may also be morally compelled to disclose information about other at-risk family members for family planning purposes or for the benefit of future caretakers.

These conflicting values raise complex questions about the significance of identifying or preventing familial disability that is occasioned by genetic anomalies, although the questions are not uniquely provoked by disabling genetic conditions. Detecting tuberculosis in one family member, for example, may benefit or harm other family members. It may result in treatment or prevention of a disabling condition, but it may also result in the individual being quarantined or otherwise isolated, or being denied employment or insurance because of her disability. In fact, one of the Supreme Court's earliest disability discrimination cases involved a person who was fired from her job because of a diagnosis of tuberculosis (21). Similarly, detecting cancer in a member of a family exposed to environmental carcinogens may indicate the likelihood of cancer in other family members. In this instance, the testing that confirms the cause of one family member's illness can expose other members to insurance and employment discrimination. So, while diagnosing a genetically occasioned disabling condition in one member of the family may have deleterious personal or social effects on other family members, once again there is no reason to believe that these problems arise from the nature of the test rather than the typical social responses to disability.

Genetic diagnostics may also serve as a vehicle for paternalistic medical intervention in the detection of disabling conditions shared within families. Physicians sometimes provide unsolicited information to individuals about their health status or that of other family members, believing that is in the best interests of the patient or of society. Physicians may suggest testing to one patient for the benefit of her family members who are also patients, where such a suggestion indicates that some or all of them are at risk for a disabling condition. Similarly, communities may urge testing on their members, as Cypriots do for thalessemia and certain Jewish communities do for Tay-Sachs disease. When recommendations for such testing are made in order to dissuade certain kinds of at-risk individuals from reproducing, or to dissuade individuals from carrying certain kinds of at-risk fetuses to term, the specter of negative eugenics is raised.

Genetic Testing as Negative Eugenics

Genetic testing that is aimed at preventing the existence of certain sorts of people carries the suggestion of negative eugenics. Negative eugenics programs aim to elevate the level of collective human performance by eliminating underachieving performers. Historically, these programs typically targeted people suspected of having inheritable inferior characteristics. The milder ways of practicing negative eugenics denied desirable employment, immigration, and other opportunities to people in categories associated with certain types of physical or mental limitations (22). The more menacing practices prevented these people from reproducing by prohibiting them from marrying, sterilizing them, or forcing them to terminate pregnancies (23). The most malignant practices euthanized people, both children and adults, whose performance limitations or behavioral infelicities were perceived as burdensome to themselves or injurious to society (24).

Genetic testing has consequences for disability in ways that broadly coincide with the familiar objectives of negative eugenics programs. There is no surprise here, for while genetic testing is not the sole way of facilitating eugenics, the advent of this genetic technology could permit these programs to be executed in a greatly refined way. A history of endorsing programs of this sort is the reason eugenics became identified as a genocidal practice conducted by dominant or strong classes and aimed at eliminating people, such as the disabled, who belonged to inferior or weak classes (25).

In the absence of knowledge about the causes of various impairments, virtually all members of certain disability categories were believed to suffer from limitations that would be inherited by their progeny. For example, because blindness was observed to run in some families, people who could not see were sterilized regardless of whether their blindness resulted from retinitis pigmentosa (a genetic condition) or opthalmia (an infection, sometimes acquired during birth). Genetic testing allows for somewhat more accuracy than was possible previously in identifying who will become disabled or will pass along a disability to future generations.

Nevertheless, genetic testing does not seem to have unique negative consequences for disability. Each of the concerns it evokes has an analogue in problems that previously have been found in the practice of medicine. Further, prenatal testing discriminates on the basis of disability only if its predominant use is to eliminate fetuses that are at risk for disability, and fetuses are accorded the moral status of persons (26). Still, putting genetic technology to this use exposes the entire field to the familiar fearful reactions provoked by negative eugenics. As Jonathan Glover remarks, "What is controversial is to eliminate or prevent ... disability by eliminating or preventing the existence of the person who has the disability. This controversial policy is the basis of screening programs" (27). Glover subsequently suggests that, while negative eugenics is destructive, there are positive eugenic applications of genetic technology that are not so. According to this view, uses of genetic technology that transform people with disabilities by alleviating or eliminating their functional differences are incontrovertibly positive and beneficial.

COMPENSATING FOR INHERITED DISABILITY: GENETIC ENGINEERING

Testing is not the only application of genetic technology that can reduce the proportion of the population that is disabled. Gene transfer technology has the potential to transform at least some persons whose disabilities are occasioned by genetic impairments into individuals who are temporarily or permanently free of disabling biological limitations, thereby promoting their existence. This new capability also promises to reduce the proportion of the population that is disabled, but not to do so through preventing the birth of or euthanizing individuals with genetic diseases or genetic anomalies.

Gene Transfer as Positive Eugenics

Programs that manipulate biological inheritance to promote, rather than prevent, the existence of certain sorts of people sometimes are characterized as positive eugenics. Selective breeding programs implement positive eugenics, as do some kinds of interventions that enhance congenital health or raise the levels of human performance. Unless we subscribe to the view that every person's existence substantively prevents the existence of others who otherwise might have replaced her, so that to engender a stronger or smarter child necessarily eliminates the weaker or duller person who otherwise would have been born in her place, positive eugenics does not collapse into negative eugenics. Neither do programs aimed at facilitating the flourishing of one kind of person necessarily disadvantage or damage other kinds of people. Thus, although positive eugenics programs may aim at genetically transforming people who have current or potential disabilities, such practice is not necessarily a form of disability discrimination (28).

Applying genetic technology to make people with inherited disabling conditions healthier can be a form of positive eugenics. Single gene anomalies appear to be obvious candidates for therapies that apply gene transfer technology. To illustrate, there is a hereditary form of retinoblastoma which causes multiple tumors in both eyes. Our current therapeutic interventions may damage the retinas or require removal of the eyes and thus may result in blindness. Moreover, retinoblastoma patients have an increased susceptibility to develop sarcomas in later life, either as sequallae of therapeutic radiation or chemical interventions, or as another manifestation of their genetic anomaly. In this form of genetic impairment, the Rb gene is missing from the chromosome (29). In principle, repairing the chromosome by adding the Rb gene would heighten the probability of long-term good results.

Achondroplasia, cystic fibrosis, Duchenne's muscular dystrophy, hemophilia A and B, Huntington's disease, and sickle cell disease are among the many other disabling conditions that have been attributed to single gene anomalies. As of this writing, the success of most gene therapy trials remains uncertain. However, there are many cases of genetically produced disability where addressing the underlying genetic anomaly conceivably could be more therapeutic than other approaches to curing or mitigating the disadvantages associated with the condition.

What is the ethical status of therapies aimed at normalizing such anomalies by "fixing" chromosomes or by otherwise changing them in order to mitigate their impact? Glover thinks there is a commonsense answer to this question: "When disorders are caused by the absence of a gene or the presence of a 'wrong' gene, it is attractive to think of inserting or deleting genes in embryos as required. . . .The day may come when this sort of gene therapy can be performed without harmful side effects. Choosing between a normal baby and one with a disability will become a genuine possibility" (27, pp. 128–129). In contrast to testing programs, Glover thinks, "One kind of intervention against disability is uncontroversially right. This is any treatment that does not prevent the existence of the person with the disability but aims to alleviate or cure the disability" (27, p. 129).

In a similar vein, LeRoy Walters and Julie Gage Palmer comment: "To people with disabilities that are diagnosable at the prenatal or preimplantation stages of development, the message of selective abortion and selective discard may seem . . . threatening. The message may be read as, 'If we . . . had known you were coming, we would have terminated your development and attempted to find or create a nondisabled replacement.' . . .[G]ene-therapy best accords with the health professions' healing role and with the concern to protect rather than penalize individuals who have disabilities" (30, p. 82).

Thus, at first glance, gene transfer therapy may seem to be the antithesis of a technique for eliminating currently low functioning members of the human collective. Far from doing so, the techniques of molecular genetic medicine promise to increase these individuals' presence in the population by countering genetic anomalies that heretofore have prevented the individuals in whom they are expressed from reproducing. Severe Combined Immune Deficiency (SCID) has prevented individuals with this condition from reproducing because it has killed them before puberty. To give another illustration, males with cystic fibrosis usually are infertile. By mitigating such effects, gene transfer therapies would increase the progeny of individuals with these inheritable conditions.

In addition to enabling them to reproduce, applications of gene transfer technology may also enlarge such individuals' opportunities for social participation by enhancing their capabilities for productive performance, for instance, by reducing or eliminating the physical deterioration characteristic of hemophilia, sickle cell disease, muscular dystrophy, and Huntington's disease. Currently the effectiveness of available therapies for such genetic conditions as cystic fibrosis, hemophilia, and sickle cell disease is limited. Rather than cure genetically anomalous individuals by removing the cause of their impairment, current nongenetic therapies merely slow or otherwise reduce the impact of the impairment.

To give another illustration, osteogenesis imperfecta (OI, or brittle bone disease) currently is treated with the traditional medical techniques for healing broken bones. In one form of OI, collagen molecules are not

anomalous but fewer than usual are made, while in another a dysfunctional collagen molecule is produced and incorporated into bone tissue. Current gene transfer research includes attempts to remove bone cells, alter them in vitro, and return them to the body; to apply the mechanism found in the first form of OI to reduce the production of dysfunctional molecules found in the second form; and to introduce functional collagen genes into cells (31).

As these illustrations indicate, existing therapeutic techniques for individuals with disabling genetic conditions often do not free them from their role as patients. In contrast, gene transfer technology has the potential to transform genetically impaired individuals into fully functional ones and thus to preclude the divisive conceptualization of strong and weak classes that drives negative eugenics (32). It therefore appears plausible to assess applications of gene transfer technology as potentially complying with the goals of positive eugenics, namely to promote the existence of certain people by making them more capable. In this scenario, the target group whose capabilities are to be enhanced through genetic transformation consists of those at risk for genetically induced disabilities, people who otherwise are in danger of living lives burdensome to themselves as well as others.

Despite the widely held view among medical professionals that gene transfer therapy holds comparatively little threat and some promise for people with disabilities, the fact that disability activists increasingly have mounted protests against at least some instances of it suggests that this technology is not as unproblematic as the authors cited above suppose. The reasons for the activists' alarm are subtle but worth considering because of the influence their views sometimes command, especially outside the United States. For instance, early in the era of recombinant DNA technology the Council of Europe called for "explicit recognition . . . of the right to a genetic inheritance which has not been interfered with" (33). Are there precautions and prohibitions that could be put in place to ensure that curative gene transfers do not abridge any rights people with disabilities have to their genetic heritage?

Prohibiting Germ-Line Alteration of People with Disabilities

Initially the obvious response would seem to be to ban any therapeutic protocol with the potential of affecting the patient's germ-line. Tom Murray proposes that "The most important question in the debate over the ethics of gene therapy is whether gene therapy is ethically distinctive from other forms of medical therapy. . . .The ethically distinctive element of gene therapy is only characteristic of germ-line manipulation" (34, p. 484). Some commentators agree with Murray. For instance, Glover and Walters and Palmer agree in holding it is always right to repair a disabling genetic impairment. However, Glover believes that we should refrain from any intervention that might alter the germ line, for fear of sliding into eugenics (27, pp. 134–135), while Walters and Palmer believe that we should encourage interventions that alter the germ-line in order to achieve eugenics. For Walters and Palmer, germ-line repair is obligatory if it advances "the effort to cure and prevent serious disease or premature death . . . the noblest of all human undertakings" (30, p. 85), while for Glover it raises controversial questions about the characteristics to be encouraged or discouraged, and incurs the risk of making mistakes about matters fundamental to other people's futures (27, p. 134).

Although, as of this writing, the National Institutes of Health endorses only therapeutic protocols that alter somatic cells (35) and the pharmaceutical industry has directed its efforts at somatic-cell alteration, there is no reason why germ-line engineering might not occur — if not intentionally, then by accident during an administration of somatic-cell therapy. Accidents like this can happen. For example, an early effort to repair defective eye color in fruit flies by inserting the gene for a missing enzyme inadvertently created a heritable repair (36). It also has been shown that when a retrovirus vector with a transgene is administered to fetal lambs, the lambs' offspring can inherit it (37). Somatic-cell therapies for conditions like retinoblastoma and adenosine deaminase deficiency might best be administered at the fetal stage because these conditions seriously compromise neonates, but doing so increases the probability of altering germ cells. Furthermore the success of somatic therapies may have outcomes that make subsequent germ-line intervention attractive to cost-conscious policy makers.

As LeRoy Walters observes, successful somatic-cell therapies eventually may permit many more individuals with dysfunctional biological conditions to live to an age at which they can reproduce (38). However, even in the absence of effective somatic-cell gene therapy for these specific conditions, medical advances in areas such as the mechanical assistance of breathing and the control of respiratory infections already have reduced early (pre-reproductive) mortality in individuals with such conditions as cystic fibrosis and spinal muscular atrophy, and presumably will continue to do so. Thus, whether prompted by the advent of successful somatic-cell therapies or by improvements in nongenetic therapies, such an eventuality will necessitate either greater health care resource expenditures due to the increasingly large number of individuals inheriting these conditions, or else the introduction of germ-line alteration that can transmit therapeutic benefits to succeeding generations. As Walters and Palmer argue, germ-line, but not somatic-cell, alterations could reverse this effect and reduce the incidence of inherited disadvantageous anomalies in the human gene pool (30, p. 81).

Parenthetically it sometimes is argued that the most benign approach consistent with such considerations of efficiency would be to eliminate genetically defective embryos, and to provide sperm or egg donation for parents for whom the probability of defective embryos is too high (39). Parents may quite reasonably prefer transmitting their own advantageous characteristics to their offspring, however, rather than risk the child's inheriting a less gratifying aggregate of traits from one or more donors, as long as they can avoid transmitting unfavorable genetic dispositions to their offspring. While germ-line alterations pose uncertain risk (as well as uncertain benefit) to future children, it is becoming clear that egg donation and egg reception procedures

also carry risks. Among other problems, we have not had the opportunity for longitudinal study of the effects on women of the various interventions that implement these procedures. Thus we cannot assume that germ-line alteration automatically puts people at more peril, or puts more people at peril, than reproductive technologies that are used for preventing the transmission of genetic defects by individuals who want to have their own (in some sense) biological children.

Contrary to Murray's supposition, then, if gene transfer technology is ethically distinctive from other forms of medical intervention, the ethical issues it raises apply to both somatic-cell and germ-line applications. One is not inherently more problematic than the other. For several reasons, including the pressure the success of somatic-cell fixes might bring to bear to pursue the permanence of germ-line repair, it is difficult to have confidence that the basic arguments for alteration of somatic-cells to treat genetic anomalies will ultimately be less persuasive in regard to the latter application. One is not inherently more or less morally problematic than the other. So, if there are purposes for which altering genes is morally suspect, the assessment is likely to be equally apt with respect to somatic-cell and germ-line intervention.

Pursuing Species-Typicality through Genetic Transformation

Consequently we may ask whether genetic alteration programs really differ from genetic testing programs in their eventual impact on disability. Both contribute to reducing the percentage of people who have disabilities. The prospect of doing so, by whatever method, provokes disability activists who insist that genetically altering individuals so as to surmount inherited biological flaws is tantamount to denying the moral worth of people with disabilities (40). Some also contend that altering their inherited traits changes who they are by denying them their identity as the progeny of ancestors who carried genes for these traits (41).

Further, they charge, programs that aim at adjusting the molecular genetic conditions occasioning these traits cannot help but promote an oppressively exacting standard of biological function. Such a standard, they say, is impossible for anyone with a corporeal or cognitive impairment to meet, and thus it threatens not only people whose genetic heritage incorporates a heightened potential for being corporeally or cognitively impaired (42), but whoever is currently impaired, whether or not the origin of the impairment lies in their genetic material. It is interesting to note here that concerns about the potential of gene transfer technology to elevate standards of human performance are not voiced exclusively by people with disabilities. In 1982 the U.S. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research assessed "interventions aimed at enhancing 'normal' people." These were considered to be problematic because "the difficulty of drawing a line suggests the danger of drifting toward attempts to 'perfect' human beings once the door of 'enhancement' is opened" (42, pp. 1–3).

The prospect is of a future in which individuals who come by advantageous biological traits as the result of natural processes will be rare. This is a future where it will be common to acquire genetically engineered biological advantages as a result of social forces that determine the development and allocation of the technology. Heretofore, in our democratically organized competitive society in which vigor, industry, and talent supposedly constitute the primary determinants of success, individuals' natural biological and moral endowments have been regarded as relatively impervious to the influence of social position and thereby as offering an antidote to artificially induced social privilege. Gene transfer technology purportedly makes humans' most fundamental biological characteristics so malleable that biological superiority could be realigned so as to be a product of, rather than a constraint upon, the privilege social rank or power bestows. From a perspective influenced by our growing power to uncouple individuals' destiny from their biological endowment, laboring under a biological disadvantage could become identified with inferior social rank. In other words, as the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research warned in 1982, genetic engineering could transform the "natural lottery" into a new kind of "social lottery," (42) one in which being left with an unrepaired genetic disability (and therapy lacking species-typicality) is the sign of a social loser.

Despite the confidence in the positive effects of gene transfer therapy expressed by such commentators as Glover, and Walters and Palmer, alleviating disability by altering genes continues to be controversial and to provoke concerned response from representatives of the disability community. Such interventions are seen as being motivated by the narrow and coercive standard of species-typicality. They thus are condemned as instruments of a social policy that promotes biological homogeneity to the detriment of people with disabilities. For instance, representatives of disabled people's organizations in 27 countries signed a statement "demanding an end to the bio-medical elimination of diversity, to gene selection based on market forces and to the setting of norms and standards by non-disabled people." The question then comes down to whether molecular genetic medicine must be governed by the goal of promoting species-typicality (43).

Norman Daniels's influential discussions of health care policy exemplify the view of medicine that is objectionable from a disability perspective. Daniels is persuaded that achieving common or normal functioning for all citizens is a moral imperative for medicine. Further, providing the resources to transform people with anomalous biological conditions to species-typicality is a civic obligation grounded in the principles of democratic morality (44). Daniels's interpretation of these principles entails that normality is a preeminent social value and thus that making everyone species-typical is a desirable social policy. In urging that democratic theory requires health care policy designed to ensure that all citizens exhibit baseline physical and mental normality, Daniels intends to promote interventions that improve particular people's lives by adjusting their health. The proposed trigger for such intervention is an individual's deviation from biological normality.

How clear is this account of health, and how fair is it as a standard for authorizing medically induced genetic alterations? To be in a healthy state is assumed by Daniels and many others to equate with being typical of one's group or species and, thereby, with being normal. Further, the functional organization typical of the human species is imagined to be the one best suited to meet our biological goals, so that to fall away from this standard is to be definitively disadvantaged in meeting these goals (42). These ideas lead to designing social arrangements to accommodate normal, not extraordinary, people. These or similar valorizing assumptions about the biological superiority of species-typicality are brought to bear to establish the moral importance of medically transforming biologically anomalous individuals into species-typical ones.

The assumptions promoting species-typicality as the goal of medicine are neither perspicacious nor unbiased. First, as to their clarity, discussions that take species-typicality to be a standard for medical intervention appear to conflate criteria that reflect three different levels of outcomes: standardizing biological states, engendering familiar modes of performing functions, and bringing about typical levels of functional outcome. Thus, it is not clear whether the goal of medical intervention is to standardize human genetic configurations, to ensure that people all can execute functions in the familiar ways, or to enable genetically diverse individuals to exhibit common levels of functional outcomes.

Second, because normal performance is defined with reference to familiar modes and levels of performance, what is thought of as normal often is artificially skewed by patterns of social domination that favor some kinds of performances over others and consequently ensure that those will continue to be the most common and seemingly normal ones. For instance, the domination of individuals for whom text is the most efficient conveyer of information has led to social arrangements that presume the ability to read text is normal. This bias causes people who perform the activities of reading and writing texts badly to be disadvantaged, even when they excel at aural, haptic, or pictorial communication. When this bias is medicalized, energy and expenditures are applied to developing therapies that can alter such individuals so that they can perform as normal readers and writers, whereas it may be more efficient for them to communicate in alternative ways. Similarly, in treating children exposed prenatally to thalidomide, medical professionals biased toward manipulating objects with upper rather than lower extremities insisted, to the disadvantage of their patients, that mechanical hands necessarily are better than fleshly feet for manipulating objects.

To suppose that anomalous performance must be functionally inferior to species-typical levels and modes of performance is to make two mistakes about human biological functioning. It is, first, to assume that human biological organization is functionally rigid, when it is instead immensely adaptive. Second, it is to assume that human social organization also is functionally rigid, when it can be flexible in expanding the opportunities of different kinds of people.

One way of grasping how the drive to normalize can lead to misperceptions about dysfunction is to note that acknowledging or ignoring differences in social context affects whether a genetic anomaly is counted as a genetic disease. To illustrate, it is well known that mild mental retardation does not disable women in low technology environments or simple societies in which a woman's role is to clean, cook, and bear children. Writing about equitable health care resource allocation in the very influential *The Global Burden of Disease*, Christopher Murray insists that, nevertheless, it would be inequitable to allocate resources to rich societies to prevent mental retardation but not to poor ones. Consequently, he thinks, we must assess the burden of mental retardation uniformly from nation to nation, regardless of significant national differences in how mental retardation affects people's lives. Murray acknowledges that the same impairment may be differentially disabling, or not at all disabling, depending on the environment. In fact, "in many cases," Murray says, "allocating resources to avert disability could exacerbate inequalities." Regardless of the realities of functioning in different social environments, he thinks, egalitarianism demands uniformity in assessing the burden of an impairment lest differences in context suggest that exalted functioning is reserved for the most privileged (45).

Contrary to Murray, however, there is no inequity in refraining from intervening where intervention has no benefit. What is inequitable is to treat people as dysfunctional when they are not, and on that biased basis to impose medical treatment that alters them. To the extent that medical perspectives assume species-typicality to be the goal, people with disabilities will be concerned about biased medical use of gene transfer technology.

So it is far from clear that we should build a dominating standard of normalization into the policy that governs the development of medical interventions. What is evident, however, are the dangers attendant upon policies that fail to disentangle natural functional disadvantage from artificial social disadvantage. Public policy should not privilege, and so fix and fortify, common modes of functioning by favoring medical practice that privileges the modes of functioning typical of our species over anomalous but comparably efficient modes. The prevalence of this kind of mistake in discussions about the deployment of medical technology is an understandable cause of wariness among people with genetic disabilities.

They have some reason to fear being subjected to genetic homogenization so as to relieve society from the responsibility of adjusting to their differences, even where their differences do not make them especially ill or dependent. Historically, accepting the dominant class's fashion of functioning as the biologically preferable mode has been a source of negative eugenics. To illustrate, Herbert Spencer, the social-Darwinist sociologist whose writing promoted eugenics, wrote about a group he labeled as "weak" because its members' "defective" biology prevented their functioning in the fashion of the dominant group. Rescuing them from their "natural" subservience would result in "a puny, enfeebled and sickly race," he warned in his repeated explanations

of how the biology of the female human made her inferior to human males. Translated into social policy, this perspective on women's differences resulted in an array of discriminatory exclusions attributed to "women's disabilities" (46). Imagine what mischief could have been done had it been possible, a century ago, to treat such disabilities by altering women so that their biological makeup became more like that of men.

Compensation, Enhancement, and Functional Diversity

This illustration suggests that distinguishing benign from baleful genetic alterations requires avoiding interventions made primarily for the purpose of homogenization. In this regard molecular genetic medicine must avoid confusing difference with dysfunction. In other words, interventions that alter genes should be governed not by the aim of imposing species-typicality but instead by the goal of enhancing functionality by whatever strategy is most effective.

The outline of this comprehensively flexible approach to disability was presented by the World Health Organization in the recently released beta revision of the International Categorization of Impairments, Disabilities, and Handicaps. For policy purposes, this document attempts to integrate the medical and social interpretations of disability. Here dysfunction is understood to emerge from a mismatch between the individual's mode and level of biological performance and the demands of the environment. No single strategy is made central to the medical interventions recommended for disability (47). The strategy of repair, which restores dysfunctional individuals to species-typical mode and level of functioning, is acknowledged to be one, but not the only, medical approach to disability. Compensatory strategies are equally recommended. A compensatory strategy differs from a reparative strategy in that effective functioning may be achieved without seeking or accomplishing restoration of species-typicality. In compensatory efforts, anomalies remain, but these need not compromise functional success.

Thus, for example, individuals who experienced successful postpolio rehabilitation often developed an exquisite sense of balance to compensate for one of the disease's typical sequellae, namely the marked disparity of strength between right and left sides, and upper and lower parts of the body. Similarly non-oral communicators often surpass most others in their ability to express themselves in body-language (48). Applications of biotechnology can be similarly compensatory. Gene transfer that boosts the low density lipoprotein receptor above normal range compensates for the effects of familial hypercholesterolemia, and gene transfer that induces capillary formation compensates for the effects of arterial blockage.

Whether all of an individual's modes of functioning are species-typical is not decisive for the individual's capabilities, for biological anomalies are not inherently dysfunctional. While restoring patients to genetic normality once seemed to be the obvious medical goal, responding to anomalous genetic configurations by enhancing the functionality of alternative modes of performance (to compensate for a deficit in another aspect of the individual's functioning) should be recognized as an equally promising strategy. Already of importance at the level of rehabilitative medicine, compensatory interventions appear to be similarly apt at the level of molecular genetic medicine.

There appears to be no medical basis for thinking that gene transfer applications designed to restore individuals to species-typical biological condition have a natural priority over compensatory interventions, that is, over gene transfer applications that improve the level of one biological performance to compensate for dysfunctional deficit in another (49). Further, although disability advocates sometimes fear that engineering their genes would threaten disabled people's identities (41), compensatory biotechnology no more does so than wheelchairs, sign language, or talking computers. All of these enhance the functionality of one kind of performance (arm movement, gesturing, listening) to compensate for limitations in another kind of performance (leg movement, speaking and hearing, seeing). It thus appears that compensatory genetic alterations are compatible with both medical and disability perspectives. However, a question remains about the ethics of using genetic technology to enhance human functioning.

It sometimes is argued that enhancing any biological function above its species-typical level, regardless of whether such an intervention is compensatory, brings about differences that are dangerous. In its influential analysis of the moral and social dimensions of "biologists' newly gained ability to manipulate. . .the material that is responsible for the different forms of life" the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research opposed "interventions aimed at enhancing" because these permit people to escape the limits imposed on them by the "natural" lottery (42).

Of course, no one appears prepared to abandon all attempts to apply medical technology to improve on nature. Immunization programs are a familiar example of a medical strategy that enhances our physical performance beyond the level that is native to our species. Immunization compensates for the fact that certain organisms or events can initiate damaging biological processes in the body by enhancing other biological processes so that they become capable of defending against such damage. Although now considered to be unexceptional and to provide the protection deserved by fragile populations like the elderly and the very young, vaccination originally was denounced as unnatural, immoral, and irreligious because it boosted individuals' resistance to disease above the level that then was typical of the species.

Not only have enhancements such as immunization become standard interventions for traditional medicine, but it has become unexceptional for gene therapeutic protocols to attempt to raise performance of one function above the common level to compensate for the adverse impact of other performances. For instance, research into genetic analogues of vaccination abounds (50). To give another illustration, one approved gene transfer protocol enhances the low density lipoprotein receptor to above normal to achieve an acceptably low level of cholesterol in the context of the genetic condition hypercholesterolemia.

This protocol is neither less natural nor more threatening than the SCID protocol that restores a species-typical genetic sequence in patients whose biological inheritance omitted it (51).

Ironically, refraining from interventions that enhance performance at the molecular level may further disadvantage those who have not been favored by the natural lottery. Promoting the primacy of natural limitations over biotechnically engineered enhancements is not necessarily fair to the disabled. To illustrate, at one time lower leg amputees were deemed noncompetitive because their prostheses made them run too slowly. Now, however, new materials and designs have created specially springy sports prostheses that permit their wearers, when very skilled and talented, to run faster than can be done with fleshly feet. So now, using these corrective devices is banned in competitive running to prevent unfairly disadvantaging nondisabled runners in the competition. Of course, some shapes and strengths of fleshly feet are better for running than others; yet athletes who have the best kinds of feet are not banned for disadvantaging the common-footed runner. Arguably, it is unfair to exclude runners when prosthetics render them uncompetitive, and also to exclude them when better prosthetics make them very competitive (52).

Similar issues conceivably can arise in molecular genetic medicine. To illustrate, let us imagine a gene transfer intervention that has been developed to decrease the deformation of erythrocytes in sickle cell anemia and thalessemia, but that also optimizes the hemoglobin in nondeformed cells. As Tom Murray has noted in respect to a similar kind of intervention, the procedure also might be effective when there is no diminution of species-typical functioning. The procedure could serve as a genetic version of blood doping, which temporarily increases the oxygen-carrying capacity of the blood. According to Murray, both blood-doping through infusion and its genetic analogues are ethically suspect in competitive sports (34). Does this consideration generalize so as to cast suspicion on interventions that enable individuals to perform better (on the molecular level) than is typical of the species in one respect, in order to compensate for their deficient performance in another respect?

Let us pursue this question in the context of the gene transfer technique described above. For people with certain genetic conditions, the intervention results in fewer erythrocytes deforming and in the hemoglobin being optimized in some cells that do not. Although more cells are species-typical than in the untreated patient, deformed cells remain, so the patient has not been restored to species-typicality. The erythrocytes have not been restored to species-typical performance because some of them perform below the species-typical level, while others perform above the common level.

Such a procedure is compensatory rather than reparative because it does not restore patients to a species-typical state. Whatever level of functionality is attained will result from the enhanced performance of some cells compensating for the limited performance of others. In other words, the performance of these cells will not be species-typical. Some patients will remain in deficit with respect to the oxygen present in their bodies, although on balance they will be better off, but for others the aggregated oxygen-carrying capacity of their blood will be elevated above that of the average person. Surely, however, the fact that some patients benefit from the intervention by attaining superior rather than species-typical functioning in no way makes it ethically suspect. Indeed, the same elevation of performance that might permit an athlete who is erythrocytically improved to triumph in a game might give an otherwise biologically typical individual the stamina to save another person's life, or an otherwise biologically compromised individual the energy to be productive in the service of others. In sum, whether prosthetic or genetic technology is in question, intervening medically to enhance human capabilities can be the morally appropriate thing to do.

CONCLUSION

Although bioethicists like Glover and Walters and Palmer assume that genetic interventions that alleviate or cure a disability are uncontroversially good, altering individuals for this purpose can be threatening from a disability perspective. The reason, however, is not that altering genes inherently endangers the disabled but that bioethicists and policy makers seem to focus on reparative applications that promote the standard of species-typical functioning. In the past, this standard has inspired practices oppressive to people with disabilities who, by definition, cannot meet it.

Reparative medical strategies aim to restore people to normal modes and levels of functioning. However, there seem to be no clinical reasons for making such an approach the preeminent strategy in molecular genetic medicine. In contrast to reparative interventions, compensatory medical strategies promote exceptional modes and levels of functioning to secure capabilities that otherwise would be in deficit because of an individual's biological anomalies. Thus the adoption of compensatory strategies in gene transfer interventions counters the homogenization that focusing on reparative genetic alteration tends to bring about. Indeed, as long as reparative approaches do not dominate compensatory ones, molecular genetic medicine is as capable of promoting as it is of preventing the existence of people with genetic anomalies.

From the medical perspective on disability, it is tempting to imagine, as Glover does, that the paradigm for genetic engineering is a reparative process in which people are made species-typical by inserting good genes or deleting bad ones. However, medicine itself is beginning to acknowledge the importance of compensatory medical strategies. The recent revision of the World Health Organization's International Categorization of Impairment, Disability, and Handicap exhibits the compatibility of compensatory medical strategies with disability perspectives. Further, compensatory approaches already constitute a familiar strategy in the development of gene transfer research.

Some bioethicists and medical policy makers discourage compensatory genetic interventions because these may promote exceptional rather than species-typical

functioning and may enhance performance above the common level. Their objections distract us from the question most important from a disability perspective, namely whether applications of molecular genetic medicine must be governed by a standard of normality that is inherently dangerous to people with disabilities. As we have seen, however, compensatory genetic alteration can improve disabled people's functioning without normalizing them. Thus gene transfer technology need not promote nor impose species-typicality, and consequently it is not inherently threatening to people with disabilities.

BIBLIOGRAPHY

1. A. Silvers, D. Wasserman, and M. Mahowald, *Disability, Difference, Discrimination: Perspectives on Justice in Bioethics and Public Policy*, Rowman & Littlefield, Lanham, MD, 1998.
2. N. Daniels, in D. Van DeVeer and T. Regan, eds., *Health Care Ethics: An Introduction*, Temple University Press, Philadelphia, PA, 1987.
3. E-mails from Bill Baughn, Theresia Degener, and Gregor Wolbring, January 11, 2000, January 12, 2000, and January 18, 2000//www.egroups.com/subscribe/Bioethics
4. J. Trent, *Inventing the Feeble Mind: A History of Mental Retardation in the United States*, University of California Press, Berkeley, CA, 1994.
5. M. Burleigh, *Death and Deliverance: "Euthanasia" in Germany c. 1900–1945*, Cambridge University Press, Cambridge, England, 1994.

   T. Duster, *Backdoor to Eugenics*, Routledge, New York, 1990, pp. 29–31.

   H.G. Gallagher, *By Trust Betrayed: Patients, Physicians, and the License to Kill in the Third Reich*, Holt, New York, 1990.
6. *Buck v. Bell*, 274 U.S. 200 (1927).
7. N.A. Holtzman and M.S. Watson, eds., *Promoting Safe and Effective Genetic Testing in the United States: Final Report of the Task Force on Genetic Testing*, Johns Hopkins University Press, Baltimore, MD, 1998, p. 6.
8. D.C. Wertz and J.C. Fletcher, *Clinical Obstetricts Gynecol.* **36**, 554 (1993).
9. D.S. Davis, *Hastings Center Rep.* **27**(2) (1997).

   A. Silvers, Hastings Center Rep. **27**(5) (1997).
10. Murdoch Institute, Victoria, Australia, Annual Report, 1997.
11. R. Ricker, Web site of Little People of America, Inc. 1995.
12. H.S. Reinders, *Bioethics Forum*, 3–10 (1996).
13. *Sutton v. United Airlines*, Inc., 119 S.Ct. 2139 (1999).
14. D. Nelkin and M.S. Lindee, *The DNA Mystique*, Freeman, New York, 1995.
15. C. Newell, in G. O'Sullivan, E. Sharman, and S. Short, eds., *Goodbye Normal Gene*, Pluto Press, Lane Cove, 1999, pp. 59–74.
16. R.A. Crouch, *Hastings Center Rep.* **4**, 14–21 (1997).
17. D. Hoar, *Biomedical Ethics and Fetal Therapy*, Wilfred Laurier Press, Calgary, Canada, 1988, p. 56.
18. A.B. Satz, Paper presented at the Australian Bioethics Association Conference on Autonomy, Community, and Justice at the University of Queensland, Australia, September 27, 1995.
19. R. Chadwick, in R. Chadwick, M. Levitt, and D. Shickle, eds., *The Right to Know and the Right not to Know*, Aveberry Brookfied, VT, 1997, p. 19–20.
20. Wiggins et al., *N. Engl. J. Med.* **327**, 1401–1405 (1992).
21. *School Board of Nassau County, v. Arline*, 480 U.S. 273 (1987).
22. M. Miller and Marvin, *Terminating the "Socially Inadequate": The American Eugenicists and the German Race Hygienists, California to Cold Spring Harbor, Long Island to Germany*, Malamud-Rose, Commack, NY, 1996.
23. T. Duster, *Backdoor to Eugenics*, Routledge, New York, 1990.
24. M. Pernick, *The Black Stork: Eugenics and the Death of "Defective" Babies in American Medicine and Motion Pictures since 1915*, Oxford University Press, New York, 1996.
25. D. Wertz, *Gene Lett.* **3**(2) (1999).
26. A.B. Satz, *Monash Bioethics Rev.* **18**(4), 11–22 (1999).
27. J. Glover, in J. Fishkin and P. Laslett, eds., *Justice between Age Groups and Generations*, Yale University Press, New Haven, CT, 1992, pp. 127–143.
28. A. Silvers, in R. Ankeny and L. Parker, eds., *Medical Genetics: Conceptual Foundations and Classic Questions*, Kluwer, Amsterdam, 2001.
29. D.R.J. Macer, *Shaping Genes: Ethics, Law and Science of Using New Genetic Technology in Medicine and Agriculture*, Eubios Ethics Institute, 1990, pp. 10–20.
30. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997.
31. Osteogenesis Imperfecta Foundation Home Page, Profiles in OI Research, *Dr. Rowe Focusing on Genetic Treatments for OI*, copy on file with author, bonelink@oif.org.
32. A. Silvers, in A. Donchin and L. Purdy, eds., *Embodying Bioethics: Feminist Advances*, Rowman & Littlefield, Lanham, MD, 1998, pp. 177–202.
33. Council of Europe Parliamentary Assembly, 1982, Recommendation 934; J.F. Keenam, *America*, October 20, 1990, 262–263.
34. T. Murray, *FASEB J.* **5**(1), 55–60 (1991). Reprinted in J. Arras and B. Steinbock, eds., *Ethical Issues in Modern Medicine*, Mayfield Publishing, Mountain View, CA, 1995, pp. 479–487.
35. Department of Health and Human Services, *Recombinant DNA Techn. Bull.* **9**, 221–242 (1986).
36. Office of Technology Assessment, 1984, 17–18.
37. *Science*, December 9, 1994, 1631; *Economist*, April 25, 1992, pp. 95–97.
38. L. Walters, *Nature* **320**, 225–227 (1986).
39. Arras and Steinbock: 428.
40. D. Kaplan, in K. Rothenberg and E. Thomson, eds., *Women and Prenatal Testing: Facing the Challenges of Genetic Technology*, Ohio State University Press, Columbus, OH, 1994, pp. 49–61.
41. Campaign against Human Genetic Engineering (CAHGE): 1998; e-mail distributed to disability-related listservs.
42. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life: The Social and Ethical Issues of Genetic Engineering with Human Being*, U.S. Government Printing Office, Washington, DC, 1982.
43. Disabled People's International, "The Right to Live and Be Different" Solihull, U.K., February 13, 2000.
44. N. Daniels, *Just Health Care*, Cambridge University Press, Cambridge, England, 1985.
45. C. Murray and A. Lopez, *The Global Burden of Disease*, Harvard School of Health, Cambridge, MA, 1996.
46. R. Miles, *A Women's History of the World*, London, Michael Joseph, London, 1988, p. 180.
47. World Health Organization: 1999, Versions of ICIDH-2, www.who.int/msa/mnh/ems/icidh/icidhtrg/sld033.htm

48. M. Corker, *Deaf or Disabled or Deafness Disabled?*, Open University Press, Buckingham, U.K., 1998, p. 48.
49. A. Silvers, in E. Parens, ed., *Enhancing Human Capacities: Conceptual Complexities and Ethical Implications*, Georgetown University Press, Washington, DC, 1998, pp. 95–123.
50. L. Morgenthaler, *Barron's*, September 20, 1993, pp. 10–11; H. Carey and G. McWilliams, *Business Week*, March 30, 1992, pp. 78–82.
51. Recombinant DNA Advisory Committee (RAC): 1997, "Discussion Regarding the Use of Normal Subjects in Human Gene Transfer Clinical Trials" Minutes of March 6–7, 1997, Bethesda, MD, National Institutes of Health.
52. A. Silvers, in L.P. Francis and A. Silvers, eds., *Americans with Disabilities: Exploring Implications of the Law for Individuals and Institutions*, Routledge, New York, 2000.

See other EUGENICS entries; REPRODUCTION, ETHICS, PRENATAL TESTING, AND THE DISABILITY RIGHTS CRITIQUE; REPRODUCTION, LAW, IS INFERTILITY A DISABILITY?

EDUCATION AND TRAINING, PUBLIC EDUCATION ABOUT GENETIC TECHNOLOGY

JOSEPH D. MCINERNEY
Foundation for Genetic Education and Counseling
Baltimore, Maryland

OUTLINE

INTRODUCTION

Of the many branches of modern science that compete for the public's attention and that claim to have significance for individuals and society, why should we expect that biotechnology, and particularly that aspect of biotechnology concerned with genetics and molecular biology, will be of interest to the average person? Consider the following. At the conclusion of his 1977 book *The First Three Minutes* (1), in which he discusses the nature of the universe immediately after the Big Bang, physicist and Nobel laureate Steven Weinberg makes the following assertion: "The more the universe seems comprehensible, the more it also seems pointless." There is, according to Weinberg, no purpose, no inherent plan, no evidence of design in the universe, even if there is great order amid the complexity. Weinberg's is not an isolated opinion among cosmologists (2), yet his profound challenge to humanity's deeply rooted belief in a guiding, supernatural hand went virtually unremarked in public discourse.

Two decades later, in the wake of voluminous, often sensational media coverage of the cloning of Dolly (3), the now-famous Scottish sheep, Chicago physicist Richard Seed declared his intention to clone a human, even if he had to leave the United States to avoid the applicable legal strictures on such research (4). Seed even offered to clone ABC's Ted Koppel, who interviewed Seed on *Nightline* (5). Despite the lack of any credible evidence that Seed had the knowledge, skills, or resources to pursue his fantasy, the reaction was rapid and strenuous, with scientists and social critics outlining the penalties for hubris and cautioning against "playing God." Even before Seed's pronouncement, the President's National Bioethics Advisory Committee met to prepare guidelines for research on cloning, and specifically placed cloning of humans out of bounds (6), even though the scientific and technical prospects for such research remain remote.

Why was there no public outcry about the challenges that physics presents to long-standing and cherished beliefs, no call to stop research in cosmology, while even the remote prospect of human cloning caused such an uproar? Perhaps the distinction lies in the methods and materials of cosmology as compared with those of biology. The remoteness and immensity of the universe are incomprehensible to most of us, especially when explanations of their origin and nature are packaged, as they often must be, in the complex language of mathematical models. True, each of us is made of "star stuff"—the elements that had their origin in the Big Bang and stellar evolution—but the stars matter little to us personally; we feel no kinship with them. Biology, however, is a different matter. Biology deals with living things, including us. This is stuff we know about, and there is a special quality to it, a quality that we have been taught to respect, even revere, notwithstanding that we cannot define its essence (7).

Public reaction to the steady and stunning progress in biology, and especially genetic technology, reveals a mixture of optimism and anxiety, the former driven perhaps by the promise of benefits to personal and public health, the latter by public ignorance of the technology and the underlying science and by the unsettling sense that progress in this arena consistently challenges many of society's most cherished values and beliefs. Many articles and books on the implications of genetic technology call for public education to help individuals and society understand and accommodate the application of new knowledge in genetics and molecular biology. Indeed, public education has been one of the central objectives of the Human Genome Project's (HGP) Ethical, Legal, and Social Implications (ELSI) program since the project's inception, and the ELSI five-year plan developed in 1998 (8) makes clear the need to expand and improve education for the public (see Table 1). Like most calls for education about genetic technology, however, the new ELSI plan provides little guidance about the content of such education, the instructional approaches, the desired outcomes of educational programs for the public, or even the reasons for educating the public about such a complex area of scientific and technological inquiry. This article provides an overview of those topics by addressing four significant challenges to public education about genetic technology and by recommending how educational programs can meet those challenges.

CHALLENGE 1: SELECTING THE CONTENT

As in other branches of science, the amount of information generated by research in genetics overwhelms even the discipline's practitioners. How then do we make the content and implications of genetics comprehensible to the nonspecialist? The answer to that question requires a clear understanding what we want the public to know about genetics in the first place, and why. Most often, that

Table 1. Ethical, Legal, and Social Implications (ELSI) Research: Excerpts from Goals and Related Research Questions and Education Activities for the Next Five Years of the U.S. Human Genome Project, October 1998

a. Examine the issues surrounding the completion of the human DNA sequence and the study of human genetic variation.

Examples of Research Questions and Education Activities:

- Will the discovery of DNA polymorphisms influence current concepts of race and ethnicity? (*e.g., How will individuals and groups respond to potential challenges to or affirmations of their racial and/or ethnic self-identification, based on new genetic information?*)
- What are the most effective strategies for educating health professionals, policy makers, the media, students, and the public regarding the interpretation and use of information about genetic variation?

b. Examine issues raised by the integration of genetic technologies and information into health care and public health activities.

Examples of Research Questions and Education Activities:

- What are the clinical and societal implications of identifying common polymorphisms that predict disease susceptibility or resistance? (*e.g., Will genetic testing promote risky behavior in persons found to be genetically resistant to particular pathogens, such as HIV, or environmental hazards, such as cigarette smoke?*)
- What are the best strategies for educating health care providers, patients and the general public about the use of genetic information and technologies? (*e.g., What are the most effective mechanisms for educating providers, patients, and the public about the uncertainties inherent in genetic risk information?*)

c. Examine issues raised by the integration of knowledge about genomics and gene–environment interactions into nonclinical settings.

Examples of Research Questions and Education Activities:

- What are appropriate and inappropriate uses of genetic testing in the employment setting? (*e.g., Are there conditions under which it might be ethical and/or legal to use genetic testing to identify those employees who may have a susceptibility to workplace hazards? What implications does the Americans with Disabilities Act have for such testing?*)
- What are the potential uses and abuses of genetic information in educational settings? (*e.g., Is placement of students on the basis of genetic data any more or less beneficial or harmful than tracking on the basis of traditional categories or classifications?*)

d. Explore ways in which new genetic knowledge may interact with a variety of philosophical, theological, and ethical perspectives.

Examples of Research Questions and Education Activities:

- Will continuing research in molecular biology and functional genomics affect how individuals and society view the relationship of humans to one another and to the rest of the living world? (*e.g., As new genetic technologies and information provide additional support for the central role of evolution in shaping the human species, how will society accommodate the challenges that this may pose to traditional religious and cultural views of humanity?*)
- What are the implications of behavioral genetics for traditional notions of personal, social and legal responsibility? (*e.g., What role will the discovery of putative genetic predispositions to violent behavior play in criminal prosecutions?*)

e. Explore how socioeconomic factors and concepts of race and ethnicity influence the use and interpretation of genetic information, the utilization of genetic services, and the development of policy.

Examples of Research Questions and Education Activities:

- How is the impact of genetic testing in clinical and nonclinical settings affected by concepts of race and ethnicity and other social or economic factors? (*e.g., Will particular communities and groups be more vulnerable to employment discrimination based on genotype?*)

From Ref. 8.

understanding is stated in the negative: The intent is *not* to turn all members of the public into specialists in genetics. Beyond that, one finds little agreement about the purpose or content of public education, save the assertion that the public should be able to make informed decisions about the personal and societal uses of new genetic knowledge. Some scholars (9), however, question the validity even of that argument as a general rationale for improved public science literacy or a guide for determining content.

Work by the Biological Sciences Curriculum Study (BSCS) provides a framework for defining the possible dimensions of public literacy with respect to genetic technology, and consequently, some guidance on the selection of content (10). According to this framework, *nominal* biological literacy consists simply of recognizing "certain words as belonging to the realm of biology." For genetic technology, nominal literacy would constitute the public's recognizing that terms such as gene, DNA, chromosome, and the polymerase chain reaction (PCR) are related to that area of investigation. In *functional* biological literacy, a person "can define certain biological terms or concepts but has limited understanding of or personal experience with them." The third level, *structural* biological literacy, implies an understanding of major concepts and the ways in which they are related. With respect to genetic technology, for example, a structurally literate person would understand that genes comprise DNA — a virtually universal information molecule in living systems — and, further, that DNA often is organized in chromosomes — cellular structures that help to maintain genetic continuity between generations of cells and generations of organisms. In addition a structurally literate person would understand that the ability to analyze and manipulate DNA with techniques such as PCR can have significant implications for individuals and society.

What does structural literacy require of the public in terms of content knowledge? Most important, the translation for nonscientists of content knowledge in genetic technology requires decisions about the level of detail, and the most difficult decisions concern what to leave out rather than what to include. Should

nonspecialists, for example, know the fine structure of DNA? Or is it sufficient that they understand that DNA is an information molecule whose content can be analyzed to provide insights into the construction and expression of human traits? Should education focus on the details of transcription and translation of genetic information, or on the variable expression of that information? Must nonspecialists know the structure of nitrogenous bases, or should public education highlight the extraordinary variation in the sequence of those bases? It is, after all, that variation, in concert with environmental variables, that confounds our ability to make definitive statements about the role of genes in complex human characters and about the eventual role of genetic medicine in improving human health (11,12).

Unfortunately, formal science education worldwide has been plagued by the tendency to subordinate the central concepts of science to a concentration on disarticulated facts. In genetics and molecular biology especially, the rate at which new knowledge is generated is so staggering that precollege and college curricula alike often are overwhelmed by the accretion of isolated details and extensive vocabulary that do little to help students form a conceptual picture of genetics or of biology (10,13,14).

Ironically, the concept most often ignored in public education about genetics is the concept that is at the heart of the discipline: variation. The centrality of variation in understanding genetics is reaffirmed by the HGP's focus on variation in its newest five-year plan (8), yet there is almost no mention of that concept in most educational programs about genetics, largely because of a concentration on relatively rare single-gene Mendelian traits, to the near exclusion of more common, multifactorial traits. In humans most single-gene traits are disorders, and a strict focus on them conveys inappropriately that the study of genetics is related only to disease. In addition an exclusive focus on single-gene traits conveys the misconception that human traits always are straightforwardly qualitative and that there always is a clear, direct relationship between genotype and phenotype. The standard textbook treatment of Mendelian genetics, for example, teaches only that one either has cystic fibrosis or does not, this disorder being a frequently used example of autosomal recessive inheritance. The same approach holds for those disorders generally chosen to illustrate autosomal dominant (Huntington disease) and X-linked (Duchenne muscular dystrophy) inheritance.

Single-gene disorders are not as invariant as most textbook treatments indicate, and the expression and severity of traits such as cystic fibrosis are often highly variable (15). Public education should emphasize that variability rather than leave the impression that all occurrences of a given disorder have the same degree of severity or the same natural history. This shift would help reinforce the message of variation and biochemical individuality (16) that is at the heart of genetics.

Although it is true that there is considerable variation in expression for any of the single-gene disorders cited above, it is nonetheless true that the presence of the mutant allele (or alleles) results in the generally recognized phenotype. Furthermore, in many instances, geneticists have a very clear sense of the biological relationship between gene and phenotype, and they often know the biochemical details of the relationship. It is clear, for example, that the accumulation of lipids that characterizes Tay-Sachs disease results from a deficiency of the enzyme hexosaminidase A, and that the range of symptoms associated with cystic fibrosis results from impaired transport of chloride ions across the membranes of certain epithelial cells.

It is more difficult to determine the biological relationship between gene and phenotype for multifactorial traits, including many common diseases. Although it is clear, for example, that genetic factors contribute to the risk for early onset heart disease, the exact relationship is yet unclear, as is the relationship between certain genetic markers and the risk of schizophrenia or bipolar disorder. In such common, complex diseases, expression is influenced by the products of multiple genes interacting — throughout one's developmental history — with a host of environmental variables whose influences are difficult to discern.

Given current trends in research, it is important for educational programs to demonstrate the distinction between the relatively rare single-gene disorders and the more common human traits, including many common diseases, that are polygenic and multifactorial (17):

> The term "complex disease" has been coined for conditions that arise from multifaceted interactions of environmental and heritable factors.... The heritability of these disorders deviates in important ways from that of classical (mono) genetic diseases: no simple Mendelian mode of transmission is apparent, and the severity of the disorder shows quantitative, unimodal variation rather than a dichotomous distribution. Complex traits are regarded as polygenic and multifactorial, with the phenotype representing the net effect of all contributing genes and environmental factors.

Educational efforts in genetics also should demonstrate that single-gene disorders are generally severe and often take their toll in infancy, childhood, or adolescence, while common, multifactorial disorders are generally less severe and tend to express themselves later in life. The later onset of common, multifactorial diseases provides opportunities for modification of environmental variables that otherwise magnify the risks inherent in genetic predisposition, and it is this prospect for prevention that is most likely to raise the prominence of genetical thinking among primary-care providers and consumers alike.

Education that addresses complex traits will help to acquaint the public with an important aspect of human genetics and will provide opportunities to help the public understand that genes and environment are both important in the expression of many human characters. That perspective will help counter the simplistic "gene for" approach to causation that often pervades media treatments of genetics (18). Any group of individuals — children, adolescents, or adults — is a living laboratory of human variation manifested as observable multifactorial traits, and educational programs should exploit that variation. Furthermore the quantitative traits manifest in any group of individuals often are those of most

interest to the public: height, weight, intelligence, and athletic or artistic ability, for example. Discussions of the expansive range of normal variation for these traits, the complex nature of causation, and societal perceptions of normality itself (19,20), can help prepare the public for the ethical and policy debates that must follow as continued research uncovers genes that are putatively associated with complex and controversial traits such as intelligence, aggression, or sexual orientation (21,22).

Finally, a focus on variation and on quantitative characters in addition to qualitative Mendelian traits may improve the public's ability to understand evolution (23), the central organizing theme of biology. Ernst Mayr (24) has written that perhaps the most important aspect of the Darwinian revolution is the replacement of "typological thinking" with "population thinking," that is, recognition that the members of any given species do not constitute a single, fixed type but instead are highly variable with respect to virtually all traits. Indeed, disease itself is a by-product of the genetic variation required for survival of the species. In certain individuals in certain environments, some variation is expressed as disease. Furthermore familiarity with evolutionary perspectives, and especially those related to human evolution, can help students understand the distribution of disease in human populations. The genetic variations associated with common diseases, which are present in all populations, have an older evolutionary origin than do the variations associated with more rare, single-gene disorders that aggregate in certain groups. The former variations arose before *Homo sapiens* migrated out of Africa and spread across the globe.

Just as genetic variation is the sine qua non of differential selection, population thinking is central to one's understanding of evolution. Because they focus on typology rather than on variation, current practices in genetics education may impede rather than enhance that understanding, and to the extent that genetics is a piece of biology education, the current, incomplete picture of genetics for the public — including pre-college students — may be a disservice. That is especially problematic given that about 95 percent of all students take high school biology and that for most, the high school course will be their last formal exposure to the life sciences (25).

With the foregoing discussion in mind, Table 2 provides a list of basic concepts that might serve to organize public education about genetic technology, in the hope of achieving structural literacy where this field is concerned.

CHALLENGE 2: EXPLAINING THE NATURE OF SCIENCE

Among the more difficult challenges the public faces as it struggles to understand genetics is the portrayal of the discipline in the media. Genetic themes — especially those related to manipulation of DNA — have become increasingly popular, pervading television (*The X-Files*) and movies (*Jurassic Park, GATTACA*) especially, and blurring considerably the lines between fact and fiction, the possible and the fantastic. This trend exposes what is perhaps the most disturbing deficiency in general science literacy: the public's failure to distinguish science as a way

Table 2. Proposed Set of Basic Concepts in Genetic Technology for Public Education

A. Concepts related to biological variation

1. Genetics is the study of biological variation. Genetic medicine is the study of human genetic variation that is associated with mortality and morbidity.
2. Individual genetic variation (biochemical individuality) results from the variable sequence of the four bases that are central components of the DNA molecule. Mutations introduce additional variation, although mutations rarely have biological significance. Some mutations can be deleterious, while other mutations can provide selective advantages that are central to evolution by natural selection. There would be no differential selection, and therefore no evolution, without mutation and variation. This principle helps to explain phenomena such as the emergence of bacterial strains that are resistant to antibiotics.
3. Human biological variation results from the interaction of individual genetic variation with environmental variables (the experiences that one accumulates during one's developmental history, from conception to old age).
4. There is no fixed type — no archetypical individual — in a species, including *Homo sapiens.* A species comprises a population of genetically unique individuals who vary in morphology, physiology, and behavior. Disease is a by-product of the genetic variation required for the survival of our species. Some variations are manifested as disease in some people in some environments.
5. The genotype for a given trait is the gene(s) associated with that trait. The phenotype is the expression of the genotype, which is influenced by the environment.
6. Some human traits result primarily from the action of one gene. Most human traits, however, are multifactorial, resulting form the action of more than one gene in concert with the influence of environmental variables.
7. The phrase "the gene for" can be misleading because it can imply that only genetic influences are responsible for a given trait (discounting the influence of the environment) or that only one gene in particular is associated with a given trait when there may be genetic heterogeneity.
8. Genetic medicine is uniquely positioned to provide insights into prevention because it acknowledges the individuality of each patient and the biological and environmental influences that produce that individuality. Genetic medicine does not focus on the disease, but rather on the individual. It asks, "Why does this person have this disease at this point in his or her life?" (12).

B. Concepts related to cell biology

1. Classic cell theory holds that all life is made of cells and that all cells come from preexisting cells.
2. Cells pass through a series of structural and functional stages know as the cell cycle. The cell cycle is under genetic control. Disruption of that control can lead to disorders such as cancer. In that sense, all cancer is genetic, but not all cancer is hereditary.
3. Cell division is the process that produces new cells.
4. Mitosis, one part of cell division, helps ensure genetic continuity from one generation of somatic cells to the next. Human somatic cells contain 46 chromosomes (the diploid number): 22 pairs of autosomes and one pair of sex chromosomes (X and Y).

Table 2. ***Continued***

5. Human germ cells, sperm and ova, contain 23 chromosomes (the haploid number). A special type of cell division — meiosis — occurs in the precursors to germ cells. Meiosis has two major biological effects: it reduces the number of chromosomes from 46 to 23 and it increases genetic variation through the exchange of genetic material (crossing over).
6. In humans (eukaryotes), cells contain a distinct structure (the nucleus) that includes the chromosomes, the carriers of most of the genetic material (DNA).
7. Human cells contain mitochondria. Because mitochondria likely were free-living prokaryotes early in the evolution of life, they carry their own DNA. This DNA in humans has been sequenced completely; mutations in mtDNA can cause health problems, often associated with neuromuscular diseases (because of the role of the mitochondrion in energetics).

C. Concepts related to classical (Mendelian) genetics

1. Our understanding of the behavior of chromosomes during first meiosis allows us to make predictions about genotype and phenotype from one generation to the next.
2. Some traits are inherited through an autosomal dominant pattern of inheritance, others through an autosomal recessive pattern. Still others, those traits associated with genes on the X chromosome, follow somewhat different patterns of transmission because the male has only one X chromosome.
3. Traits, not genes, are dominant or recessive, but we refer to genes as dominant or recessive because it is convenient.
4. Aberrations in the behavior of chromosomes during meiosis can result in structural or numerical alterations that have serious consequences for ontogeny and subsequent growth and development. Some of these aberrations are associated more frequently with advanced maternal age. We can detect many chromosomal aberrations prenatally.
5. Our understanding of the movement of genes through populations allows us to make predictions about the frequency of genes in given population groups and, therefore, about the frequency of disease phenotypes.
6. During the last two decades, research has uncovered genetic mechanisms that extend our understanding of inheritance and that provide biological explanations for heretofore unexplained observations. These mechanisms include imprinting, uniparental disomy, mitochondrial inheritance, and the expansion of trinucleotide repeats.

D. Concepts related to molecular genetics

1. DNA and RNA are information molecules; they store biological information in digital form in a well-defined code.
2. DNA is the primary information molecule for virtually all life on earth; this is evidence for the relatedness of all life by descent with modification.
3. The structure of DNA lends itself to replication. DNA replicates with great fidelity, which is critical to the maintenance of genetic continuity.
4. Sometimes mistakes arise during DNA replication. Evolution has produced mechanisms that repair such mistakes. In fact those mechanisms are conserved evolutionarily all the way back to *Escherichia coli*. When these mechanisms fail, mutations arise. Some cancers result from the failure of DNA repair mechanisms.

Table 2. ***Continued***

5. In most biological systems, the flow of information is: DNA–RNA–protein. The process by which this occurs is known as the central dogma of molecular biology: replication, transcription, translation (into protein).
6. DNA is a fragile molecule; it is easily damaged by a host of environmental insults. It is especially important to avoid such insults during pregnancy.
7. The damage that occurs to our DNA during the course of our lives can contribute to the onset of cancer.
8. A gene is a segment of DNA (although the segment may not be contiguous). Some genes code for the production of structural proteins (collagen) or enzymes (lactase). Other genes are regulatory, helping to control such processes as prenatal development.
9. A gene occupies a particular place on a chromosome — a locus. A gene can have two or more alternate forms — alleles.

E. Concepts related to new genetic technologies

1. Advances in technology allow us to analyze and manipulate the genetic material in ways that were not possible even a few years ago. These technologies include PCR, RFLPs, direct and indirect (linkage) analysis of DNA, and quantitative trait loci.
2. These technologies allow us to identify, isolate, and test for genes associated with disease.
3. Like all technologies, genetic technologies are fallible, can have unintended consequences, and often serve the interests of entities apart from the patient.
4. The growth of information technology in concert with the expansion of genetic technology is a great boon to genetic medicine and to basic research, but it also raises concerns about the use of genetic information.

of explaining the natural world from other explanatory systems, particularly those that appeal to supernatural events. This deficiency leaves the public susceptible to fanatics and pseudoscientific charlatans, ranging from radical animal-rights activists and creationists to purveyors of questionable cures for illness and disability. Even Pope John Paul II, in his 1998 encyclical, "Faith and Reason" (26), warns against the abandonment of reason and rational thought: "It is an illusion to think that faith, tied to weak reasoning, might be more penetrating; on the contrary, faith then runs the grave risk of withering into myth or superstition."

Because of the near-homogeneous — and generally incorrect — portrayal of science in science textbooks, the public may view the science as a set of invariant steps in "the scientific method." To help the public make sense of the rapidly moving science of genetic technology and of the ways in which new knowledge supplants old, educational programs must reflect more realistically what science tries to do. Above all, education should emphasize the habits of mind that distinguish scientific inquiry from other ways of knowing and should provide the public with the skills required to make sound judgments about the validity of information they encounter about genetic technology. The following material, excerpted from two programs produced by BSCS (27,28), proposes some major concepts that should pervade public education about genetic technology.

Concept 1: The laboratory seldom provides definitive or final answers to scientific questions.

Although laboratory investigations in molecular biology may provide concrete data, it is a mistake to assume that those data always provide final and immutable answers to complex questions about life on earth. Uncertainty — resulting from indeterminacy and from the emergent properties of organisms at higher levels of biological organization (29,30) — makes simple extrapolations from molecular data to complex characteristics difficult at best. Growing interest in functional genomics (8), which seeks to integrate our understanding of sequence data into the biology of whole organisms, reflects recognition among biologists that it is difficult to derive much helpful information about complex systems from an analysis of DNA sequences alone. One hesitates to enter into a protracted debate about the relative merits of reductionism in the context of educational programs for the public, but it seems only judicious that education about genetic technology include a strong caveat about assuming that we understand life on earth simply because we have access to its constituent molecules.

Concept 2: Scientific explanations are based on empirical observations or experiments.

Scientific inquiry assumes that the universe is explainable without appeals to supernatural phenomena. Evidence in science includes empirical data and existing explanations about related phenomena that are supported by independent data and are publicly observable. The fact that others can confirm or refute what one investigator claims to observe provides a crucial test of the scientific validity of that observation. Rumor, speculation, mystical experiences, and other unsubstantiated information are not accepted as credible scientific data. Public understanding of the distinction between credible and questionable information becomes ever more important as the amount of unvalidated information available via the Internet grows.

Concept 3: Scientific explanations are tentative.

Explanations can and do change. There are no scientific truths in an absolute sense, and scientists often suspend final judgment on the answers to scientific questions. Advances in genetic technology, in fact, have caused geneticists to revise their definition of a gene (31) and to extend their view of inheritance to encompass genetic mechanisms such as extranuclear inheritance, trinucleotide repeat expansions, and genetic imprinting. Additional research likely will result in additional revisions.

One key point to consider is the public's perception of the word "theory." Many people mistakenly think that this word means an ephemeral guess or a hunch; they see a theory as an unsubstantiated idea and, as in the case of the evolution/creation controversy, try to dismiss as "only a theory" what scientists recognize as a powerful explanatory framework. As scientists use the term, a theory refers to a large-scale explanation, or series of explanations, that describe the causes of many natural processes (32). An accepted scientific theory is very well substantiated by evidence, has been built logically upon valid assumptions, and has been tested extensively. A scientific theory is neither established nor refuted on the basis of personal opinion that fails to follow the discipline of scientific methods. Rather, a theory is an explanation of far-reaching significance so well tested and supported by such an abundance of credible evidence that it becomes a broadly accepted and fundamental scientific concept. There are a number of powerful theories in biology: for example, cell theory, chromosome theory, germ theory, and the theory of evolution. Each is supported by overwhelming amounts of evidence.

Concept 4: Scientific explanations are probabilistic.

A statistical view of nature, not an absolute view, is fundamental to science. The probabilistic view of explanations is evident implicitly or explicitly when stating scientific predictions of phenomena or explaining the likelihood of events in actual situations. Geneticists long have faced the problem of conveying a statistical view of nature to the public, and the impending identification of large numbers of human genetic variations will complicate the problem still further as geneticists try to explain to the public the expression of those variations in different environments.

Concept 5: Scientific explanations assume cause–effect relationships.

Cause–effect relationships are fundamental to making sense of phenomena, and much of science is directed toward determining causal relations and developing explanations for interactions and linkages among objects, organisms, and events. Distinctions among causality, correlation, coincidence, and contingency separate science from pseudoscience.

It is not easy to establish causality. People who do not employ scientific reasoning, however, sometimes mistakenly assume that one event is the cause of a second event solely because it precedes the second event. Science requires a much stronger association to establish a casual process. A classic example from folklore states that eating strawberries while pregnant can cause birthmarks on one's newborn child. The timing of the putative cause (eating strawberries) precedes the observed result (birthmark), but there is no credible and relevant evidence to support the first event as a *cause* of the second one.

Concept 6: Pseudoscientific explanations do not meet the requirements of science.

Pseudoscience can be defined as the promotion of unsubstantiated, allegedly scientific opinions. Some of these opinions may be very appealing to the public (33) and thus gain popularity, but they lack supporting, credible evidence. Pseudoscientific ideas have not been tested reliably. Often they are built on inaccurate premises, or they do not follow logically from what is observable. Pseudoscience often involves claims for which it is almost impossible to provide scientific evidence. Just as a poorly formulated hypothesis cannot easily be tested, pseudoscientific claims usually are so vague, ill-formed,

and undetailed that they make no specific predictions and cannot be tested through credible experiments.

Ideas based in pseudoscience generally are inconsistent with other, well-tested concepts. Indeed, pseudoscientific claims may not even be internally consistent, but without rigorous criteria for evidence and reasoning, proponents of pseudoscience are not likely to recognize the inconsistencies. We may find ourselves drawn toward the ideas put forth by pseudoscientists because they have an emotional appeal, but when we do so, it is often because we are being intellectually lazy. Science is not easy, but it does provide sound results. New scientific findings, including those in genetics, sometimes challenge comforting, long-standing assumptions about the natural world, but scientists must go where the data take them even if the destination is a bit unsettling.

Concept 7: Science cannot answer all questions.

Some questions simply are beyond the realm of science. Questions involving the meaning of life, ethics, and theology are examples of questions that science cannot answer. Genetic technology, for example, might help us determine the processes by which *Home sapiens* arose from its earliest mammalian ancestors and the chronology of descent with modification, but it is powerless to determine why we are here, in the sense of ultimate causes (34).

It is important, however, to distinguish for the public those questions that science cannot answer from those that science likely can answer, but has not yet. The origin of life is a case in point. Science cannot explain why there is life on earth, and it has yet to provide a complete explanation for life's origin. Failure to provide a completely naturalistic explanation for the latter question does not, however, mean that scientists should throw up their hands in resignation and ascribe the origin of life to supernatural causation. Indeed, such explanations are inadmissable in science, and scientists assume that the problem of life's origins ultimately will yield to the same methods of inquiry and habits of mind that have successfully stripped the mystery from so many of nature's mechanisms.

Concept 8: Science is not authoritarian.

Ecologist Garrett Hardin (35) reminds us that "science is ineluctably married to doubt." Although the public may view disagreements among scientists as evidence that the science in question is somehow flawed, disagreements and multiple competing hypotheses are essential to the health of the scientific enterprise.

The collective nature of science encourages objectivity in the field. Scientists must report their findings, and communication must use standardized descriptions so that results are meaningful to any informed person to whom they are communicated. Science is not monolithic; it attempts to be authoritative, but it is not authoritarian. Certainly recognition of an "authority" in a particular area of study encourages other scientists to pay attention to what is reported by that individual or research group, but the requirements for supporting evidence and all of the other criteria of valid science still apply. A famous name in itself does not establish disciplined methods nor produce the credible evidence required to substantiate an idea presented to the scientific community. The rules of science are the same for everyone, and anyone who proposes a new explanation for a natural phenomenon ultimately must answer the same two questions from any other scientist: "Where are your data?" and "How do you know they are sound?"

Table 3. Snapshot of the Goals and Methods of Science

1.	**What does science try to do?**
	Science tries to provide causal explanations for natural phenomena.
2.	**How do we know that causal explanations are on the right track?**
	The causal explanations demonstrate predictive power. Other observations (evidence) rule out competing causal explanations.
3.	**What counts as evidence in science?**
	Empirical data are included as evidence in science. Existing explanations about related phenomena are included if they are supported by independent data.
4.	**What makes science objective?**
	Science is conducted by a rigorous set of methods. Authority and fame by themselves are not sufficient to establish the scientific validity of an explanation.
5.	**How and why does scientific knowledge change?**
	Scientific explanations always are open to change. New evidence may show that an existing explanation is inadequate and that it needs to change.
6.	**What is the difference between pseudoscience and science?**
	Pseudoscience fails to meet the intellectually rigorous requirements of science. Internal consistency sets scientific knowledge apart from pseudoscience.

From Ref. 27.

The establishment of valid scientific explanations depends on review and critique by the scientific community at large, but the critique begins with the individual scientist. Scientists are trained to be critics of their own work. Indeed, British biologist Peter Medawar noted that "most of a scientist's wounds are self-inflicted" (36). In addition, at any given time, a minority view on some important scientific concept usually exists. With time, if sufficient evidence is brought to light, even an unpopular view can be validated.

Table 3 provides an overview of the goals and methods of science that can serve as guidance in the development of educational programs in genetic technology.

CHALLENGE 3: THE PRINCIPLES OF TECHNOLOGY

Genetics may be the science of the future, but lay persons almost never will encounter that science directly. They are much more likely to encounter genetically based technologies, ranging from chorionic villus biopsy and PCR to FISH and pre-implantation diagnosis. This distinction is especially true of the Human Genome Project, which

is driven by a wide range of technologies, from those employed in mapping and sequencing to those central to the creation and maintenance of international genomic databases (37). The details of the underlying science are likely to remain inaccessible to the average person, partially because those details are extremely complex, but also because they are unimportant for the nonspecialist.

The technology-dependent nature of genetics and of genetic medicine highlights a more recent challenge to public education: the need to include technology as a focus of serious study (13,14,28). If cherished tenets of medical ethics such as informed consent and nondirective counseling are to be more than topics of discussion at genetics meetings, public education must acquaint potential consumers of genetic medicine with some basic principles of technology. The AAAS, in its publication *Benchmarks for Science Literacy* (13), provides helpful guidance about those concepts of technology that should be central to public education. The brief discussion below paraphrases several of those concepts and demonstrates their relationship to education about genetic technology.

Concept 1: Technologies extend our senses.

Many technologies associated with science help us see, hear, or measure objects or phenomena that we would be unable to experience otherwise. We must, however, understand the limitations of our technologies and the role of inference in the interpretations of information we derive from them. For example, we can infer the presence of mutations in the genes for muscular dystrophy or cystic fibrosis on the basis of restriction fragments displayed on Southern blots. We can neither see nor touch the mutations themselves, however, and the accuracy and predictive value of our technologies are partially a function of the soundness of our inferences.

Concept 2: Technologies often have unintended consequences.

Almost all technologies are developed for specific purposes, yet many have side effects that are unintended, and worse, undesired. DNA analysis, for example, provides insights into long-standing biological questions such as gene regulation and evolution and allows us to detect genes associated with disease. But DNA analysis also has raised questions of privacy and discrimination to levels heretofore thought unimaginable (38,39).

Furthermore the impact of unintended consequences multiples rapidly with the introduction of public-health initiatives such as voluntary or mandatory genetic screening and testing. The sickle cell screening programs of the early 1970s, which resulted in discrimination against heterozygotes, demonstrated well the unintended consequences of technologies applied in the absence of thorough planning, education, and counseling. The more recent screening programs for Tay-Sachs disease benefited from those errors and reduced unintended consequences.

Concept 3: All technologies are fallible.

With respect to medical genetics, the public must know that diagnostic, laboratory, and treatment techniques can fail for reasons that range from those inherent in the technologies themselves to those associated with human error. When technologies fail in the context of personalized genetic medicine, the results can be tragic for individuals and families. When they fail in the context of broadly based public-health initiatives such as genetic screening, the negative consequences are likely to be much more pervasive. The public must be aware that the extent to which society embraces technologies such as genetic screening and testing is the extent to which society also embraces the risk of technological failure.

Concept 4: All technologies serve the interests of particular individuals, groups, or agencies.

Sometimes those interests compete. For example, genetic testing for breast or colon cancer can identify individuals at risk for those diseases, but insurance companies might use the same information to restrict coverage for those found to be at risk (40). Similarly population screening for carriers of mutations for cystic fibrosis can identify at-risk couples. Screening programs, however, can enrich companies that provide the tests and analyses, and those companies might push for wholesale screening before the laboratory tests meet appropriate standards.

CHALLENGE 4: THE PERSONAL AND SOCIAL IMPACT OF SCIENCE AND TECHNOLOGY

Progress in genetic technology demonstrates that we need public education not only because we can do new things (e.g., detect mutations associated with breast cancer or compare base sequences between humans and chimpanzees), but because the new things we can do raise profound, sometimes troubling questions for individuals, families, and society (41). Although the HGP has not necessarily raised issues of ethics and policy that are new to genetics, it has accelerated the rate at which once-hypothetical issues are likely to become reality. Furthermore the HGP is likely to result in more widespread public awareness of genetics in general and in increased use of molecular medicine and its clinical applications (11,42).

Such complex issues challenge scientists and educators to provide serious and rigorous educational treatments of ethics and public policy. These treatments should address concrete applications of genetic technology, such as genetic screening and testing, and as genes come to be associated with specific behaviors, conceptual issues such as notions of normality, societal views of what it is to be human, perceptions of free will, and even biological and cultural perceptions of race (8,43–45).

Many educational programs address ethical and policy issues related to genetic technology, and experience indicates that the most effective instruction includes the features described below.

Feature 1: A clear recognition that controversy is inherent in such instruction, and clear recommendations for dealing with it.

Progress in genetic technology invites controversy almost as a matter of course. Whether the issue is genetic screening for predisposition to breast cancer or the

development of chorionic villus sampling, which allows first-trimester detection of certain birth defects, genetic technology challenges traditional values and traditional views of the world. The incredible rate of progress in science and technology ensures that there always will be great disparity between what is possible and what people find acceptable.

The public generally encounters two major categories of scientific controversy: debates within the scientific community and debates about the use of science and technology that extend into society as a whole. The public must understand that debates between scientists are essential, or there will be no scientific progress. Science is a dynamic, self-correcting enterprise that continually tests new information and ideas in open, sometimes confrontational, debates. Debates within the scientific community demonstrate that the concepts under scrutiny are intellectually viable.

The current debate between evolutionary biologists who espouse gradualism and those who espouse punctuated equilibria is a good example, as is the dispute over whether *A. africanus* is ancestral to *A. afarensis,* and vice versa. Neither debate questions the validity of the theory of evolution. The former is a debate over the pace of evolutionary change, the latter, a disagreement over the sequence in which our hominid ancestors diverged. Each debate has a healthy effect on evolutionary biology because the scientists involved must work harder and be more creative and insightful to establish sound arguments.

Although debates within the scientific community may or may not attract public attention, other issues derived from biological progress are certain to do so, and the public encounters a number of those issues with respect to genetic technology. Such issues call into question many long-standing values and moral traditions, but the attendant controversy does not necessarily mean that those values and moral traditions will be found wanting. It does mean, however, that new knowledge and new techniques raise what once were intellectual abstractions to the level of hard, often painful, reality for individuals, families, and policy makers.

Feature 2: A clear conceptual framework for ethical analysis.

Sound instruction about bioethics must specify the underlying criteria for making and evaluating arguments to reduce the likelihood of unproductive discussions. The criteria need not be esoteric, and science educators, in conjunction with specialists in ethics, have developed a number of frameworks for ethical analysis that work effectively with nonspecialists, including the general public and precollege students. For example, the use of competing interests or contrasting goals, rights, and duties can focus analysis and highlight sources of disagreement.

Feature 3: A clear structure for discussions.

Ethical analysis and argument are forms of public discourse, but there is a misperception among nonspecialists that ethical discourse consists of a rather free-form sharing of ideas among the participants. Although that view is incorrect, the prospect of such unstructured discussions often is unsettling, especially to educators who are justifiably uncomfortable with deliberations that may have no structure or resolution. Productive ethical analysis requires structure and an objective. A well-structured model for discussion and analysis can help participants make arguments that lead to important insights and conclusions, while learning that there can be competing, well-made arguments about complex aspects of an area such as genetic technology. Even more important, a well-structured model for discussion can help the public realize that even seemingly intractable issues are amenable to analysis, and that civil, respectful discourse is essential if one wants to understand conflicting views.

Feature 4: A clear understanding that ethical analysis is a form of rational inquiry.

Instruction in bioethics and related policy questions often asserts that there are no correct answers to bioethical dilemmas. That statement may be correct, but it does not go far enough. Well-designed instruction in bioethics conveys clearly that there are well-reasoned and badly reasoned arguments, just as there are in science. Sound ethical analysis, for example, does not permit conclusions that are unsupported by the facts of the case, any more than science would allow such conclusions.

Feature 5: A clear demonstration of the connection between ethics and public policy.

Ethics is vital to public policy because it provides the concepts and terminology for the carefully organized debate that can result in well-reasoned conclusions about what society should or should not do. This inquiry is valuable in and of itself. Once society identifies a well-reasoned conclusion, however, it is reasonable to ask whether it should be enacted into formal public policy. Sometimes the best response is *not* to enact new policies in response to a controversy but rather to allow individuals, institutions, and society to act in the manner they choose.

One example (46) of an educational framework that helps make the connection between ethics and policy uses the conditions of urgency, means, and effectiveness to assess whether any conclusion of a well-reasoned ethical argument should become formal public policy. This analysis (see Table 4) assumes that it is not reasonable to enact new policy if conclusions from ethical arguments do not satisfy those conditions. Instead, ethical inquiry should continue, and public policy should remain at the de facto level.

CONCLUSION

As knowledge of genetics expands, perhaps the greatest challenge to public education will be the integration of new data into a cohesive picture of biology for the average person, much as the greatest challenge of the HGP will be the integration of the complete physical map of three billion bases into a cohesive understanding of the biology of *Homo sapiens.* The scientific community must play a central role in this integration, working with educators to determine how best to make this difficult, sometimes disconcerting discipline understandable to the average

Table 4. Framework for Connecting Ethics to Public Policy

Condition 1.
The situation is urgent: There is immediate risk of serious, far-reaching, and irreversible harm if legislation is not enacted or an existing law is not changed.
Immediate means that there are reasonable scientific grounds to conclude that impairment of interests will occur in the near future, for example, access by unauthorized individuals or institutions to human genome databases will deny or violate individuals' rights of privacy.
Serious means that the risk involves potentially grave injury to interests, for example, some private insurance carriers may deny coverage to an individual who has a genetic predisposition to a particular disease.
Far-reaching means that the impact of severe impairment of interests may be widespread, for example, a public policy of mandatory genetic testing of all pregnant women or children at birth would threaten the interests of millions of individuals.
Irreversible means that the serious damage to interests likely will be permanent, for example, labeling infants as having a genetic predisposition to learning disabilities could well have a permanent impact on their future education.
Conditions 2 and 3.
There are effective means to address the urgency of the situation: Scientifically valid or technologically practical means are available to prevent, reduce, or avoid risk of serious, far-reaching, and irreversible harm. The public policy must work and be enforceable — resources must be available to implement the public policy. An example would be that there are enough scientifically and technologically qualified individuals to carry out populationwide genetic screening. Public policy works when it enjoys broad-based political acceptance, that is, when few, if any, individuals or groups disagree strongly with the public policy. An example of a policy that is controversial in this respect is the requirement that all job applicants submit their individual genetic profile to prospective employers. Public policy is enforceable when few, if any, individuals or groups will disobey the public policy. An example of policy that is controversial in this respect is a requirement that all citizens submit their individual genetic profile to the federal government for inclusion in a national database.

From Ref. 46.

person and to situate genetics in the larger context of life on earth and, especially, of human biology. Because genetics and its associated technologies will continue to raise difficult questions of ethics and public policy, scientists and educators also must develop mechanisms for the rational consideration of those questions and must help develop in the public at large the skills and knowledge essential to informed, dispassionate analysis.

The translation of science for the public has never been easy. We must overcome the inherent difficulty of the information itself and, especially in genetics, the anxiety that arises when new information forces us to confront and perhaps revise long-standing assumptions about what it is to be human. The good news is that the public appears endlessly interested in genetics, and especially in its applications to personal and public health.

BIBLIOGRAPHY

1. S. Weinberg, *The First Three Minutes*, Basic Books, New York, 1977.
2. P. Davies, *The Last Three Minutes: Conjectures About the Ultimate Fate of the Universe*, Basic Books, New York, 1997.
3. I. Wilmut et al., *Nature (London)* **385**, 810–813 (1997).
4. J.B. Elshtain, *The New Republic*, February 9, p. 9 (1998).
5. K. Kestenbaum, *Science* **279**, 315 (1998).
6. National Bioethics Advisory Commission, *Cloning Human Beings: Report and Recommendations*, NBAC, Washington, DC, 1997.
7. E. Mayr, *This Is Biology: The Science of the Living World*, Belknap/Harvard University Press, Cambridge, MA, 1997.
8. F.S. Collins et al., *Science* **282**, 682–689 (1998).
9. J. Gregory and S. Miller, *Science in Public: Communication, Culture, and Credibility*, Plenum Trade, New York, 1998.
10. Biological Sciences Curriculum Study (BSCS), *Developing Biological Literacy*, BSCS, Colorado Springs, CO, 1993.
11. L.B. Andrews, J.E. Fullarton, N.A. Holtzman, and A.G. Motulsky, eds., *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994.
12. B. Childs, *Genetic Medicine: A Logic of Disease*, Johns Hopkins University Press, Baltimore, MD, 1999.
13. American Association for the Advancement of Science (AAAS), *Benchmarks for Science Literacy*, Oxford University Press, New York, 1993.
14. National Research Council (NRC), *National Science Education Standards*, National Academy Press, Washington, DC, 1995.
15. L.B. Jorde, J.C. Carey, and R.L. White, *Medical Genetics*, 2nd ed., Mosby-Year Book, St. Louis, MO, 1998.
16. C.R. Scriver and B. Childs, *Garrod's Inborn Factors in Disease*, Oxford University Press, Oxford, UK, 1989.
17. K. Lindpaintner, *N. Engl. J. Med.* **10**(332), 679–680 (1995).
18. D. Nelkin and M.S. Lindee, *The DNA Mystique: The Gene as a Cultural Icon*, Freeman, New York, 1995.
19. M. Rutter and R. Plomin, *Br. J. Psychiatry* **171**, 209–219 (1997).
20. P. Kitcher, *The Lives to Come: The Genetic Revolution and Human Possibilities*, Touchstone Books, New York, 1996.
21. R. Plomin, J.C. Defries, G.E. McClearn, and M. Rutter, *Behavioral Genetics*, 3rd ed., Freeman, New York, 1997.
22. J.D. McInerney, *Q. Rev. Biol.* **71**(1), 81–96 1996.
23. M.U. Smith, H. Siegel, and J.D. McInerney, *Sci. Edu.* **4**, 23–46 (1995).
24. E. Mayr, *The Growth of Biological Thought*, Belknap/Harvard University Press, Cambridge, MA, 1982.
25. Council of Chief State School Officers, *State Indicators of Science and Mathematics Education*, Council of Chief State School Officers, Washington, DC, 1995.
26. Pope John Paul II, *Fides et Ratio*, The Vatican, Rome, 1998.
27. Biological Sciences Curriculum Study (BSCS), *The Puzzle of Inheritance: Genetics and the Methods of Science*, BSCS, Colorado Springs, CO, 1997.
28. Biological Sciences Curriculum Study (BSCS) and Social Sciences Education Consortium (SSEC), *Teaching about the History and Nature of Science and Technology: A Curriculum Framework*, BSCS and SSEC, Colorado Springs, CO, 1992.

29. E. Mayr, *Toward a New Philosophy of Biology: Observations of an Evolutionist*, Belknap/Harvard University Press, Cambridge, MA, 1988.
30. M. Blois, *N. Engl. J. Med.* **318**, 847–851 (1988).
31. P. Portin, *Q. Rev. Biol.* **68**(2), 173–223 (1993).
32. A. Rosenberg, *The Structure of Biological Science*, Cambridge University Press, Cambridge, UK, 1985.
33. C. Sagan, *The Demon-Haunted World: Science as a Candle in the Dark*, Random House, New York, 1995.
34. National Academy of Sciences, *Teaching about Evolution and the Nature of Science*, National Academy Press, Washington, DC, 1998.
35. G. Hardin, *Am. Zool.* **25**, 469–476 (1985).
36. D. Pyke, *Perspect Biol. Med.* **39**(4), 555–568 (1996).
37. N.G. Cooper, ed., *The Human Genome Project: Deciphering the Blueprint of Heredity*, University Science Books, Mill Valley, CA, 1994.
38. M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997.
39. T.H. Murray, M.A. Rothstein, and R.F. Murray, Jr., eds., *The Human Genome Project and the Future of Health Care*, Indiana University Press, Bloomington, IN, 1996.
40. J. Beckwith and J.S. Alper, *J. Law, Med., Ethics* **26**, 205–210 (1998).
41. B. Gert et al., *Morality and the New Genetics*, Jones & Bartlett, Sudbury, MA, 1996.
42. F.S. Collin, *N. Engl. J. Med.* **341**, 28–37 (1999).
43. J.D. McInerney, *Am. Biol. Teacher* **60**(3), 168–173 (1998).
44. M.R. Gorman, *Perspect. Biol. Med.* **38**, 61–81 (1994).
45. K.J. Kevles and L. Hood, eds., *The Code of Codes: Scientific and Social Issues in the Human Genome Project*, Harvard University Press, Cambridge, MA, 1992.
46. Biological Sciences Curriculum Study (BSCS) and the American Medical Association (AMA), *Mapping and Sequencing the Human Genome: Science, Ethics, and Public Policy*, BSCS, Colorado Springs, CO, 1992.

See other entries FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, ELSI; FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, NATIONAL BIOETHICS ADVISORY COMMISSION; PROFESSIONAL POWER AND THE CULTURAL MEANINGS OF BIOTECHNOLOGY; PUBLIC PERCEPTIONS: SURVEYS OF ATTITUDES TOWARD BIOTECHNOLOGY.

EUGENICS, ETHICS

GLENN MCGEE
DAVID MAGNUS
University of Pennsylvania
Philadelphia, Pennsylvania

OUTLINE

INTRODUCTION

Francis Galton, an Englishman and cousin of Charles Darwin, who synthesized the word from the Greek *eugenes* (wellborn), coined the term "eugenics" in 1883 (1). Galton described a "science of improving human heredity over time," through the systematic, social and even governmental application of human knowledge about the hereditary roots of desirable and undesirable traits. The history of the twentieth century eugenics movement has been widely chronicled, and it is associated with what were eventually thought of as nefarious political aims. However, policies and arguments that could be described as eugenic both predate and antedate that movement. With every significant advance in reproductive genetic technology, fears of eugenics are revived in social and political institutions around the world. Eugenics is both rooted in the history of biology and tied to contemporary debates, and is thus always complicated by its past and future.

HISTORY

That eugenics is a difficult concept to define is in part a result of its disparate origins and uses. Long before Galton described systematic eugenics, cultures had devised strategies for the regulation of reproductive relationships. Society has always exercised a measure of control over reproduction: in sexual recombination, it takes two to reproduce, and those two choose each other under the influence of family, economic, political and other community values. We learn what counts as attractive, successful, and desirable within the ethos of a community that has models of the successful family and of health. At many times and in various places that ethos has been fairly emphatic. Long before Galton, different cultures were telling families not to have children or with whom reproductive behavior is authorized.

Early versions of what would later be called eugenics were deployed in three ways, each of which was designed to prevent reproduction: exposure or infanticide, abortion, and sterilization. These three techniques are distinguished from milder, more positive means of regulating reproduction in only two ways. First, agreement to terminate or forestall pregnancy is achieved by a political or social instrument, such as a law sanctioning sterilizations or a medical protocol for therapeutic abortion. Second, the techniques involve surgical intervention into the bodies of citizens, rather than acts of verbal coercion.

History reports numerous cases in which societies discussed and employed both clinical and nonclinical techniques. The Spartans left their unwanted offspring to the elements. Plato wrote in the *Republic*: "those of our young men who distinguish themselves . . . [should receive] . . . more liberal permission to associate with the women, in

order that . . . the greatest number of children may be the issue of such parents." Contemporary scholars of Greek history argue that while his utopian scheme was not to be realized, systematic control of procreation in Greece was an objective on several occasions.

In the first half of the twentieth century, the eugenics movements in the United States, Great Britain, Germany, and other countries included leaders from across the political spectrum. In 1865 Galton published a series of magazine articles and eventually a book (*Hereditary Genius*) that developed the key theses of early eugenics. First, intellectual and moral qualities could be inherited. Second, through appropriate policies, humans could take advantage of burgeoning (quantitative) biological knowledge to produce an improved version of humanity.

Galton's eugenic ideas were taken up by many of the followers of his biometrical work, including most notably Karl Pearson. In the late nineteenth and early twentieth century, eugenics was largely a movement of scientists and intellectuals, and in fact the history of eugenics and the development of genetics are deeply intertwined. In Great Britain, the group that initially took Galton's mantle (under Pearson's leadership) was largely socialist. Pearson, George Bernard Shaw, and Havelock Ellis used eugenics as the basis of an attack on class distinctions. They argued that these social distinctions created artificial barriers to the "natural" breeding between fitter individuals which otherwise take place. Thus, in the name of good breeding, we should tear down class distinctions.

In the United States, the leader in the eugenics movement was Charles B. Davenport. As in Great Britain, the United States eugenics movement was initially tied to developments in the biological study of heredity. But, whereas in England there was a sharp divide between biometricians (Galton, Pearson) and Mendelians, in the United States, no such deep division occurred. Consequently Davenport was able to draw on both traditions as he helped establish the Cold Spring Harbor Laboratories (one of the most prestigious institutions in the world for biological research to this day) and the Eugenics Record Office (ERO) in Cold Spring Harbor. The ERO carried out studies of "important" character traits and their hereditary pattern on a grand scale accumulating an enormous amount of data (though much of suspect scientific value). Politically Davenport was a conservative and his political approach was to have a far greater impact on eugenic policies in the United States and Great Britain than that of the socialists.

In those early days of eugenics, the possible benefit of science to the improvement of human conditions though better breeding served to unite conservatives like Davenport, radicals like Pearson, and progressives like David Starr Jordan (first president of Stanford University and a leading biologist and politician). Jordan argued that eugenics provided the best possible ground for pacifism. In times of war, countries send out their fittest individuals to die, leaving the unfit behind to reproduce. There were anecdotes that the average height in France had fallen by 6 inches after the Napoleonic wars.

In the early twentieth century there was a parallel discussion of euthenics — the scientific study of how environmental influences could be manipulated to improve humanity, and euthenics was pursued by many of the same people who were prominent in the eugenics movement. However, the focus on the role of heredity eventually led to the popularization of the eugenics movement. This took place not through the positive enthusiasm for eugenics' potential to improve conditions. Rather, a growing fear that society was becoming overrun with the unfit offspring of those with less than desirable hereditary backgrounds motivated the rise of eugenics into a full blown political force. Works such as Henry Goddard's classic on the Kallikak family gave rise to concern about the "menace of the moron." Goddard estimated the number of "feeble-minded" in the United States ranged from several hundred thousand to a million individuals. Given the perceived rapid rate of reproduction among the feeble-minded relative to the disturbingly low reproductive rates among elites, there was a great fear that the country was would be overrun with degenerates. "Fitter family" contests and displays demonstrating the dangers of the breeding of the unfit were regular displays at state fairs (2–4). Estimates of the cost to society of allowing the "least fit" to breed were estimated and became a hot political and social issue. The Jukes and the Kallikaks became household names and policies were soon developed to "solve" the problem of "race suicide" which the popular press increasingly decried.

Increasingly conservative solutions soon became promulgated. Laws allowing the forced sterilization of those deemed less fit became commonplace in most states and resulted in the forced sterilization of over 60,000 individuals. In the famous 1927 Supreme Court decision which upheld these laws (*Buck v. Bell*) Justice Oliver Wendell Holmes argued that "three generations of imbeciles is enough."

At the same time, concern grow that undesirable, inferior immigrants (with large birth rates) were rapidly contributing to the race suicide. The U.S. Immigration Restriction Act of 1924 favored immigration from Northern Europe and greatly restricted the entry of persons from other areas referred to as "biologically inferior." Intelligence testing and the rating of families became a national mania and an important aspect of immigration restriction.

Argentina, Austria, Brazil, Canada, China, Finland, France, Italy, Japan, Mexico, Norway, and Sweden each had eugenic initiatives. In Germany, the fledgling eugenics movement had existed since 1904 with the creation of the Archives of Race-Theory and Social Biology by Dr. Alfred Ploetz, and the creation in 1905 of the German Society of Racial Hygiene. The German movement could only look in envy at the growing movement in the United States and Great Britain (which ironically had now become a way of curbing the growing ranks of the lower classes). That was soon to change as the Nazis to rose to power and began to imitate and eventually "surpass" their American and British counter parts. The Kaiser Wilhelm Institute of Anthropology, Human Heredity, and Eugenics was created in 1927. German sterilization laws were enacted in 1933, requiring compulsory sterilization "for the prevention of progeny with hereditary defects" in cases of "congenital mental defects, schizophrenia,

manic-depressive psychosis, hereditary epilepsy ... and severe alcoholism (5).

The counselor of the Reich Interior Ministry called sterilization "an exceptionally important public health initiative ... we go beyond neighborly love; we extend it to future generations" (1, pp. 117–118). Under the Nazi law, physicians reported all "unfit" persons to Hereditary Health Courts, established to determine the sorts of persons who ought not to procreate. Decisions could be appealed to a "supreme" eugenics court, whose decision was final — and could be carried out by force. Within three years, German authorities had sterilized some 325,000 people, more than 10 times that in the previous 30 years in America (6). Marriage or sexual contact between Jews and other Germans was banned. According to Muller-Hill, a few hundred black children and 30,000 German gypsies were sterilized. Eventually the eugenics movement culminated in the Holocaust.

The science behind the eugenics movement was often of a suspect nature. The movement authorized some dramatic, if ill-informed inferences about the hereditary roots of a variety of behaviors and traits. Davenport, for example, argued that family pedigrees established the existence of the trait "thalosophia" or love of the sea. He traced multiple generations of sea faring in families and concluded that this was a simple Mendelian trait — and a sex-linked one (women did not seem to go to sea).

After the German experience, eugenic thought was at its nadir. Eugenics was associated with terrible images and widely discredited in the media, scholarly literature, and policy. However, whether eugenics continued to be practiced or still remains a prominent feature of modern medical genetics and counselling depends on how one understands the term.

Much has changed in the years since the endorsement and rejection of eugenics in overt political and scholarly institutions. Yet eugenics is constantly referenced as a danger of contemporary reproductive and genetic technology. With every advance in the ability of genetic testing to detect disease in adults, fetuses, or germ cells, the likelihood that such technologies might be used in discriminatory ways or as part of a thoughtless or diabolical public campaign is debated. Could eugenics, either in its optimistic or conservative historical incarnations be repeated? Scholars agree that the flowering of early twentieth century eugenics involved a fairly specific set of circumstances (e.g., the rise of the Nazi regime) and a primitive understanding of genetics and inheritance. The present context is very much different, and it bodes much more complex opportunities for misuse of technologies. In addition scholars such as Daniel Kevles have argued that an authoritarian politics, albeit perhaps different than the one seen in Germany, would be a necessary precursor to any state effort at sterilization.

Overt control and planning of reproduction has certainly seen a manifest increase in this century. The development of a birth control pill expanded reproductive control, but it carried new risks and ways of choosing whether to have children. Amniocentesis, ultrasonography, and chorionic villus sampling (CVS) made it possible to look into the womb to check on a fetus's condition. The possibility of doing so without risk to the fetus (amniocentesis and CVS each carried the risk of inducing a spontaneous miscarriage) through sampling of fetal cells circulating in maternal blood has begun to loom large. With the 1973 *Roe v. Wade* decision legalizing abortion, diagnosis of a fetal anomaly entailed the new option of therapeutic abortion. These events enlarged the region of reproductive control for families, physicians, and the community. Parents and health care providers were able to participate in social decisions about the traits that are acceptable in a child before that child is born. As the sensitivity and specificity of reproductive genetic testing has improved, the fetus and donor gamete have been more open to genetic testing. Society exerts influences on parents as they make new decisions about how and when and with what outcome they will reproduce, but such influences are the same sort it exerts on those who are deciding with whom to mate, whom to marry, and when to have children.

In most states, hereditary information is also available to institutions, such as employers and the life insurance industry. Companies whose workers are exposed to chemicals (e.g., Kodak and DuPont) routinely screen for hereditary sensitivities to a particular chemical in the work environment (7–10). Insurance companies and governments have begun to debate the use of detailed genetic information about applicants prior to granting health, life, or annuity policies, and laws banning such use have been passed in many U.S. states.

While sterilization by medical institutions or at the hand of state agencies, authorized to act on behalf of patients, is not presently illegal in many nations or U.S. states, existing laws do emphasize the importance of competent decision making and attenuate the role of the state in actual reproductive decisions that might be aimed at the prevention of disease. The repeal of sterilization laws and the censure of institutions that carried out eugenic policies did not render eugenics illegal, but the trend toward patients' rights and autonomy in reproductive health care have made attempts at comprehensive eugenic policies by governments markedly more difficult and visible (11).

Human genetic testing, gene therapy, choice of gametes, gestational carrying, surrogacy, DNA banking, and cloning all portend a range of possible benefits and hazards, only some of which can usefully be understood in terms of the legacy of eugenics. To understand what might be drawn from this century's experience with eugenics, we here make some distinctions among those ethical issues that have been grouped as "eugenics" and discuss their respective implications.

COERCIVE VERSUS VOLUNTARY

Past policies of eugenics became notorious when their application involved the use of brute force, and particularly against those least able to resist. Most notable were state-directed policies that forced sterilization or institutionalization of those whose reproductive capacity was seen as threatening to the public health. It is for this reason that contemporary genetic counselling has avowed

an ethic of noncoercion and even nondirectiveness. As patient autonomy and personal freedom have grown in social importance in the twentieth century, so too has an emphasis on allowing reproductive decision making to rest with the parties involved. Important U.S. regulatory and judicial tests of the rights of patients to make reproductive decisions were the legal battle over abortion rights and the willingness of states to regulate assisted reproductive technology. In both areas the United States has endorsed an enormous amount of personal and familial freedom against intrusion by the state into sexual and procreative activity, while still holding that it is the state's role to protect already born children against abuses by parents and others.

A variety of scholars have argued that these personal liberties are "negative" in nature, meaning that citizens are guaranteed only freedom from procreative interference (12). A positive liberty would entail the obligation of the community to provide procreative aid, akin to the courts' avowed responsibility to provide due process or many states' guarantee of a primary education. A negative liberty interest is obviously much narrower. It entails an emphasis by the state and other institutions on the rights of the sexually active, the procreating parent and the future parent, rather than an emphasis on the embodiment of future generations or their genetic endowment.

Interference by the contemporary state in reproductive decision making does take place. U.S. states and the federal government regulate marriage, prenatal testing, licensure for obstetrical services, and reproductive services for minors. In a few cases courts have required that pregnant women take action to protect a fetus that they intended to bear. However, the likelihood of comprehensive state sterilization or genetic discrimination is less today. The primary questions of reproductive freedom (apart from those discussed below under Population and Individual today concern *coercion*, either by government and other institutions or, more controversially, by social conditions more generally.

In the first case, agencies of government or medicine (or other institutions) may offer incentives or structure the delivery of information about reproductive decisions. Obviously it will be very difficult for those making reproductive decisions to do so carefully if physicians, nurses, employers, insurers, clergy, or the state skew or introduce bias in describing reproductive options. Similarly upstream decisions about which reproductive decisions will be covered by insurance companies or other paying organizations will have a marked effect on the ability of many to make choices. When research monies are allocated to one kind of disease or one technology rather than another, options are similarly constrained for those at the bedside. All these cases suggest some interference with an ideal state in which choices are maximally exercised by consenting and mature adults in a state of informed and reasonable decision making. However, it certainly remains to be seen that such a state ever existed, and in any case the development of upstream research and payment schemes cannot be made in such a way as to allow all choices to be made by all persons toward all desired ends.

At issue is the meaning of voluntariness. Two kinds of challenges to the voluntariness of reproductive decisions are made more likely as technology develops in this area. First, reproductive decisions can be made more difficult and less "free" when in a context of insufficient or coercive information. Second, and perhaps more controversially, economic and social pressures may create situations where reproductive decision making is constrained as if the situation were legislated. As philosopher John Dewey claimed, it makes little sense to say that people have a free choice if only one option is practically available. This is particularly important in the context of contemporary genetic testing. Lack of social support, economic security, or insurance could be important factors in determining whether a woman will abort a fetus at risk for a genetic disorder. Taken at the individual level, this is a threat to the negative right against interference with procreative decision making. More globally, entire groups of people, many of whom would share other ethnic or economic or class distinctions, might find themselves left out of genetic testing options or encouraged to utilize tests or procedures that would lessen the costs to the state. The collective impact of such patterns of economic allocation of services, or pressures toward allocation, might well resemble the impact of early experiments with eugenics. More dangerous, such decisions would not be traceable to a particular policy maker or agency but would be suffused in the economic climate of a market in genetic and reproductive technology. Put another way, if a market in genetic services became the primary mode of distribution, and that market failed to provide opportunities for all to make equal choices, the lack of equality might manifest itself in the appearance of a genetic underclass.

POSITIVE AND NEGATIVE

Another important distinction in the eugenics movement was between efforts primarily aimed at producing more people with desired traits and those aimed at eliminating undesirable traits. Many of the early eugenicists wanted to promote increased production of "geniuses" and people of great talent, through encouraging more scientific selection of mates, and more breeding by the chosen few. This was "positive eugenics." In its most extreme form, German SS officers were encouraged to reproduce with Aryan women, and offspring of the unions were placed with families chosen by the lead scientist of the program. Today positive practices include the selection of sperm and egg donors from highly selective reproductive recruiting pools, and a general emphasis on the importance of genetic relationship in the family (13–15).

In contrast, "negative eugenics" was concerned with eliminating the least fit individuals through reducing or eliminating their reproduction. Sterilization laws that were aimed at eliminating defectives from the population came to be the popular image of eugenics in the United States. In the infamous *Buck v. Bell* Supreme Court case (1927), Justice (and American philosopher) Oliver Wendell Homes articulated for many the felt need "to prevent our being swamped with incompetence." In upholding the eugenic sterilization laws, Holmes wrote, "It is better for

all the world if, instead of waiting to execute degenerate offspring for crime, or to let them starve for their imbecility, society can prevent those who are manifestly unfit from continuing their kind. . . . Three generations of imbeciles is enough." (2, p. 83). As has been mentioned, on the basis of these laws, over 60,000 Americans would be sterilized.

The analysis of contemporary policy in genetics and reproduction reveals little statutory regulation in most nations. Certainly there are few nations who express positive eugenic aims, though many have rules against incest and intrafamilial marriage. However, the increasing use of donated gametes that have been highly selected suggests a growing interest in improvement of particular offspring among those able to afford such improvement. Importantly, the interest in such technologies grows out of a felt need to be responsible for the embodiment of children. Parents who are unable to be linked by genetics to their children seek to replace that seamless bond with responsible decision making about what gifts they can instead give a child. In this way, what begin as a few choices about the health of potential gamete donors can expand to include a wide variety of traits that parents would like to be able to give their child in the absence of their own DNA.

In the background of new positive decision making is an assumption that the transmission of DNA from parent to child through sex and recombination is not only common but also medically normal. In offering treatments for infertility that aim at restoring as much as possible the sexually created DNA bond, medicine and society has embraced the assumption that being "related" is largely a function of sharing genetic information. This context emphasizes both a particular kind of relatedness and, more broadly, the importance of DNA and its stewardship. A $1 billion infertility expenditure in the United States alone is proof of the importance of this faith in the importance of DNA to the economy, but the real impact is the reification of a particular kind of positive reproductive control in which certain kinds of family life are both socially privileged and made medically normal.

POPULATION AND INDIVIDUAL

The most aggressive policies of eugenics were motivated by a concern for the public health, described in terms of the genetic makeup of the population as a whole. Improvement of the population is the explicit aim of these policies. In contrast, most recent practices that are seen as "eugenic" are motivated largely by a desire to improve people's decision-making options or to improve the health of individuals. Genetic testing and screening is intended primarily to be aids for prospective parents facing difficult decisions. However, it is important to keep in mind that there are collective consequences of individual decisions. The *effect* of individual decisions at the population level may be just as significant as the effect of decisions made with the explicit goal of producing population level changes. Worries about the collective consequences of individual decision making — of leaving choices to "the market" and the good intentions of parents is what motivates concerns about what sociologist Troy Duster referred to as "backdoor" eugenics, the regulation of reproductive patterns by class, access to medical services, and income (16).

MODERN DILEMMAS AND EUGENICS

In several developing nations amniocentesis is used to determine the sex of the fetus, with the goal of terminating unwanted females. This has resulted in skewed sex ratios in India and China, just one example of what can happen if genetic testing and reproductive technologies are utilized in unregulated or poorly structured ways. Yet how do we distinguish between the moral dilemma of an India that aborts unwanted female fetuses and parents who (perhaps aided by a genetic counselor) choose to abort a fetus destined to die an early painful death? Some patients with Huntington's disease feel that the several healthy decades of life that they have is what really matters. Other genetic traits many cause lesser health problems and risks. Will testing eventually stigmatize all those who are "unhealthy" or "abnormal" in any way? Will parents choose to test for other socially important traits, such as being thin, or tall? Will they test for homosexuality along with a propensity to develop heart disease (17,18)?

Critics of backdoor or market eugenics argue that the same prejudices and values that were problematic in the earlier state-directed eugenics are equally present in the new eugenics. They deny that it is possible to delimit the use of genetic selection and manipulation to any value-free concept of disease. Hence "citizens will end up being engineered in accordance with a dominant set of values after all, and the new eugenics will collapse into the eugenics of old (16,19,20).

In response to this, a number of authors have attempted a defense of market eugenics. Some claim that what is objectionable in the old eugenics is absent from the new eugenics and a significant moral distinction can be made based upon concepts such as disease and disability (21). Other "defenders" of backdoor eugenics argue that while there is no nonnormative basis for distinguishing medical from nonmedical (enhancements) genetic interventions, they deny that any sort of unity in values is likely to result from the marketplace (13,19). Kitcher argues that as long as decisions are left to well-intentioned, well-educated parents trying to do what is best for their children, eugenics is inevitable and unproblematic.

A third set of defenders begins by rejecting the "genetic exceptionalism" they see in criticisms of backdoor eugenics. We currently allow tremendous inequalities in access to environmental and educational circumstances that are far more likely to have a direct, measurable impact on the lives of children and their future expectations and opportunities. Exclusive focus on the potential for genetic inequalities is misleading and unjustified (22,23).

Another problem in genetics is the general lack of education about its meaning and use. Most in the world today do not understand genetic science, let alone the complex fact that genetic probabilities are always understood in terms of particular populations in particular environments. As a result many will not understand how

to interpret their risks. For example, testing positive for a BRCA mutation for breast cancer will have different implications for a patient with a family history of breast cancer than for one without such a family history. Yet economic incentives may lead to a push for population level screening before we have a good understanding of the relevant risks for most women.

CONCLUSION

Eugenics is a complex concept with a diverse set of meanings and uses and a rich (if nefarious) history. The ethical issues associated with eugenics depend crucially on the different meanings of the term, and on the particular details of the uses and the context of the reproductive practices under consideration.

BIBLIOGRAPHY

1. D. Kevles, *In the Name of Eugenics*, University of California Press, Berkeley, 1985.
2. D. Paul, *Controlling Human heredity: 1865 to the Present*, Humanities Press, NJ, 1995.
3. G. Allen, *Genome* **31**, 885–889.
4. M. Haller, *Eugenics: Hereditarian Attitudes in American Thought*, Rutgers University Press, New Brunswick, NJ, 1985.
5. B. Muller-Hill, *Murderous Science*, Oxford University Press, Oxford, UK, 1988.
6. R. Proctor, *Racial Hygienie: Medicine under the Nazis*, Harvard University Press, Cambridge, MA, 1988.
7. D. Nelkin, *Dangerous Diagnostics: The Social Power of Biological Information*, Basic Books, New York, 1989.
8. M.A. Rothstein, *Houston Law Rev.* **29**(1), 23–84 (1992).
9. E.V. Lapham, C. Kozma, and J.O. Weiss "Genetic discrimination: perspectives of consumers." *Science* **274**(5287), 621–624 (1996).
10. S.M. Wolf, *J. Law, Med. Ethics* **23**(4), 345–353 (1995).
11. P. Reilly, *The Surgical Solution*, Mcmillan, New York, 1993.
12. D. Heyd, *Genethics*, Cambridge University Press, New York, 1993.
13. P. Kitcher, *The Lives to Come*, Free Press, New York, 1997.
14. J. Glover, *What Kind of People Should Their Be?*, Oxford University Press, New York, 1996.
15. T. Murray, *The Worth of a Child*, University of California Press, Berkeley, CA, 1996.
16. T. Duster, *Backdoor to Eugenics*, Routledge, London, 1990.
17. G. McGee, *The Perfect Baby: A Pragmatic Approach to Genetics*, Rowman & Littlefield, New York, 1997.
18. J. Robertson, *Children of Choice*, Princeton University Press, Princeton, 1994.
19. N. Ager, *Public Affairs Q.* **12**, 137–155 (1998).
20. B. Muller-Hill, in A. Clarke, ed., *Genetic Counselling: Practice and Principles*, Routledge, London, 1994, pp. 133–141.
21. J. Harris, *Wonderwoman and Superman: The Ethics of Human Biotechnology*, Oxford University Press, Oxford, 1992.
22. A. Caplan, G. McGee, and D. Magnus, *B. Med. J.* **319**, 1284–1285 (1999).
23. J. Glover, *What Sort of People Should There Be?*, Penguin, Harmondsworth, UK, 1984.

See other entries DISABILITY AND BIOTECHNOLOGY; EUGENICS, ETHICS, STERILIZATION LAWS.

EUGENICS, ETHICS, STERILIZATION LAWS

PHILIP R. REILLY
Shriver Center for Mental Retardation, Inc.
Waltham, Massachusetts

OUTLINE

INTRODUCTION

During the last quarter of the nineteenth century and the first half of the twentieth century, the idea that the instruments of social policy could and should be used to protect the gene pool of a nation's population attracted broad interest in the United States and in several European nations. Beginning in 1907 a number of states and nations began to enact compulsory laws that authorized officials to order the sterilization of institutionalized (allegedly, retarded) persons. Prior to World War I, a number of state supreme courts held such laws to be unconstitutional. In 1927 the United States Supreme Court upheld a revised state involuntary sterilization law as a valid exercise of the police power. The years from 1927 to 1939 mark the zenith of such eugenic programs. In the United States at least 60,000 institutionalized persons were sterilized pursuant to state law. In Germany during the period 1934 to 1945 many hundreds of thousands of person were sterilized pursuant to a law with similar intent and broader reach. Although eugenically rationalized involuntary sterilization programs remained active, particularly in several southern states after World War II, the scope of most state programs rapidly diminished. The major reasons for the decline included the growing sophistication of genetics which revealed the intellectual bankruptcy of most eugenic ideas, revulsion from the revelations

concerning events in Nazi Germany, and the growth of the Civil Rights Movement. Eugenic thinking remains widespread in the world today. A prime example is a maternal and infant health law that was enacted in China in 1995, a statute with sections that recall the sterilization laws of the 1930s.

EARLY EUGENIC STERILIZATION LAWS

Origins and Rise of the Eugenics Movement

Although speculations about the perfectibility of humankind date at least as far back as the flowering of philosophy among the ancient Greek city states, the first serious social policy proposal to improve the gene pool of our species that was tied to scientific claims arose in England in the second half of the nineteenth century. Francis Galton, a Victorian polymath who was probably influenced by the revolutionary impact of *The Origin of Species* published in 1859 by his cousin, Charles Darwin, began investigating the inheritance of talent among eminent English families about 1864 (1). He coined the term eugenics (from the Greek, fusing words for good and birth) in 1883 in *Inquiries into Human Faculty and its Development* (2). Public interest in the notion that success and failure in life might be closely tied to the germ plasm that one inherited grew rapidly, especially in England and the United States, and the new word enjoyed a certain vogue. In 1904, when he made a major financial gift to the Eugenics Record Office at the University of London, Galton drafted an official definition of "natural eugenics" as "the study of the agencies under social control that may improve or impair the racial qualities of future generations either physically or mentally" (1).

In the United States eugenic policies, which flowered in the early twentieth century, germinated in the climate of progressive reform that took root in the last quarter of the nineteenth century. From 1850 to 1880, the states had built many prisons, hospitals, insane asylums, and colonies for the mentally retarded. Initial enthusiasm faded as funding problems arose and conditions declined. Richard Dugdale, a well-to-do Englishman who made New York City home, was an ardent social reformer and one of many who sought to improve such facilities. It was while inspecting prisons in upstate New York that he discovered a large family many of whose members seemed to be inhabiting one state facility or another (3). His book, *The Jukes* (4), based on an exhaustive study of the family, detailed the cost to taxpayers of their incarceration and support. He also championed two ideas that would become core beliefs among many Americans within a decade: that feeblemindedness, epilepsy, drunkenness, criminality, and insanity had strong hereditary influences, and that affected individuals tended to produce larger than average numbers of offspring.

Interest in eugenics was widespread in the United States just as biologists rediscovered the laws of inheritance that Gregor Mendel had postulated in 1865, but which had been reported in an obscure local scientific journal and been little noticed. Charles Davenport, a talented young biology professor, played an instrumental role in propagating "mendelism" in the United States. He was quick to apply the theory to problems of human heredity and about 1905 he secured a large gift from Mrs. E.H. Harriman (the wife of the railroad magnate) to develop and sustain a eugenics research facility at Cold Spring Harbor on Long Island, an entity which operated independently of the Station for Experimental Evolution that he had founded there a couple of years earlier. One of his most important decisions was to recruit a midwestern high school teacher named Harry Hamilton Laughlin to direct the Eugenics Record Office. In so doing, Davenport, a highly respected biologist who would become a member of the National Academy of Sciences, tied human genetics to eugenics and provided eugenics with a cloak of scientific legitimacy that it wore for more than three decades (5).

The indefatigable Laughlin became an ardent eugenicist who, from 1918 to 1939, played the premier role as a strategist for the eugenics movement in the United States. One of his early projects was to train cadres of young women as eugenic field-workers. Armed with knowledge on how to prepare detailed family pedigrees, many of these workers reviewed the records of thousands of institutionalized persons and interviewed their relatives, thereby gathering the raw material that became the basis for what might be considered the nation's first foray into sociobiology (3). Between 1910 and 1920 eugenicists working in association with the Eugenics Record Office and sometimes assisted by Laughlin's field-workers published a number of lengthy monographs with colorful names, such as *The Hill Folk: Report of a Rural Community of Hereditary Defectives*. These monographs reinforced the eugenic ideas propounded in *The Jukes* and the equally famous *The Kallikaks* (1912), written by Henry Herbert Goddard, a prominent psychologist who worked at the Vineland Training School in New Jersey and who imported IQ testing from France about 1905 (6). These and similar works caught the attention of American journalists. About 1910 articles on eugenics constituted the second most popular topic in the print media according to the Index to Periodical Literature (3).

Laughlin played a critically important role in the effort to secure enactment of federal laws to limit the immigration of persons that he and many others believed were of inferior racial stock. He conducted and published surveys that purported to show that immigrants from southern Europe and Russia were much more likely than immigrants from northwestern Europe to wind up in charity hospitals or need public assistance. In 1922 he served as the official expert on eugenics to the committee in the United States Congress charged with immigration matters. In that role he provided testimony that offered an apparent scientific basis to rationalize a legislative quota system that favored the immigration of some ethnic groups over others from 1924 to 1968 (7).

During the 1920s Laughlin also spent an immense amount of time drafting and lobbying for the enactment of laws to permit state officials to sterilize institutionalized retarded persons without their consent. He helped propagate the second wave of such laws that swept through America in the 1920s, and he provided an important deposition in the lower court proceedings that led to *Buck*

v. Bell (8), in which the United States Supreme Court ultimately upheld the constitutionality of a law he helped to craft. This opinion removed lingering doubts in many state legislatures and made possible the enactment of about a dozen new laws (discussed below) in the ensuing five years.

Impact of Vasectomy

Prior to 1910 in the United States proponents of eugenics tried a variety of methods to reduce child-bearing by persons thought to be unfit. In some states, notably California, the focus was on the insane; in others such as New York and New Jersey, there was special concern for "protecting" mentally retarded young women who, it was feared, were especially vulnerable to unscrupulous men. The typical proposals were to segregate the sexes in institutions and prohibit trysts. In society at large the problem was dealt with by the enactment of marriage restriction laws that forbade the insane, the retarded, the epileptic, public drunkards, those with tuberculosis, and even, in some states, the poor, from marrying. Unlike anti-miscegenation laws which were vigorously prosecuted, marriage restriction laws were not much enforced. Those few that were legally challenged in the years before World War I did not survive constitutional scrutiny.

A clinical advance, the development of the vasectomy, about 1897 had an obvious, material impact on the rise of sterilization laws. A.J. Ochsner, chief surgeon at St. Mary's Hospital in Chicago, published his surgical experience with several cases in the *Journal of the American Medical Association* in 1899 (9). The paper carried the remarkable title: "Surgical Treatment of Habitual Criminals." After describing the new surgical technique. Dr. Ochsner asserted: "if it were possible to eliminate all habitual criminals from the possibility of having children, there would soon be a very marked decrease in this class." Further, the same treatment "could reasonably be suggested for chronic inebriates, imbeciles, perverts and paupers." Although there was no comparable, safe operation for women, Ochsner opined that since most female criminals were also prostitutes and were highly likely to become infertile due to the impact of untreated gonorrhea, their class would produce relatively few children.

Ocshner's proposal received a large boost in 1902 when Dr. Harry C. Sharp, the surgeon for the Indiana Reformatory in Jeffersonville, reported on the follow-up of 42 prisoners who had agreed to undergo vasectomy, claiming that the patients "feel that they are stronger, sleep better, their memory improves, the will becomes stronger, and they do better in school" (10). He urged his fellow physicians to lobby the leaders of state institutions of all kinds "to render every male sterile who passes its portals, whether it be almshouse, insane asylum, institute for the feeble-minded, reformatory or prison." Policy makers, who had viewed earlier, sporadic proposals to castrate criminals with distaste, greeted the suggestion to use the much less mutilating vasectomy with enthusiasm (3).

In the years 1902 to 1912, Sharp was an outspoken advocate for vasectomy as a social tool. He spoke on the subject frequently at regional and national medical meetings (often exaggerating the salubrious effects of sterilization), wrote political pamphlets on eugenics, and button-holed state legislators. It is no accident that in 1907 Indiana became the first state (indeed, the first political jurisdiction in the world) to enact an involuntary sterilization law that had a demonstrably eugenic underpinning (3).

First Sterilization Laws

The nation's first sterilization bill was introduced in the Michigan legislature in 1897, but did not come to floor vote. In 1905 the Pennsylvania House of Representatives became the first legislative body to pass a bill proposing involuntary sterilization of certain institutionalized persons, but the governor vetoed it. On April 9, 1907, Indiana governor J. Frank Hanly, a month after a sizable majority of both houses had voted favorably, signed the nation's first eugenic sterilization law. The statute authorized the compulsory sterilization of "confirmed criminals, idiots, imbeciles and rapists" residing in a state institution if, after appropriate review, a panel of one physician and two surgeons concluded that it was "inadvisable" that the individual procreate and that there was "no probability of improvement." The new law was crafted to legitimize the program that Dr. Sharp was already vigorously pursuing in Jeffersonville, except it eliminated the pretense of obtaining consent (3).

In 1909 the legislatures in California, Connecticut, Oregon, and Washington passed similar laws. Despite overwhelming support in the legislature, the governor of Oregon vetoed the bill sent to him; the other three governors promptly signed their bills into law. In the ensuing four years (1910–1913), ten states (Iowa, New Jersey, Nevada, New York, North Dakota, Michigan, Kansas, Wisconsin, Vermont, and Nebraska) passed sterilization laws. In general, there was little opposition and most votes were lopsided. Only in the state where the vote was close (96–82 in the House), Vermont, did a governor cast a veto (3).

The California law, which launched the most active eugenical sterilization program in the United States until well into the 1920s, was slightly more sophisticated. Section 1 of the law covered institutionalized persons who had been diagnosed with "hereditary insanity," "incurable chronic manic," and "dementia," requiring that their discharge must be premised on "asexualization." Used in the early days as a synonym for vasectomy, this term was often confused with castration. Section 2 targeted recidivists in state prison. It identified three persons (the resident prison physician, the general superintendent of state hospitals, and the secretary of the state board of health) to review the cases of persons who had been convicted twice of rape or sexual assault or three times of having committed other crimes, and who while in prison continued to show evidence that they were moral or sexual degenerates. If two of the three reviewers concluded that there was no hope of "moral recovery," they could order that the prisoner be sterilized without his consent. The third section of the law directed the state to pay for the sterilization of institutionalized retarded children or

adults so long as their parents or guardians consented. This relatively enlightened section may explain why the constitutionality of this law was not challenged (11).

By 1913 involuntary sterilization programs were active in 14 states. There were significant differences in the their scope and pace, partly because in a number of the states opponents of eugenics attacked the constitutionality of the enabling laws. In every instance (Indiana, Iowa, Michigan, Nevada, New Jersey, New York, and Washington) in which the constitutionality was put at issue, the courts invalidated the laws, usually on the grounds that they violated the requirements of the Due Process Clause of the Fourteenth Amendment (12). Laws that targeted prisoners were held to violate the Eight Amendment prohibiting cruel and unusual punishments (13). In Oregon the sterilization law, which also was challenged on constitutional grounds, was repealed by public referendum only months after the governor signed it (3).

From 1907 to 1922 activities in California account for the vast majority of reported involuntary sterilizations conducted pursuant to state law. Of a national total of 3,233 operations, 2,558 were performed there, most on institutionalized, mentally ill persons. During that era a substantial fraction of the men and women who were discharged from a California state hospital were sterilized. The women usually were subjected to oophorectomy (removal of the ovaries), in those days a risky operation that resulted in several deaths. Due to the vigorous commitment of its medical superintendent, Dr. John Reily, during this era the South California State hospital in Patton sterilized 1,009 of its residents. In March 1918, Dr. Reily reported that he had sterilized 43 persons (3).

The constitutional challenge brought against the New York sterilization statute is of special interest because of the involvement at trial of some of the leading strategists of the national eugenic movement. The case arose because of disagreements among public officials about the scientific rationale that purported to justify sterilizing retarded persons to prevent the birth of children who too would be retarded. To resolve the dispute, the Board of Examiners of the Custodial Asylum in Rome, New York, ordered that a 22-year-old man named Frank Osborne be sterilized, knowing that his attorney would challenge the law. At trial Dr. Francis Bernstein, the superintendent of the facility, vigorously opposed the law, claiming correctly that if feeblemindedness was a recessive genetic disorder (a commonly held view in that time) than it was unlikely that Osborne would father a genetically retarded child. Charles Davenport also testified, no doubt disappointing many eugenicists by favoring a policy of segregation of the sexes over sterilization. Another prominent eugenicist, Bleeker Van Wagenen, favored sterilization but acknowledged it would be preferable to obtain the consent of the retarded person's guardian. The court, in part because of the weakness of the scientific arguments offered in support of sterilization, struck down the law (14).

During the years just before and during World War I, there was a hiatus in the enactment of sterilization laws. Almost certainly this was because of the failure of existing statutes to survive constitutional challenge. But the hiatus was probably also influenced by the sharp drop in immigration during that era. In the first two decades of the twentieth century American eugenicists were far more vexed by the massive influx of what they were convinced were racially inferior people than by the slowly growing number of retarded persons who were housed in state institutions. As the tide of immigration rapidly subsided during 1914 to 1919, the sense of urgency among eugenicists may have relaxed (15).

RESURGENCE OF STERILIZATION LAWS

Eugenic Thinking in the 1920s

During the 1920s the eugenics movement, which prior to the World War had begun to decline, grew and prospered. August institutions, such as Yale University, the Cold Spring Harbor Laboratory, and the New York Museum of Natural History numbered intellectual leaders of eugenics on their faculties. The Second International Congress of Eugenics met in New York City in 1921. By 1924 the New York based American Eugenics Society was lobbying in Albany against bills that its members thought were dysgenic (e.g., including those intended to provide financial assistance to poor women with school age children) (3). Across the nation, local eugenic societies flowered. In 1926 the Human Betterment Foundation, the pet project of an eccentric California millionaire named Ezra S. Gosney, emerged as a major voice for eugenics on the West Coast (3). In Cleveland in 1928 Charles F. Brush, a successful inventor, launched the Brush Foundation for the Betterment of the Human Race with one stated goal being the propagation of eugenic goals. In Topeka, J.H. Pile, a self-made millionaire who had studied at Yale, founded the Eugenic Babies Foundation. Especially in the midwest, interest in positive eugenics (the search for methods to have genetically superior children) captured the imagination. County fairs sponsored "fitter family" contests in which people, not unlike the prize hogs or cattle they showed, competed for blue ribbons based on their pedigree, physical examinations, and their children's report cards (16).

At the Eugenics Record Office, Laughlin, stung by the constitutional defeats suffered by the involuntary sterilization laws between 1913 and 1918, produced a massive tome on the societal benefits of eugenic sterilization (11). He carefully analyzed the laws that the courts had found flawed, and then drafted and circulated a model sterilization law that he hoped would satisfy the constitutional concerns. In the early 1920s his work was widely used by legislators who wanted to sponsor such bills. His polemics on sterilization found their way to Nazi Germany where Laughlin was held in such high regard that he was awarded an honorary degree by the University of Heidleberg in 1934 (7).

Beginning in 1923 there was a major resurgence of sterilization laws. After five years of legislative inactivity, new laws were enacted in Delaware, Michigan, Montana, and Oregon. Virginia adopted a law in 1924, and governors signed seven of the nine bills that were passed in 1925. By January 1926 eugenic sterilization laws were on the books of 17 states and small bands of pro-sterilization

lobbyists were urging many others to follow suit. In some states directors of state institutions permitted involuntary eugenical sterilizations to occur despite the absence of enabling laws (3).

In mid-1925 the Michigan Supreme Court upheld the constitutionality of the new law which it held was "justified by the findings of Biological Science," and was a "proper and reasonable exercise of the police power of the state" (17). This greatly encouraged other legislatures, but many still wondered how such laws would fare before the United States Supreme Court. In a bold move, pro-sterilization forces in Virginia decided to find out. The resulting decision, *Buck v. Bell* (8), was the single most important event in the history of the sterilization laws in the United States.

Buck v. Bell

In 1924 the Virginia legislature had by a nearly unanimous vote (30–0 in the Senate and 75–2 in the House) passed a law that authorized the superintendents of the five state institutions for the retarded to petition a special board for permission to sterilize their wards. A few months later, Dr. A.S. Priddy, the superintendent of the State Colony for Epileptics and Feeble-Minded, who had decided to test the constitutionality of the law (which he supported) selected an 18-year-old woman named Carrie Buck as his first candidate for surgery. By 1925 surgeons had become fairly adept at performing tubal ligations, although such surgery was certainly much more risky than was vasectomy. One reason why Priddy chose Carrie for the test case was that she was allegedly the daughter of a retarded woman and she had recently given birth to a child who was also reported to be retarded. The superintendent thought he had a clear case of hereditary mental retardation (18).

To bolster his case, Dr. Priddy asked a number of experts to examine Carrie or review her records. Among them was Harry Laughlin, who after reviewing her case file, concluded that Carrie was part of the "shiftless, ignorant, and worthless class of anti-social whites of the South." He opined that the possibility that her feeblemindedness arose from nonhereditary causes was "exceptionally remote" (18). On September 10, 1924, the review board approved the sterilization petition, and as planned, R.G. Shelton, her court-appointed attorney, immediately filed an appeal in the local circuit. The constitutional challenge was heard in November. The appellees assembled a formidable array of eugenics experts who testified on behalf of the scientific validity of sterilization programs. Attorney Shelton did not offer a single expert to rebut those claims. In April 1925 Judge Bennet Gordon upheld the law and ordered the operation to take place within 90 days. On appeal the Virginia Supreme Court unanimously upheld the law, finding that it was intended to benefit the persons who would be sterilized (19).

Shelton appealed to the United States Supreme Court. On May 2, 1927, by a vote of 8–1 the high court held the Virginia law to be a constitutionally valid exercise of the police power that did not run afoul of the Equal Protection Clause. Writing for the court, Judge Oliver Wendell Holmes, Jr. asserted: "It is better for all the world, if instead of waiting to execute degenerate offspring for crime, or to let them starve for their imbecility, society can prevent those who are manifestly unfit from continuing their kind. The principle that sustains compulsory vaccination is broad enough to cover cutting the Fallopian tubes" (8). Holmes, who in private correspondence revealed that he shared many commonly held eugenic tenants, wrote to a colleague that he was particularly proud of the social contribution he made by upholding the Virginia law (3).

Sterilization Data

Buck v. Bell put the constitutional questions to rest. The legislative response was swift. In 1929 nine states enacted sterilization laws, and by the end of 1931 laws had been enacted in 28 states. At the Eugenics Record Office in Cold Spring Harbor, it was Laughlin's pleasant task to monitor the spread of sterilization programs and keep tabs on the victories. His work and the surveys conducted annually by the Human Betterment Foundation (to which the state agencies dutifully reported their activities) (20) provide substantial documentation of the number of persons sterilized for eugenic reasons. The next 15 years were unquestionably the heyday of eugenical sterilization in the United States.

In California the number of reported sterilizations rose from 322 in 1925 to 2,362 in the period from January 1928 through December of 1929, about a 300 percent annual increase. In Minnesota in the decade 1916 to 1925 only 21 eugenical sterilizations were performed. In the ensuing decade nearly 1300 were performed pursuant to the same law. In 1938 the Minnesota School for the Feeble-Minded in Faribault admitted 452 new patients and sterilized 151. Nationally there were about 2,500 to 3,000 operations each year. The highest reported annual figure was 3,921 in 1932, but in 1940 the nation's institutions still reported in excess of 2,800 operations (3). These are well-documented, minimum estimates. The archives of the Human Betterment Foundation contains correspondence in which state officials and judges acknowledge that illegal sterilizations were being performed that for obvious reasons could not be officially reported. A Michigan probate judge wrote that he knew of 71 illegal sterilizations; the Maine assistant attorney general reported that "many more" operations were being performed than were being reported (3).

Despite the relative uniformity of the state laws, the programs differed substantially among states, among institutions within states, and from year to year. In most states the single most critical factor was the attitude of the superintendent of each institution. If the superintendent opposed eugenic sterilization, as was frequently the case in the northeastern states, few operations were performed with or without enabling laws. If the superintendent vigorously supported sterilization, the programs could be ensured of adequate budgets and political protection and could flourish.

During the late 1920s and 1930s, even as the pace of eugenical sterilization picked up, the rationales and the goals shifted. Perhaps the most significant change was that with each year young women accounted for

a higher percentage of those who were sterilized. In 1924 the cumulative data showed that 47 percent of those who had been sterilized were women. By the end of 1940 the data indicated that women accounted for 58 percent of the all time totals. Between 1927 and the end of 1941 about 18,300 women were sterilized, while 11,200 men underwent the procedure (3). Several factors contributed to this pattern. By 1930 American surgeons had substantial experience with tubal ligation which, being a much less risky procedure than hysterectomy and oophorectomy, they were more willing to perform. Also by this period early convictions that mental retardation was explained by a few dominant or recessive genes that obeyed Mendel's laws had dissolved, and few good physicians were willing to predict the risk that a person would parent a "defective" child. But, they were willing to predict who would be a "defective" parent.

The impact of the Depression Era economy was almost certainly very significant. Strapped by limited resources, some state officials abhorred the thought of poor, young women bearing children out of wedlock many of whom would become wards of the state. In many states the profile of the individual most likely to be admitted to the local institution was a mildly retarded teenage girl who had either born one child out of wedlock or was thought very likely to become pregnant by an unscrupulous man. In many instances these young women were admitted mainly so that could be sterilized, and rapidly thereafter returned to the community. A 1928 Wisconsin report concluded: "Many mentally deficient persons by consenting to the operation are permitted to return, under supervision, to society where they become self-supporting social units and acceptable citizens. Those inmates unwilling to consent to the operation remain segregated for social protection as well as individual welfare" (21, p. 28). Wisconsin prided itself on not performing sterilization's without consent, but seemed to overlook the fact that when the alternative is incarceration, consent cannot be voluntary.

The Wisconsin report was echoed in many other states where again and again sterilization was tied to release. Many agencies generated follow-up studies that glowingly reported that, once sterilized, midly retarded women were easily maintained at home or made wonderful live-in domestic servants. At the same time there was less interest in sterilizing retarded men because they were considered unable to have access to partners outside of the institution. Also sentiment in favor of sterilizing habitual criminals and rapists had faded as the eugenical rationale had been found scientifically wanting. At the turn of the century leading criminologists accepted a biological basis for many crimes (22), but by 1930 criminological thinking as to cause greatly favored socioeconomic forces as being of paramount importance. A genetic rationale was fading fast, but the idea that certain persons were not fit to be parents remained strong.

Another curious aspect of the history of eugenic sterilization in the United States is that the state programs seemed to flourish in different regions in different decades. During the first quarter of the century, California far outpaced the rest of the nation in number of persons sterilized. In the late 1920s and 1930s, programs in several midwestern states were the most active. In the 1950s and 1960s, three southern states accounted for more than half of the sterilizations performed on institutionalized persons in the nation (23).

Critics

Despite its legislative successes and the level of its programmatic activity in the United States during the late 1920s and 1930s, eugenical sterilization met with sustained criticism in some quarters. They included biologists and physicians, social scientists and lawyers, and, most notably, the Catholic Church. Although many scientists were sharply critical of eugenic theorists for having built social policy upon shoddy scientific data, few academic biologists were willing to take on the public role that was needed to testify against proposed bills. Two notable exceptions were the distinguished zoologist, Herbert Jennings, and Raymond Pearl, director of the Institute for Biological Research, both of whom worked at Johns Hopkins (24). The brazen manner in which Nazi Germany made its perverse eugenic ideology ever more public during the 1930s does not seem to have had much impact on academic geneticists in the United States. A significant minority of sociologists, social workers, and psychologists resisted eugenic sterilization programs from their origins, but with only modest effect. During the 1940s prominent geneticists like L.C. Dunn and Theodozius Dobzhansky argued vociferously against the eugenic proposals (25) and their influence was palpable.

Although many physicians expressed doubt about sterilization programs, only a few took up the cause. No one was more effective than Dr. Abraham Myerson, a Tufts neurologist who was especially troubled by trends in Germany. In 1936 he coauthored a report that advocated that sterilization programs should only proceed if it was based on a freely given consent, that enabling laws should apply equally to all citizens, and that advice about human sterilization should be sought only from recognized experts (26).

In many states the Catholic Church provided the only major opposition to sterilization bills. The decision by the governor of Colorado in 1927 to veto a sterilization bill can in part be traced to the strong opposition to the program from organized groups of Catholics. Eugenicists reported that Catholic groups had constituted the main (and, ultimately, effective) opposition to bills in New York and Connecticut. In 1930 Pope Pius XI issued *Casti Connubi* (On Christian Marriage), an encyclical that harshly criticized eugenics (27). During the 1940s the National Catholic Welfare Conference lobbied hard against eugenics (23). Well-placed eugenic strategists asserted that the organized Catholics were instrumental in defeating a sterilization bill in Wyoming (3).

EUROPE AND CANADA

England, the birth place of Social Darwinism and of eugenics, never enacted an involuntary sterilization law, nor came close to implementing eugenic social programs. Certainly it had its share of unabashed advocates, such

as Robert Reid Rentoul, a Liverpool physician who published a book entitled, *Proposed Sterilization of Certain Mental and Physical Degenerates* in 1903. During the decade before World War I the Eugenics Education Society was fashionable at Oxford and Cambridge, and its rolls numbered many scientists. But efforts to include sterilization programs in the Mental Deficiency Act of 1913 failed, and the issue was rarely, if ever again, debated in Parliament. This may have been in part because the academic geneticists in England argued that there were insufficient data to draw any inferences concerning the genetic influence on mental retardation, much less the likelihood that a particular person or couple would parent a retarded child (28).

Canada seems to have been influenced by events in the United States. Beginning in 1928 the Province of Alberta operated a sterilization program remarkably similar to the model advocated by Laughlin which resulted in the sterilization of several thousand persons. It continued to operate out of the spotlight until 1960 when government officials ended it. During the 1990s Alberta was the defendant in a class action lawsuit filed by some of those who had been sterilized (29). On November 2, 1999, the Government of Alberta reached an out of court settlement awarding $82 million to a group of 246 persons who had been sterilized pursuant to its law.

Involuntary sterilization in the name of eugenics reached its apotheosis in Nazi Germany from 1934 to 1945. Popular interest in eugenics swept through Germany early in the century. Interest was high enough that a sterilization bill was introduced in the Reichstag in 1907, but it was soundly rejected. World War I and its devastation curtailed interest, but in Germany, as elsewhere, there was a resurgence during the 1920s. The first university professorship in eugenics was established in Bavaria in 1923. In 1921 the German Society for Race Hygiene adopted a 41 point manifesto on eugenics that favored the right of defective persons "to be sterilized by their own wish" (30). About this time Adolph Hitler was writing *Mein Kampf*, in which he urged that "to prevent defective persons from reproducing equally defective offspring is an act dictated by the clearest light of reason. Its carrying out is the most humane act of mankind. It would prevent the unmerited suffering of millions of persons, and, above all, would, in the end, result in a steady increase in human welfare" (3). This language is strikingly similar to the opinion in *Buck v. Bell*.

A comprehensive German eugenic sterilization law was enacted on July 14, 1933. Pursuant to it, the nation set up a network of Hereditary Health Courts empowered to sterilize persons about whom, in "the experience of medical science, it may be expected with great probability that their offspring may suffer severe physical damage." At first persons with any one of nine conditions were targeted: inborn feeblemindedness, schizophrenia, manicdepressive insanity, hereditary epilepsy, Huntington's chorea, hereditary blindness, hereditary deafness, severe hereditary physical deformity, and severe habitual drunkenness." Each special court had three members: a district judge, a local public health official, and a physician deemed to be expert in making the evaluations of the individuals thought to be at risk (31).

The scale of the eugenic sterilization program in Nazi Germany dwarfed those in all other nations, including the United States. In 1934 more than 200 courts received 84,500 petitions to sterilize. These were sometimes filed by doctors or local public health officials, but often they were filed by one family member about another. In one of the more extreme examples of patriotism, substantial numbers of deaf persons volunteered to be sterilized as a show of support for the "fatherland." Of the 64,499 petitions that were heard, the courts decided for sterilization in 56,244 for a eugenic conviction rate of 87 percent. By 1935 more than 150,000 sterilizations had been approved, many based on judicial proceedings that must have taken under an hour (32).

Over the ensuing years the scope of the law was broadened. For example, in 1934 it was amended to apply to non-Germans living in Germany. During the 1940s people were often sterilized on the weakest of pretenses, such as being half-Jewish. In 1951 the Central Association of Sterilized People in West Germany estimated that the Nazi programs had sterilized 3,500,000 persons, although it is not possible to document the claim (3).

Nazi Germany also operated a positive eugenics program known as Lebensborn that fostered reproduction by ideal Aryan couples selected for that purpose by public health officials (33). Such couples and their offspring received a variety of extra public benefits (tax breaks, funds for child support). At about the same time in the United States a private organization, The Pioneer Fund, with the approval of federal officials, operated a similar program that offered financial aid to any officer in the U.S. Air Corps with three children who had another one during the year 1940. About a dozen children eventually received such scholarships (34).

Eugenical sterilization laws were enacted in other European nations as well. Norway and Sweden adopted programs in 1934 and Finland did so in 1935. The program in Sweden, which remained active into the 1970s, was not anchored to unsupported genetic assumptions concerning the transmission of genes that caused mental retardation. Instead, it targeted individuals who were not thought fit for parenthood, arguably a much larger cohort (3). During the early 1990s Sweden repudiated its sterilization program. In November 1999 Swedish officials disclosed that more than 500 persons who alleged that they had been involuntarily sterilized had filed for compensation, and that it had awarded and would continue to award 175,000 crones ($21,250) to each person who it determined had been sterilized without consent.

STERILIZATION IN THE UNITED STATES AFTER WORLD WAR II

Although many state sterilization programs went into sharp decline or ceased during World War II, it is inaccurate to assert, as has frequently occurred, that revulsion over the crimes committed by the Nazis was the major impetus for this change. The large reduction in sterilizations from 1942 to 1945 was mainly due to the unavailability of trained nurses and surgeons to serve the state institutions. The urgency of the war effort put

the programs on hold. The war almost certainly speeded a decline that was inevitable. Also, during this hiatus and into the late 1940s and 1950s, genetics continued to mature. Early notions that a phenotype as complex as mental retardation could be explained by a few purely mendelian alleles were now regarded as at best quaint and at worst dangerous. Justice Holmes dictum that "three generations of imbeciles are enough" rang scientifically hollow and was embarrassing.

Nevertheless, in some quarters sterilization programs flourished. Of particular interest is the work of the Human Betterment Leagues, the brainchild of Dr. Clarence Gamble, a physician and an heir to the Proctor & Gamble soap fortune (35). In 1944 Gamble joined Birthright, a successor to the Sterilization League of New Jersey, an unabashedly proeugenics group run by a former social worker, Marion Norton. From 1941 to 1942 this group had tried, but repeatedly failed, to secure a sterilization law for New Jersey. Gamble arranged for the Cincinnati-based Gamble Trust Fund to grant Birthright a gift of $10,000 for educational work. These funds were used, in effect, to decentralize Birthright by founding a number or local programs in other states (3). These Human Betterment Leagues sprung up in several states in the midwest and south. The effort had its greatest impact in North Carolina from 1945 to 1963 (36).

In 1945 Gamble funded a study of IQ among rural residents of North Carolina and reported the results to public health officials. The disturbing findings led the state officials to permit a pilot voluntary sterilization program in Orange county in which trained sociologists identified women whom they felt were incompetent to be mothers and offered them the option of sterilization. After two years officials considered the program a great success, and by 1948 social workers throughout rural North Carolina were searching for appropriate candidates. From 1948 to 1955 about 186 women (mostly from rural areas), half of whom never had resided in a state institution, were sterilized (36). In Iowa a similar effort undertaken from 1944 to 1947 resulted in about 50 allegedly voluntary sterilizations each year.

During the years from 1952 to 1958 from 50 to 75 percent of all eugenical sterilizations conducted pursuant to state law were performed in just three states — Georgia, North Carolina, and Virginia. In 1958 those states reported sterilizing 574 persons, 76 percent of the nation's total.

There is, with one exception, no evidence that programs in the southern states were racially motivated. In fact state sterilization programs in the deep south (Alabama, Georgia, and South Carolina) originally targeted whites (23). This was because during the 1930s when eugenic sterilization programs in the deep south commenced, most state institutions for the retarded and the mentally ill provided comparatively few or no beds for blacks. It also reflected the overriding concern among southern eugenicists to purify the Caucasian race. In the end, as segregated state facilities were made available to blacks, they and whites were sterilized in numbers that roughly mirrored the composition of the general population. Of course they were sterilized in racially segregated facilities.

South Carolina is the only state in which there is evidence that sterilization programs were aimed directly at black persons residing in state institutions. Although the state facility for the mentally retarded did not admit blacks, the facility that housed the incurably mentally ill did. During the years from 1949 to 1960, 104 inmates of the later institution were sterilized, of whom 102 were black (3).

Georgia offers a good example of the idiosyncratic manner in which state sterilization laws were often implemented. Although covering all state institutions for the retarded and/or insane, the vast majority of sterilizations were performed on the residents of only one, the huge (12,000 residents) Milledgeville State Hospital for the mentally ill. Among residents at Milledgeville, schizophrenia was the most prevalent single diagnosis. During the heyday of sterilization, most Georgia physicians and alienists (psychiatrists) held the thesis that schizophrenia was largely hereditary. For well over a decade physicians at Milledgeville sterilized more than 200 persons each year. Unlike officials at other state institutions, authorities at Milledgeville kept the sterilization effort quiet. This was one reason it continued into the early 1960s (23).

During the 1960s the number of reported eugenical sterilizations, already much below the numbers reached in 1930s, declined to very low levels. There is some evidence that inside the gates of a few state institutions sterilizations continued to occur, but these were usually done at the behest of relatives who feared that a retarded young woman would be seduced or raped and become pregnant. From June 1970 through April 1974, at least 23 sterilizations were performed on residents of North Carolina institutions pursuant to the eugenics statute (37).

Although one cannot point to a moment in which state-sanctioned eugenical sterilization in the United States ended, a satisfactory date is 1983 when a class-action lawsuit brought by women in Virginia who had been sterilized without their consent while in state facilities was settled. The case, *Poe v. Lynchburg*, was filed by the American Civil Liberties Union (ACLU) in December 1980 on behalf of five plaintiffs and their class. The allegations of one plaintiff are illustrative. She claimed that she was not retarded (38). She had been admitted to the Lynchburg Training School in 1949 after giving birth at age 14 to a child whom she said was conceived when she was raped by her stepfather. At the facility she was told she needed an appendectomy; shortly after undergoing the operation she was discharged. A few years later she married but did not discover that she was infertile because she had undergone a tubal ligation until more than two decades later. The records of many state institutions include summary surgical statistics that support this, for they indicate that a incomprehensibly large number of appendectomies were performed on young women.

The lawsuit, which among other things, challenged the constitutionality of the Virginia sterilization law, had little chance of success, for it was taking on the very statute that had been upheld in *Buck v. Bell*. However, under the 1983 settlement the state of Virginia agreed to attempt to locate all living persons who had been sterilized pursuant

to the law, and to provide them with modest compensation. Relatively few came forward (3).

Sterilization in the Courts

Given the legitimacy conferred upon eugenical sterilization programs by the United States Supreme Court in *Buck v. Bell*, courts played a relatively unimportant role in the demise of state sterilization laws in the United States after World War II. In 1942 the high court had an opportunity in *Skinner v. Oklahoma* to revisit eugenical sterilization when it considered a constitutional challenge to the Oklahoma Habitual Criminal Sterilization Act, a 1935 statute that authorized the state to sterilize any person upon his or her third felony conviction of crimes involving "moral turpitude" (39). The statute exempted certain crimes that today we might call "white collar" (e.g., income tax evasion, embezzlement) from the reach of the law, a feature that the Supreme Court found to violate the Equal Protection Clause. No doubt influenced by practices in Nazi Germany, Justice William O. Douglas, writing for the Court, warned that such laws "in evil or reckless hands" . . . can cause races or types that are inimical to the dominant group to wither and disappear" (39). But the Court decided the matter narrowly, declining to review the broader issues raised in *Buck v. Bell*. Thus the constitutional status of laws targeting institutionalized retarded persons was unaffected by the 1942 holding.

The Supreme Court has only considered the power of the state to sterilize persons without their consent on one other occasion — growing out of a judicial proceeding in a county court in Indiana in 1971. At issue in *Stump v. Sparkman* was whether an Indiana judge could be sued for granting a petition brought by a woman to have her mildly retarded 15-year-old daughter sterilized because the mother feared that the daughter would become pregnant (40). The central issue was the scope of judicial immunity for actions taken on the bench. Finding that no Indiana statute forbade the judge to consider such petitions and noting in passing that Indiana had enacted an eugenical sterilization law, the high court ruled that the judge could not be sued by the woman who had been sterilized once she discovered that fact (40). As was so often the case, she had been told that she would undergo an appendectomy, and only learned the true nature of the surgery after she married and realized she was infertile.

Perhaps the most significant decision against sterilization practices was handed down by the federal district court for Washington, DC, in 1976, resolving a lengthy battle over procedures to be followed in permitting federally funded medical clinics to sterilize individuals. Growing out of a widely reported incident in which several poor, black teenage girls were sterilized without having first given informed consent, the case, *Relf v. Weinberger*, forced the U.S. Department of Health and Human Services to formulate strict guidelines concerning procedures (including informed consent for adults and a prohibition of the sterilization of minors) that must be adhered to if a sterilization was requested (41).

During the 1970s and 1980s, a period in which there were virtually no statutorally based sterilizations (although only a few of the enabling laws were repealed), a new question was presented to the supreme courts of more than a dozen states. Under what circumstances, if any, may a noninstitutionalized retarded person be sterilized? Typically such cases arose as sterilization petitions brought by the mothers of mildly to moderately retarded teenage daughters whom they feared would become sexually active or raped. Faced with such questions state courts had two options: to hold that, absent express legislative authorization, they lacked power to decide the matter or to rule on the matter. Most of these cases wound up in the state supreme courts where in a series of cases in the 1970s the courts embraced a "best interests" test. That is, they would approve the sterilization petitions if they were convinced that the sterilizations were intended primarily to benefit the young woman (the cases never involved men) who were the subject of the petitions. The courts almost invariably found a judicial path to approve sterilization petitions brought by caring parents. During this era, a period in which society was struggling to integrate persons with disabilities into everyday life, the courts essentially held that, just as did persons of normal intelligence, persons with developmental disabilities had a right to be sterilized if it was in their best interests. In the 1970s nine state supreme courts articulated this right, followed by several more in the 1980s. A New Jersey decision, *In the Matter of Grady*, is among the most thorough discussions of the matter (42).

EUGENIC LAWS TODAY

Globally speaking, government-sponsored sterilization programs that were premised on the value of negative eugenic policies (e.g., sterilization and immigration restriction laws) declined rapidly after World War II. In the United States about five states have repealed their laws, but in most cases the laws remain on the books, although no programs are active. The law in Alberta, Canada was repealed in 1972 (43) and Sweden repealed its law in the early 1990s. Japan enacted a "Eugenic Protection Law" in 1948 which permitted persons affected with or at risk for a litany of disorders, some of which were correctly identified as genetic, to obtain sterilization. The law was not compulsory. The statute was recently amended, and the term eugenic was dropped (44).

No modern European state has a law authorizing authorities to sterilize persons without their consent. Indeed, many have enacted laws that expressly or implicitly forbid state-supported eugenic sterilization. These include France, Germany, Norway, the United Kingdom, Spain, and Switzerland. The Council of Europe's Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Medicine (1997) opposes eugenic programs. The United Nations has recently endorsed a Universal Declaration on the Human Genome and Human Rights (1997) which a eugenic sterilization law would clearly contravene (45).

In the modern era practices in India and China have raised concerns that eugenic thinking is alive and well (46). Although now officially forbidden by the governments, there are states in India and provinces in China where it

is relatively common practice to use medical technology and selective abortion to avoid the births of girls. This, together with the once not uncommon practice in China of denying lifesaving treatments to infant girls who are ill, has led to claims that as many as 100 million girls are missing from the Asian continent.

In 1983 Lee Kuan Yew, the autocratic Prime Minister of Singapore, launched a social program based on unscientific assumptions concerning positive eugenics. Essentially he authorized programs to encourage well-to-do, highly educated persons to marry and have large families. The unspoken, but clear, message was that such persons would be more likely to produce bright talented children than would persons in the lower classes (47).

China enacted a Maternal and Infant Health Care Law in 1994 that contains many laudable elements (48). However, it also includes language that many Western observers have interpreted as revitalizing long discredited eugenic notions (49). The law requires medical counseling before marriage for people whose families have a relative with one of a listed group of conditions (including mental retardation, epilepsy, and mental illness) that the law seems to presume are hereditary. Related language has been interpreted to require sterilization or the monitored use of long-term contraception as a precondition of marriage if a person is determined by a doctor to be at risk for parenting such children. However, the law includes no penalty for noncompliance, and some have interpreted it as expressing more an ethical obligation than a legal requirement. Although the law is not generally used to restrict child-bearing, it appears to be a public health policy in China to discourage reproduction by mentally retarded persons (49).

CONCLUSION

Ever greater understanding of the human genome has led to ever greater certainty that complex phenotypes such as intelligence emerge from countless interactions between genes and the environment in which a person develops and lives. The naive application of mendelism to complex human conditions that was common in the first three decades of the twentieth century is no longer scientifically accepted. For this reason social programs built on unsupported quasi-genetic tenets have virtually no adherents among biologists and physicians.

The late twentieth century has witnessed in the West at least a remarkable surge of concern for the well-being and rights of persons with developmental disabilities. Persons who once were housed inside the walls of state institutions for the retarded are today living and working in the community. The majority of citizens applaud this change. The Americans with Disabilities Act, arguably the most important civil rights legislation enacted in the United States since the 1960s, reifies a national commitment to treat disabled persons as equals.

Today the notion of involuntary eugenical sterilization of a person to prevent the infinitesimally small contribution to the gene pool that would be caused by his or her reproducing is scientifically ludicrous. But one cannot discount the possibility that misinformed and prejudiced persons and political entities will choose to rationalize their acts with eugenic arguments. The twentieth century drew the millennial curtain on a world in which "ethnic cleansing," a political goal frighteningly similar to the Nazi ideology of the 1930s, was being attempted in several nations on several continents.

While it is highly unlikely that state-supported eugenic sterilization programs will reassert themselves in Western nations, it is likely that eugenic thinking will manifest itself in other ways. In 1971 Nobel laureate Williams Shockley suggested to the American Psychological Association that persons of low intelligence (as measured by IQ scores) should be offered financial incentives to be sterilized with the incentive growing as the IQ score dropped (16). Comparable ideas have been floated with some regularity in every decade.

Much more important to consider is the impact of prenatal diagnosis. As this technology includes an ever larger array of tests that ever more women will use, an ever larger number of fetuses that would in earlier times have been born with disabilities will be aborted. This trend is already well underway in respect to the fate of fetuses ascertained through screening programs designed to warn women about the risk of bearing children with spina bifida (50). Similarly widespread use of prenatal screening coupled with selective abortion is causing a significant decline in the number of children born with Down syndrome. Such outcomes are the result of free choices made by thousands of women when confronted with knowledge delivered to them by the application of new tools. The results are not the product of a state law, yet they may represent a new form of eugenics. The tools are used and the choices are made in a climate that seems to accept those born with disabilities while promoting efforts to avoid such births.

The challenge before us is to mobilize advances in genetics to maximize benefits for individuals while blocking the efforts of malevolent or ignorant persons to misuse those tools in the name of a false science. In this regard there is no weapon as powerful as education. Only when we all understand that humans result from an ultimately unfathomable complex of gene-environmental interactions and that it makes (with rare exceptions) no sense to attempt to predict human phenotypes will we truly be confident that another sad episode in the history of eugenics does not lurk in our future.

BIBLIOGRAPHY

1. D.W. Forrest, *Francis Galton: The Life and Work of a Victorian Genius*, Taplinger, New York, 1974.
2. F. Galton, *Inquiries into the Human Faculty and its Development*, Macmillan, London, 1883.
3. P.R. Reilly, *The Surgical Solution: A History of Involuntary Sterilization in the United States*, Johns Hopkins University Press, Baltimore, MD, 1991.
4. R.L. Dugdale, *The Jukes: A Study in Crime, Pauperism, and Heredity*, 4th ed., Putnam, New York, 1910.
5. C.E. Rosenberg, *Bull. Hist. Med.* **35**, 266–276 (1961).
6. H.H. Goddard, *The Kallikak Family*, Macmillan, New York, 1912.

7. F.J. Hassencahl, *Harry H. Laughlin, Expert Eugenics Agent for the House Committee on Immigration and Naturalization, 1921 to 1931*, Ph.D. Dissertation, Case Western Reserve University, 1970, University Microfilms, Ann Arbor, No. 50.
8. *Buck v. Bell*, 274 U.S. 200 (1927).
9. A.J. Ochsner, *J. Am. Med. Assoc.* **53**, 867–868 (1899).
10. H.C. Sharp, *N.Y. Med. J.* 411–414 (1902).
11. H.H. Laughlin, *Eugenical Sterilization in the United States*, Chicago Psychopathic Laboratory, Chicago, 1922.
12. *Smith v. Board of Medical Examiners of Feebleminded*, 88 Atl. 963 (1918).
13. *Mickle v. Henrichs*, 262 F. 2D 687 (1918).
14. *In re Thompson*, 103 Misc Rep. 23, 169 N.Y.S. 538 (1918).
15. R.L. Garis, *Immigration Restriction*, Macmillan, New York, 1927.
16. D.J. Kevles, *In the Name of Eugenics: Genetics and the Uses of Human Heredity*, Knopf, New York, 1985.
17. *Smith v. Probate*, 231 Mich. 4409 (1925).
18. R.J. Cynkar, *Columbia Law Rev.* **881**, 1418–1461 (1981).
19. *Buck v. Bell*, 143 Va. 310 (1925).
20. E. Gosney and P. Popenoe, *Sterilization for Human Betterment*, Macmillan, New York, 1929.
21. State Board of Control of Wisconsin, *Nineteenth Biennial Report: Period Ending June 30, 1928*, Democratic Printing, Madison, WI, 1928.
22. H.M. Boies, *Prisoners and Paupers*, Putnam, New York, 1893.
23. E.J. Larson, *Sex, Race, and Science: Eugenics in the Deep South*, Johns Hopkins University Press, Baltimore, MD, 1995.
24. H.S. Jennings, *The Biological Basis of Human Nature*, Norton, New York, 1930.
25. L.C. Dunn and T. Dobzhansky, *Heredity, Race, and Society*, New American Library, New York, 1952.
26. A. Myerson, *Eugenical Sterilization: A Reorientation of the Problem*, Macmillan, New York, 1936.
27. *Five Great Encyclicals*, Paulist Press, New York, 1939.
28. G.R. Searle, *Eugenics and Politics in Great Britain, 1900–1914*, Noordhoof International, Leyden, The Netherlands, 1976.
29. *Muir v. Alberta*, Atla. L.R. 3d 305 Alt. Q.B. (1996).
30. F. Lenz, Eugenics in Germany, *J. Heredity* **15**, 223–231 (1924).
31. P. Popenoe, *J. Heredity* **26**, 257–260 (1935).
32. R. Cook, *J. Heredity* **26**, 485–489 (1935).
33. M. Hillel and C. Henry, *Of Pure Blood*, McGraw-Hill, New York, 1976.
34. D.A. Blackmon, A Breed Apart. *Wall Street J.*, August 17, p. 1 (1999).
35. C. Gamble, *J. Am. Med. Assoc.* **141**, 773 (1949).
36. M. Woodside, *Sterilization in North Carolina*, University of North Carolina Press, Chapel Hill, NC, 1950.
37. *North Carolina ARC v. North Carolina*, 430 F. Supp. 451 (1976).
38. *Poe v. Lynchburg*, 518 F. Supp. 789 (1978).
39. *Skinner v. Oklahoma*, 316 U.S. 535 (1942).
40. R. Macklin and W. Gaylin, eds., *Mental Retardation and Sterilization: A Problem of Competency and Paternalism*, Plenum Press, New York, 1981.
41. *Relf v. Weinberger*, 565 F. 2d 722 (1977).
42. *In the Matter of Grady*, 426 N.E. 2d 467 (1981).
43. T. Caulfield and J. Robertson, *Alberta Law Rev.* **35**, 59–80 (1996).
44. T. Tsuchiya, *Newslet. Network Ethics Intellectual Disabil.* **3**(1), 1–4 (1997).
45. B. Knoppers, *Nature Genet.* **22**, 23–26 (1999).
46. D.C. Wertz and J. Fletcher, *Soc. Sci. Med.* **46**(2), 255–273 (1998).
47. Chan, *Int. J. Health Serv.* **15**(4), 707–712 (1985).
48. Law of the People's Republic of China on maternal and infant health care (official translation) 1994, Legislative Affairs Commission of the Standing Committee of the National People's Congress of the People's Republic of China, Beijing.
49. Board of Directors of the American Society of Human Genetics, *Am. J. Hum. Gen.* **64**, 335–338 (1999).
50. H.S. Cuckle and N.J. Wald, *Prenat. Diag.* **7**, 91–99 (1987).

See other entries DISABILITY AND BIOTECHNOLOGY; EUGENICS, ETHICS.

F

FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE

MICHAEL J. MALINOWSKI
Widener University School of Law
Wilmington, Delaware

OUTLINE

INTRODUCTION

The 1990s transitioned life science from a century of profound discovery and commercial development into a new millennium that holds unprecedented potential to improve health care through technology. The Food and Drug Administration (FDA), which approved more new drugs in 1996 to 1999 than in any three-year period since 1962 (1) continues to implement a comprehensive modernization mandate from Congress to accelerate the accessibility of health care innovation (2). The pipeline of applied life science innovation, from laboratory to market, never has been filled with so much promise to alleviate suffering from debilitating diseases and generally to improve human health. "The industry is now delivering its second generation of products, a generation that includes humanized monoconal antibodies, protease inhibitors, traditional small molecules, proteins that serve as drugs, and combinations of delivery devices/drugs that uniquely link diagnostics and therapeutics" (3).

The first generation of biotech products has reached the clinic and market. Close to 100 biotech drugs now are commercially available, and FDA approved 24 biotech products in 1997 alone — a 12-fold increase since 1994 (4). "In fact, of the 350 new biotech drugs moving through clinical trials toward FDA approval, some 30% are in the late stages of testing. More than a third of those target various types of cancer; the rest target AIDS-related diseases, autoimmune disorders, diabetes, infectious diseases, and other ailments, according to a 1998 survey published by the Pharmaceutical Research and Manufacturers of America" (5).

These accomplishments tower high above the expectations of most health care providers at the commencement of the Human Genome Project (HGP) in 1990. Nevertheless, biotechnology's achievements are merely the beginning of markedly more rapid scientific progress that will change health care fundamentally and comprehensively during the early decades of the twenty-first Century (6). Notably scientists are coupling biotechnology and informatics to identify the intricacies of protein interactions and their impact on cell function and disease pathways, a field known as proteomics. People's individual genotypes are now being taken into account in pharmaceutical clinical trial design. Increasingly, the delivery of health care, including the prescription of pharmaceuticals, will become tailored to personal genotypes through pharmacogenetic profiling.

Change is the theme of this article. The first part summarizes the FDA's official approach to the regulation of biotechnology and presents a primer on how the FDA generally regulates the major groupings of products developed through biotechnology. Parts III and IV offer discussion of present changes to, within, and around the agency, and changes the agency will face as biotechnology and health care fully integrate over the next several years.

BACKGROUND: FDA REGULATION OF PRODUCTS DEVELOPED WITH BIOTECHNOLOGY

As a matter of federal policy, the United States evaluates and regulates products, including products derived through biotechnology, based on what they are rather than according to the processes used to make them (7,8). This official approach to the regulation of biotechnology, known as the Coordinated Framework for Regulation of Biotechnology and adopted in the mid-1980s (9), has distinguished the United States:

> Where other countries have tried to write entire new bodies of jurisprudence in response to recent medical advances, American lawmakers have said that questions raised by biotechnology can all be answered within the body of existing law. As a result, while other nations' biotech industries have become mired down in legal wrangles, the industry in America is booming, with 1997 sales of $13 billion . . . (10).

Although agency compliance with the Coordinated Framework policy has not been uniform, the FDA generally has adhered to the policy. In fact the FDA was instrumental in the policy's formation and adoption: "During the biotech regulation formation process, the FDA determined that

its regulatory infrastructure could handle biotechnology while EPA and USDA concluded that rDNA techniques introduce, per se, an incremental risk in new products" (7,9,11).

As discussed later, during the 1990s the FDA went through a period of public questioning fueled by collaborations among the research-driven life science industries, a Republican Congress, and patient groups. This questioning inspired new legislation, coupled with self-assessment and reform from within the agency. The Prescription Drug User Fee Act of 1992 (PDUFA) (12), which was proposed and heavily supported by industry, significantly expanded the staff of FDA through drug sponsor fees, and thereby accelerated review and approval times. Although PDUFA did not directly address clinical development requirements, the Food and Drug Agency Modernization Act of 1997 (FDAMA), which reauthorized the collection of user fees, includes several provisions intended to reduce drug development times (13). The net effect of these reforms has been documented by the Tufts Center for Drug Development: FDA approved 108 new chemical entities in 1996 to 1998, the largest total number of approvals in a three-year period since 1962 (1). "The 1990s value represents a 47% increase over that of the 1980s" (13).

The following is a primer on FDA review of major biotech product groupings. The groupings addressed include therapeutics, diagnostics and medical devices, vaccines, tissue products, and food.

Therapeutics

FDA and sponsors of therapeutic products are drawn together when new chemical entities advance from animal studies into human clinical trials (14). The U.S. Food, Drug, and Cosmetic Act (FDCA), which prohibits the introduction of drugs into commerce in the absence of data sufficient to establish safety and efficacy, includes an express exception for clinical experimentation that complies with regulatory precautions to protect human subjects (15). To exercise this exception, sponsors must submit an investigational new drug application (IND), and FDA must approve that application (14). FDA authority over clinical trials also is bolstered by their expense and the fact that sponsors are proceeding with the objective that FDA ultimately will accept the resulting data in conjunction with an application and find it persuasive of both safety and efficacy. Human clinical trial highlights are set forth in Table 1 (14).

Table 1. Human Clinical Trial Highlights

Characteristics	Descriptions
Phase I	Closely monitored studies
Primary objectives	Determine *toxicity* and whether the drug generally is safe for human use Determine the *preferred route of administration* Determine the *safe dosage range*
Secondary objective	Make a preliminary determination of effectiveness
Subjects	Small number of subjects (less than 100) and, in the U.S., usually healthy volunteers
Time frame (U.S.)	From six months to one year
Prerequisites (U.S.)	Approval of an investigational new drug application (IND) Protocol approval by an institutional review board (IRB) and For gene therapies, special protocol approval by FDA and perhaps also by the National Institutes of Health (NIH)
Prerequisites (EU)	Ethics committee approval and/or approval of the equivalent of an IND
Phase II	Placebo-controlled and double-blind
Primary objectives	Develop *dosage and toxicity data* Assess the *risks of administration* Obtain preliminary evidence of *effectiveness*
Subjects	Several hundred subjects (usually 100 to 300) who are patient volunteers with the target condition
Time frame (U.S.)	Two years
Phase III	Randomized, double-blind studies
Primary objectives	Verify *effectiveness* Determine the incidence of *adverse reactions over time* Overall, gather enough data to make a *meaningful risk-based assessment*
Secondary objectives	*Refine dosage* and administration ranges Determine appropriate *labeling* Perhaps address *pharmacoeconomic considerations* (cost-benefit analysis of the drug for targeted consumers)
Subjects	Approximately 1000 patient volunteers
Time frame (U.S.)	Three years (subject to acceleration, especially for fatal conditions without alternative treatments)

Historically FDA has drawn a regulatory distinction between new drugs and new biologics. Although all new drugs, referred to as new chemical entities (NCEs), are regulated under FDCA (15), biologics have been subject to sometimes onerous additional requirements under the Public Health Services Act (PHSA) (14,16). The primary objective of PHSA is to control manufacturing processes, meaning that, relative to sponsors of traditional drug products, sponsors of biologics have been subjected to additional licensing requirements for manufacturing (14). FDA's definition of biologics is broad enough to encompass virtually all biotech therapeutics: "A biologic drug is a virus, therapeutic serum, toxin, antitoxin, vaccine, blood component or derivative, allergenic product or analogous product, or arsphenamine or its derivatives (or any other trivalent organic arsenic compound) used for a therapeutic purpose" (17).

FDA bureaucracy reflects this historic division in the regulation of drugs and biologics for, administratively, there are two regulatory pathways to approval — the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER). Sponsors petition for the classification — CBER or CDER — they desire, and FDA has 60 days to respond in writing (2). If FDA fails to respond on time, the applicants' recommendation becomes binding. FDA may modify the classification only with the consent of the application or based upon public health reasons (2).

During the 1990s the historic distinction between drugs and biologics has been blurring in favor of harmonization between CBER and CDER. The introduction of the biologics license application (BLA) constitutes a major bridge between the Centers. Prior to the BLA, drug sponsors filed new drug applications (NDAs) with CDER while biologics sponsors had to file product license applications (PLAs) coupled with establishment license applications (ELAs) and other license applications, and those had to be filed with CBER. BLA is a single application covering all biologics (14,17). Today drug sponsors still file NDAs with CDER, but biologics sponsors may petition to file BLAs with either CBER or CDER (2,17). The advent of the BLA also marks a much more fundamental harmonization between CBER and CDER through the FDA reform movement and modernization:

> The combined impact of the Food and Drug Administration Modernization Act (FDAMA) of 1997, reinventing government (REGO) initiatives, improved bioanalytical methods, FDA's increased familiarity with recombinant DNA product safety, and good manufacturing practices (GMPs) have allowed for more harmonized review between the Center for Drug Evaluation and Research (CDER) and the Center for Biologic Evaluation and Research (CBER) for specified biologics (18).

The net effect of harmonization between CBER and CDER is the lessening of incentives on the part of NCE sponsors to attempt to steer products to one division or another and force them into product classifications to achieve this purpose. Ideally the end result of the new consistency will be less gaming of the system by sponsors, and more honesty, transparency, consistency, and predictability throughout the process, which in turn should improve the reliability of the review process and lower transaction costs for industry. Implementation of FDAMA is still underway, so time will tell.

Diagnostics and Medical Devices

"Medical device" has served as a regulatory catch-all for health care products for human use and consumption other than pharmaceuticals, including components, parts and accessories of devices; diagnostic aids, including reagents, probes, and antibiotic sensitivity discs; and test kits for use in laboratories. In fact, under FDCA, the term "device" encompasses all health-care products that do not achieve their primary intended purpose through chemical action in or on the body, or by being metabolized (15).

FDA Classification and Review of Devices. The FDA has established a comprehensive system to regulate the safety and effectiveness of medical devices under the FDCA as amended by the Safe Medical Devices Act of 1990 and Medical Device Amendments of 1992 (15). The most fundamental difference between the regulation of therapeutics and devices is that devices are subject to classification at the outset, and the classification is correlated with varying degrees of scrutiny and regulation (14). Virtually all new and changed devices are classified at the outset pursuant to FDA's premarket notification (PMN) or 510(k) notification process (15). Devices are classified I to III based on the concerns about safety and effectiveness, with Class I devices subject to only general controls, Class II devices subject to special controls, and Class III devices generally subject to full premarket review and approval to ensure safety and effectiveness (15).

Before developers make medical devices available to the public even for research, they must apply for an investigational device exemption (IDE) — the device analogue to an IND for new drugs (15). However, developers may be able to circumvent the IDE requirement by establishing that there is an independent means to confirm the validity of their tests — such as by establishing that the product is the substantial equivalent of a previously marketed product (meaning the 510(k) clearance process discussed below) or by just directly obtaining a premarket approval (PMA).

Devices, depending on their classification and other considerations, may reach the commercial market through any of the following: (*1*) a PMA, (*2*) a 510(k) exemption with a PMN, or (*3*) a 510(k) exemption and a PMN exemption under Section 510(m) of FDAMA (2,14). The 510(k) exemption is for diagnostics that are the substantial equivalent of others already approved or otherwise exempted under section 510(k) of the FDCA (15,19). Under FDAMA (2), the FDA has been downgrading device classifications, and many devices are entering the market even without PMN reporting. FDA now publishes lists of Class I and Class II devices that qualify for this double (510(k) and PMN) exemption.

Any change in a device's design triggers additional review (20), and FDA also regulates device construction and manufacturing pursuant to good manufacturing process (GMP) requirements. These requirements, which are tailored to all stages of the manufacturing process,

can be detailed. FDA monitors compliance through factory inspections (at least once every two years for Class IIII products) (15) and postmarketing reports, such as Medical Device Reports on adverse incidents (21).

FDA Review of Biotech-Based Diagnostics and Predictive Tests. FDA has adopted an ad hoc approach to biotechnology-based products, which has made classification of biotechnology-based diagnostics and other tests, including predictive genetic tests, extremely unpredictable (14). Also, due to a number of ongoing federal regulatory initiatives and multitude of state legislation, the manufacturers of gene-based diagnostics and other tests must expect additional requirements, including mandatory counseling requirements and specified informed consent requirements.

As suggested at the outset of this article, in addition to breakthrough therapeutics, understanding the function of genes offers the promise of tailoring health care to individual genotypes and radically increasing the efficacy and capabilities of medicine through genetic profiling and pharmacogenetic testing to predict individual reactions (positive or adverse) to drug interventions (22). FDA regulation of this technology depends on whether genetic testing is offered as a kit for others to perform or as a testing service performed internally by the manufacturer. While kits are regulated by FDA as devices, testing services do not fit squarely within the FDA regulatory infrastructure. Companies and laboratories that perform testing services, sometimes by accepting samples through the mail, are not subject to FDA regulation to ensure safety, effectiveness, and market responsibility. Federal regulation is limited to the Clinical Laboratory Improvement Act and Amendments (CLIA) (23), which requires only that laboratories performing these tests meet standards for technical competence (24). However, laboratories also are licensed on the state level, and laboratories associated with academic and research institutions usually are subjected to oversight by Institutional Review Boards (IRBs) and adhere to prescribed human subject protocols. Similarly, some private laboratories have voluntarily subjected themselves to IRB oversight (24).

At this time, most genetic testing takes place in academic settings. According to a study reported in September 1997 by the Task Force on Genetic Testing, in comparison with biotech companies, twice as many nonprofit organizations are engaged in genetic testing (25). Beyond the context of newborn screening programs and diagnostic use for (monotype) conditions, regulatory uncertainty has raised reservations among health care providers and the general public, and thereby impeded the business of commercial predictive genetic testing. Notably, *predictive* genetic testing has raised a multitude of issues, such as the potential for employment and insurance discrimination based upon resulting information even when that information is obtained in an authorized manner.

National regulatory infrastructure for commercial predictive genetic testing may be introduced soon, but perhaps not by FDA, and certainly not in a voluntary manner. Although FDA regulates test kits and the quality of analyte-specific reagents used by clinical laboratories (26), clinical testing services too closely approximate the practice of medicine. FDA holds broad authority to regulate under the Medical Device Amendments to FDCA (15), but the Agency is expressly prohibited from interfering with the practice of medicine (24,27–29). Moreover FDA is preoccupied with pressing demands associated with the implementation of FDAMA. If needed regulatory infrastructure is introduced, it is likely to rise out of the work of the Secretary's Advisory Committee on Genetic Testing (SACGT) established in 1998 by the Secretary of the Department of Health and Human Services (HHS), or in conjunction with the mandate for federal medical privacy protection under the Health Insurance Portability and Accountability Act of 1996 (HIPAA) (30). Pursuant to HIPAA, on October 29, 1999, HHS released Proposed Standards for Privacy of Individually Identifiable Health Information (31) for public comment. Response to the Secretary's proposal also could move Congress to craft a legislative solution.

In contrast to the limited presence of genetic testing in the commercial consumer market, the technology is notably present in commercially sponsored research, including the design of clinical trials. With the increased role of genomics, informatics, and proteomics in drug development, the technologies are rapidly advancing in the research context. Presumably these genetic testing technologies will enter commerce in conjunction with the therapeutics they are instrumental in developing. Herceptin, a drug to treat advanced breast cancer that was introduced commercially by Genentech, Inc. in 1998, is representative of the next generation of pharmaceuticals and the first full generation of commercially viable genetic tests (22). Herceptin is tied to expression of the HER2 gene associated with an aggressive form of breast cancer, which is found in 25 to 30 percent of breast cancer patients. Genentech, in conjunction with another company, Dako, developed a genetic test to identify carriers of the HER2 gene. Although the HER2 precision of Herceptin has narrowed the drug's labeling and limited the size of the market for this drug, that same precision is enabling Genentech to sell the drug at a premium price of approximately $19,000 per treatment, roughly twice the cost of Taxol, another innovative cancer treatment (4).

Vaccines

Vaccines are reviewed and regulated within the context of therapeutics regulation as set forth above. However, traditionally the methodology for vaccine development has been to introduce weakened strains of the target viruses. The risk associated with this approach has necessitated extensive study of the animal and human cells in which the weakened viruses are grown, followed by animal trials that are generally more comprehensive than drug trials (32). FDA regulation of vaccines postapproval also is more intense than for drugs. FDA has established a comprehensive adverse event reporting system for vaccines, known as the Vaccine Adverse Event Report System (VAERS), which was introduced in response to the National Childhood Vaccine Injury Act of 1986 and

Table 2. Representative Vaccines Developed with Biotechnology

Name	Indication	Approval	Developer
Engerix-B	Prevention of hepatitis B in individuals suffering from chronic hepatitis C	August 1998 for the U.S. market	SmithKline Beecham
LYMErix	Preventions of Lyme disease	December 1998 for the U.S. market	SmithKline Beecham
Primavax	Active immunization (primary vaccination and booster) against hepatitis B, dipththeria, and tetanus in infants	February 1998 for Europe	Pasteur Merieux

is managed by both FDA and the national Centers for Disease Control and Prevention (33).

Only a handful of biotech-based vaccines have reached the commercial market. Some of those most recently approved are identified in Table 2 (34).

Nevertheless, biotechnology is revolutionizing vaccine development. The technology is being used to develop vaccines for myriad infectious diseases, including herpes, tuberculosis, and meningitis, and illnesses such as cancer (5). Biotechnology's primary contribution is to enable vaccines to be developed that do not expose patients to actual viruses. "Scientists can now produce a single viral protein that tricks the immune system, getting it ready to defeat the virus should it show up.... Strictly speaking, these are not 'vaccines' but 'therapies' because they are given after a patient comes down with the disease; but the process is the same" (5). Primary examples are AIDSVAX, an AIDS vaccine being developed by VaxGen (Brisbane, California) that has been reported to be close to FDA approval (5).

Unfortunately, advancement of these innovative vaccines now may be impeded by some recent events. Notably in October 1999 a Federal health advisory panel, the Advisory Committee on Immunization Practices to the Centers for Disease Control, withdrew a recommendation that all infants be immunized with RotaShield against rotavirus, a virus that causes a severe form of diarrhea. FDA had approved Rotashield for the U.S. market in August 1998 for oral administration to infants in a three-dose series at 2, 4, and 6 months of age (34). Subsequently Rotashield was linked to a painful and potentially fatal bowel obstruction known as intussusception (35). The federal government had licensed the vaccine from the manufacturer, American Home Products, a year earlier, and much of the work to develop the vaccine over 23 years was done at NIH (35). This incident closely followed other adverse incidents, including vaccine-associated polio and the death of an infant soon after receiving a dose of a pertussis (whooping cough) vaccine (32).

Tissue Products

The multidisciplinary field of tissue engineering couples cellular biology and chemical engineering, and much of the potential of this field is attributable to the ability of genetically engineered cells to assimilate into the environment in which they are placed (14). Cells are modified and cultivated to create body parts, including both implants and replacements. Potential and actual commercial applications include artificial skin, tendons, bone, corneas, bioartificial organs, blood and blood vessel substitutes, heart valves, expansion of bone marrow stem cells, neurological implants, tissue-engineered vascular grafts, various orthopedic devices such as tissue-engineered cartilage, and the use of artificial skin and its equivalents for toxicity testing (14). Scientists also are studying the use of modified human cells to treat viral infections, Parkinson's disease, and diabetes, as well as other disease conditions. The ultimate in tissue engineering may be gene therapy and xenotransplantation, or animal-to-human transplants (36).

Under the prevalent existing technology, for most of these applications, either cells are genetically modified to achieve a very specific genetic assimilation that will achieve a targeted therapeutic impact (e.g., gene therapies) and/or living tissue cells are applied to biodegradable plastic scaffolds. Apligraf, engineered by Organogenesis of Canton, Massachusetts, is one of the first commercially available products to contain human cells. Apligraf is a substitute skin approved by FDA in May 1998. Beyond applications such as Apligraf, the pipeline from this technology includes the possibility of tissue implants that deliver therapeutic drugs and hormonal secretions, and myriad gene therapies. At this time more than 3000 patients are in gene therapy trials (5). "From a health care perspective, the need for tissue-engineering products is unquestionable: 'In the United States each year there are 20,000 transplants [and] there are 2,000,000 implants.... The need for implantable parts and devices is staggering, and this need cannot be met through [traditional] organ and tissue transplantation'" (14,36).

FDA announced on September 30, 1999, that the Agency intends to establish a comprehensive new system for the regulation of human cellular and tissue-based products (37). The agency's proposals include amending current GMP regulations that apply to human cellular and tissue-based products — whether regulated as drugs, medical devices, and/or biologics — to incorporate into existing GMP regulations new (*1*) donor-suitability procedures (i.e., improved screening against the epidemiological dangers associated with conditions such as HIV and hepatitis) and (*2*) procedures for the proper handling, storage, and processing of these products (37). According to FDA, (*1*) the regulations will embody appreciation for the fact that there is a wide spectrum of products in this category that carry varying degrees of risk, (*2*) the regulations will be responsive to the specific level of risk of communicable disease

involved, and (*3*) this approach will avoid unnecessary and duplicative regulations. As stated by FDA (37).

> [T]he agency has tailored the proposed testing and screening requirements to the degree of communicable disease risk associated with the various types of human cellular and tissue-based products. The testing and screening for donors of cells and tissues that pose a high degree of communicable disease risk will be more extensive than for donors of cells and tissues with lesser risk. Where the risk is quite low (e.g., cells or tissues used autologously), FDA will recommend testing and screening, but will not require them; however, certain labeling will be required.

Food

Through the Center for Food Safety and Applied Nutrition, the FDA regulates $240 billion worth of domestic foods, $15 billion worth of imported foods, and $15 billion worth of cosmetics sold across state lines (38). FDA coordinates regulation of food with the United States Department of Agriculture, which regulates field testing (39), and the Environmental Protection Agency, which has jurisdiction over genetically modified plant and microbial pesticides under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) (14,40).

FDA has required notification and consultation from manufacturers before they market bioengineered foods, but FDA has required an official approval and labeling only in certain circumstances — such as when the foods contain (*1*) a known toxic substance, (*2*) nutrients different from those found in the unmodified version, (*3*) other new substances, (*4*) a known food allergen, or (*5*) antibiotic-resistant genes (41). However, FDA regulation of bioengineered foods presently is the subject of public and political scrutiny and debate, and is susceptible to change. Domestic pressure is mounting in response to media coverage of consumer concerns in Europe, which have fueled a trade war. The influence of consumer concern in the European Union (EU) is augmented by the EU's lack of a counterpart to FDA for food regulation (42,43). In response to these pressures, in October 1999 the FDA announced public meetings on bioengineered foods and an openness to reevaluation based on new scientific information (44).

Moreover the United States is expected to comply with the Biosafety Protocol (BP) agreed to by 120 to 30 nations in February 2000, following the November to December 1999 World trade Organization (WTO) meeting debacle in Seattle, Washington (45). The BP (*1*) requires exporters to label shipments that "may contain" bioengineered commodities, and (*2*) allows countries to block imports of genetically modified organisms (GMOs) on a precautionary basis in the absence of sufficient scientific evidence about their safety (45).

CHANGES TO, WITHIN, AND AROUND THE FDA

The accomplishments of contemporary life science place extraordinary external demands on FDA for change at a time when the agency is moving into a new era under the leadership of Commissioner Henney. Foremost, the Agency is working to implement a comprehensive collection of reforms required under FDAMA that are the ongoing focus of intense industry and patient group lobbying. Under FDAMA, the FDA's longstanding role of protector and promoter of public health now includes increased recognition of a responsibility to make innovative products available as expeditiously as possible, albeit without sacrificing safety and efficacy (2).

> "[P]ublic confidence in safety and effectiveness of the drug supply depends on a system of extensive premarket testing, but business competition, the expense of drug development, and, increasingly, the vocal advocacy of patient groups, create countervailing pressures on manufacturers and regulators to shorten the investigational period and to speed new drugs to market" (46).

Many FDA changes reflect responsiveness to breakthroughs associated with biotechnology, and there is sincere enthusiasm at the agency for this technology. However, concern that speedier agency approvals may be leading to a higher rate of adverse drug events (ADEs) and safety-related drug withdrawals (47) has increased tension between the goals of patient accessibility and safety. In June 1999, an FDA Task Force on Risk Management reported that both premarket and postmarketing safety-review programs would be bolstered (47).

The challenge before FDA to meet its restated mission is exacerbated by the fact that information technology now is integrating with all stages of drug development and health care to increase the pace of innovation and patient access to experimental treatments. The College of American Pathologists (CAP), a national medical specialty society with a membership of board-certified physicians that has significant interface with FDA on clinical laboratory issues, supports increased communication with all interested parties (47). CAP has recommended (47):

- Greater communication with and inclusion of stakeholders in all FDA activities — including development of policy and regulations and advisory committees;
- Better internal communication among FDA offices and between FDA centers;
- Clearly defined policy on the authority of guidance documents and the processes and circumstances surrounding their issuance; and
- More efficient use of the FDA's Web site for distribution of information to the general public.

The maturation of biotechnology will continue to have a profound impact on FDA well into this century. Moreover the agency continues to change in ways that directly affect biotech research, development, and commercialization. The primary forces of change presently bearing upon FDA and the agency's regulation of biotechnology are highlighted below.

Commissioner Henney

Dr. Jane E. Henney, an oncologist who served as FDA Commissioner Kessler's Deputy Commissioner for

Operations from January 1992 to March 1994, was appointed Commissioner of FDA on November 30, 1998. Prior to her appointment, Commissioner Henney served as Deputy Director of the National Cancer Institute and was Vice President of Health Sciences at the University of New Mexico.

Commissioner Henney succeeds Dr. Kessler, who left the agency in 1997. During the interim, Deputy Commissioner Michael Friedman negotiated PDUFA and FDAMA. Now Dr. Henney must fully implement the reforms associated with this legislation. In fact she has identified her priorities to include implementing the "letter and spirit" of FDAMA and strengthening the FDA's science base (49). Internally, Commissioner Henney must lead the transformation of the agency. During the period of self-reform preceding FDAMA, the need for extreme leadership, especially in the absence of a commissioner, necessitated heavy staffing in the Office of the Commissioner. Commissioner Henney now must shift FDA staff away from her office and refocus the Agency. Rather than self-assessment and self-reform, the agency must implement Congressionally mandated reforms while maintaining (even increasing) the speed of product review and approval, which is closely monitored by industry and patient groups. In addition, despite all these pressures and the challenges associated with escalating scientific innovation, Commissioner Henney must ensure that there is no catastrophic, widespread adverse event under her watch.

Modernization of the Agency

The passage of FDAMA in 1997 with virtually unanimous bipartisan support marked the culmination of years of intense lobbying by a coalition comprised of the biotechnology industry, pharmaceutical industry, patient groups, and the academic life science establishment. Industry and its allies used FDA's dependence on user fees under the Prescription Drug User Fees Act and the public's enthusiasm for accomplishments in life science to modernize the agency through codification of comprehensive changes designed to increase the agency's predictability, speed, accountability, and constructive communication with drug sponsors (4,14,42).

Implementation of FDAMA is an ongoing process, the outcome of which will be determined largely by the commitment of the interests responsible for its passage. FDAMA-mandated reforms focus on five general areas:

1. Reauthorization of the Prescription Drug User Fee Act of 1992 ("PDUFA II") with new performance goals for the FDA
2. Enhanced collaboration between manufacturers and the FDA throughout the approval process
3. Expansion of expedited drug and device approval tracks
4. Improvements in the economy and efficiency of manufacturing
5. Increased access to information for practitioners, health care organizations and consumers (14,50)

Moreover, through FDAMA, Congress has given FDA a mission that encompasses an obligation to promote public health by making new products available as quickly as possible, albeit without sacrificing safety and efficacy. Under FDAMA (2),

> The Administration shall—
>
> (1) promote the pubic health by promptly and efficiently reviewing clinical research and taking appropriate action on the marketing of regulated products in a timely manner;
> (2) with respect to such products, protect the public health by ensuring that—
> (A) foods are safe, wholesome, sanitary and properly labeled;
> (B) human and veterinary drugs are safe and effective;
> (C) there is reasonable assurance of the safety and effectiveness of devices intended for human use;
> (D) cosmetics are safe and properly labeled; and
> (E) public health and safety are protected from electronic product radiation;
> (3) participate through appropriate processes with representatives of other countries to reduce the burden of regulation, harmonize regulatory requirements, and achieve appropriate reciprocal arrangements; and
> (4) as determined to be appropriate by the Secretary, carry out paragraphs (1) through (3) in consultation with experts in science medicine and public health, and in cooperation with consumers, users, manufacturers, importers, packers, distributors, and retailers of regulated products.

The fees collected under PDUFA II provide an ongoing incentive for FDA to implement FDAMA while bestowing on the agency the financial means to maximize its performance; the FDA will collect an estimated $740 million in fees under PDUFA II (12,51). In accordance with PDUFA II and its FDAMA-prescribed mission, FDA has demonstrated initiative to engage in constructive interaction with industry by introducing regulatory infrastructure that would make such interaction standard operating procedure. For example, the FDA published a draft guidance in 1999 to introduce procedures for requesting, scheduling and conducting meetings between the FDA and sponsors (52). This draft guidance suggests that the agency is receptive to increased collaboration in the design of clinical trials, especially for making outcome determinations and establishing effectiveness.

FDA also has been responsive to the call under FDAMA and PDUFA II for the agency to establish performance goals that expedite the review processes. Several of the implementation regulations promulgated by the agency introduce response time lines (42). For example, FDA has issued a final rule that requires the agency to respond to a Humanitarian Device Exemption application (explained below) within 75 days of receipt (53). The agency also has issued a final rule stating that it will respond within 30 calendar days to a sponsor's written requests to remove a hold placed on the sponsor's clinical trails of a drug or biologic product (54).

Several key FDAMA provisions are especially beneficial to biotechnology. For example, FDAMA harmonized BLA

application procedures with those used for NDAs to introduce more uniform evaluation of products going to market. Now sponsors may file a single BLA, rather than a PLA and ELA (55). Also Section 112 of FDAMA introduced a "fast track" for new drugs that are "intended for the treatment of a serious or life-threatening condition and [demonstrate] the potential to address unmet medical needs for such condition" (2). A disproportionate number of biotech products qualify for fast-track designation, for biotech products are more inclined to be innovative and address unmet health care needs and serious health conditions as defined by FDA (7).

Orphan Drug Act and Humanitarian Devices

The biotechnology industry has been the commercial beneficiary of the Orphan Drug Act (56), which is subject to ongoing interpretation and modification (57). This act is applicable to drugs that treat rare diseases when the target patient population is not significant enough to make development of the drug economically feasible, hence the "orphan" label. "Before 1983, the development of these drugs was left to the benevolence of the drug companies, who would occasionally develop them as a public service" (56,58).

The Orphan Drug Act bestows direct economic incentives and incentives that facilitate the development process, many of which are highlighted in Table 3. Notably, the first applicant to obtain such designation and product application approval in the United States is entitled to market exclusivity for a period of seven years — meaning that no other company can market a molecularly identical drug for the FDA-approved use for seven years following approval of the original orphan (7). The period of exclusivity, which begins once the drug is approved, offers sponsors an opportunity for maximum pricing.

The Orphan Drug Act has been accomplishing its policy objective to the extent that it has resulted in a significant increase in the number of drugs available to treat rare diseases, and the sponsors of those drugs are largely biotechnology companies (58). However, the act is being criticized on the grounds that, in light of the benefits bestowed to sponsors, experimental treatments are not being made readily available to patients (59).

FDA approved nine orphan drugs in 1998, six of which were sponsored by biotech companies (7). The biotechnology industry's utilization of the act reflects that (*1*) many of the products being developed by the biotechnology industry are for genetic diseases that also happen to be rare diseases, and (*2*) the scope of protected markets under the Orphan Drug Act, although insignificant by most pharmaceutical company standards, often constitute a meaningful incentive for small biotech companies. Despite "orphan" designation, several of these drugs have reaped enormous profits for their sponsors — such as AZT (HIV infection /AIDS); pentamidine isethionate (pneumonia associated with AIDS); human growth hormone (hGH) (improper growth in children lacking the enzyme); erythropoietin (EPO) (anemia associated with end-stage renal disease); and CeredaseTM (Gaucher's disease) (58).

Sponsors are increasingly applying for orphan drug status, which means ongoing interpretation and regulatory modification, and the possibility of conditions being placed on orphan drug status where there is unexpected profitability. Moreover the EU now is close to fully implementing a counterpart to the act, and FDA has introduced a device counterpart in the form of analogous benefits for "humanitarian devices." Specifically, FDA has discretion to grant an exemption from the effectiveness requirements of Sections 514 (performance standards) and 515 (premarket approval) of the FDCA (15) after finding that a device:

- is designed to benefit a target disease population of not more than 4,000 individuals in the U.S.;
- would not be available otherwise;

Table 3. Benefits Under the Orphan Drug Act

Incentive	Description
Marketing exclusivity	First drug sponsor to have its application approved by FDA receives a 7-year period of market exclusivity against all other sponsors of the same drug approved for the same condition.
	Exclusivity is still subject to revocation based on whether (*1*) the sponsor fails to produce enough of the drug to meet demand, or (*2*) a competitor demonstrates clinical superiority — e.g., based on improved efficacy or diminution in adverse reactions.
Tax credits	Sponsors receive a tax credit for qualified clinical testing expenses, meaning a tax credit for 50% of the amounts spent conducting clinical trials.
Protocol assistance	FDA assists to distinguish exactly what tests and experiments sponsors need to complete in order to secure drug marketing approval.
	Objective is to enable sponsors to overcome the increased difficulty in structuring trials associated with the smallness of the sizes of orphan drug patient groups.
Grants	Sponsors may receive grants to defray the costs of qualified testing; this provision covers all testing after a drug is designated an orphan drug.
Open protocols	Sponsors may make drugs available to people not participating in their clinical trials while those trials are still ongoing; this creates an opportunity for sponsors to recoup some costs through charges to users before full approval.

- has no alternative available to treat or diagnose the disease or condition;
- will not expose patients to unreasonable or significant risk of illness or injury; and
- presents benefits to health from its use that outweigh associated risks.

However, there are some conditions. Devices granted this exemption may only be used at facilities that have an established institutional review board. Also, the humanitarian use must be approved by the board before studies begin (60).

Judicial Developments

In 1998, the pharmaceutical industry increased direct-to-consumer (DTC) advertising 23 percent to reach $1.32 billion (61). During that same year, the United States District Court for the District of Columbia enjoined FDA from enforcing pre-FDAMA guidance documents for off-label drug use (i.e., uses for purposes not reflected in FDA-approved labeling), concluding that FDAMA's labeling/marketing restrictions unduly burdened commercial free speech in violation of the First Amendment (62). Although the guidance documents permitted some dissemination of off-label information, they also restricted manufacturer use of textbooks, journal reprints and other educational materials that promote off-label use of drugs. FDA allowed distribution of printed and graphic materials addressing the safety and effectiveness of off-label drug use, provided that the manufacturer applied for FDA approval of the new use within six months of initial distribution. FDA was attempting to strictly limit dissemination through these media to avoid companies telling physicians about benefits but not risks. FDA wanted authority to review published studies before companies could give them to doctors. The agency also desired the power to require drug makers to give doctors additional studies of the drug, and the authority to limit distribution of studies until firms entered the federal review process for the additional uses.

The court deemed that the proposed FDA regulation was more extensive than necessary given less burdensome alternatives — for example, full, complete, and unambiguous disclosure by the drug's manufacture — and unduly burdened commercial free speech (42,62). The court granted summary judgment and issued a permanent injunction against the FDA restricting dissemination of information on off-label use of drug and medical devices. The court subsequently issued an amended order staying final determination of the legality of off-label drug use regulations implementing FDAMA, pending the filing of further submissions by FDA and the Washington Legal Foundation (63). On July 28, 1999, the District Court ruled that the FDAMA provisions requiring a supplemental application for the approval of dissemination of information concerning off-label drug use violate First Amendment commercial speech protections (42,64). The court has enjoined FDA and U.S. Department of Health and Human Services from prohibiting or restricting the following services:

- Dissemination of articles to doctors and medical professionals about off-label uses for drugs or medical devices when those articles are published in a bona fide peer-reviewed professional journal, regardless of whether the article focuses on non-FDA-approved uses.
- Dissemination of reference textbooks published by a bona fide independent publisher, regardless of whether they discuss nonapproved uses.
- Suggestion of content or speakers to independent continuing medical education (CME) program providers regardless of whether they discuss non-approved uses.

The net effect of these rulings is that pharmaceutical companies have greater leeway in the promotion of drugs for unapproved uses. Companies may give doctors copies of published medical studies that highlight the uses of drugs not approved by the FDA. However, the studies companies provide to doctors cannot be false or misleading, drug company sales representatives must disclose any association between the company and researcher, and the company must disclose whether the treatments detailed in the studies are FDA-approved.

From a policy perspective, these ruling have brought DTC marketing issues to the forefront and raised a call for empirical data that can provide guidance. An FDA survey by the Division of Drug Marketing, Advertising, and Communications (DDMAC) released in 1998 found that DTC advertisements for prescription drugs increased patient compliance with drug therapies. According to this survey, a solid description of the benefits and risks of a drug, compared to just name identification, induced greater consumer confidence (65).

CHALLENGES IN THE NEW MILLENNIUM

Biotechnology has raised both capabilities and expectations for life science to an unprecedented height, especially for medicinal applications. For biotechnology to realize its potential, both regulators and developers of this technology will have to overcome a multitude of emerging obstacles. FDA and those engaged in product development share the challenge of meeting expectations. They also are mutually dependent on each other's success in overcoming their respective challenges.

FDA

In the absence of a tragic mistake of a magnitude that shifts public opinion (Phen-Fen criticism has been mostly sponsor-specific), FDA will be under continued pressure to make new technology available. Moreover, as we enter an era of increased transparency and industry interaction, technology will continue to deluge the agency and demand greater scientific expertise. Given that these pressures will be coupled with accelerated patient access through mechanisms such as the fast track, more adverse events are probable. The FDA has identified the following as key challenges in its immediate and near future (66):

- Research and development (R&D) fueled pressures on regulatory responsibilities
- Greater product complexity driven by breakthroughs in technology
- Growth in recognized adverse effects associated with product use
- Unpredictable, new health and safety threats
- More targeted needs and awareness of citizen-shareholders
- Emerging regulatory challenges in the international arena
- Increased volume and diversity of imports
- Federal budget constraints

Product Developers

After decades of chemistry-based pharmaceuticals approved for broad market use that drew tremendous revenues throughout the duration of their sponsors' patents, biotechnology is introducing precision and lessening the hit-or-miss nature of therapeutics. Genentech's Herceptin for advanced breast cancer in women with the HER2 gene (approximately 35 percent of the women who have the disease, and the most aggressive form of the disease) marks the beginning of the forthcoming generation of pharmaceuticals (67).

Approved use and product labeling will be tied to genetic profiling, meaning that therapeutics will be coupled with genetic tests to determine their likely effectiveness and also susceptibility to adverse events. In terms of development, this precision will dramatically streamline clinical trials and accelerate FDA approval. FDA's standard for demonstrating "substantial effectiveness" under Section 355(d) of Title 21 of the United States Code, FDCA §505(d), no longer will demand multiple, independent clinical trials.

However, unless genetic profiling is utilized to create reimbursable, preventive care use of these therapeutics under managed care and such use is extensive enough to offset the streamlining of therapeutic use, pharmaceutical markets will fraction dramatically. Increasingly the combination of genomics, proteomics, and informatics will result in pharmaceuticals tailored to individual genetic profiles, thereby dramatically streamlining markets by historic standards. The net effect will be the introduction of a bounty of highly safe and effective therapeutics approved for tailored patient populations. In many ways today's orphan drugs could foreshadow tomorrow's mainstream pharmaceuticals.

The transition into this new era in life science will involve a series of pressing challenges. Arguably, the first collection of challenges are the most overarching and demand some of the more difficult and fundamental changes. Foremost, pharmaceutical innovation is shifting health care expenditure from hospital and provider care, including expensive surgeries and other specialized care, to drugs. Consequently the pharmaceutical industry increasingly is being blamed for the social and economic costs of contemporary health care, and the inadequacies and inequities of the U.S. health care system (68). Moreover there now is awareness that the pharmaceutical and biotechnology industries have been integrated through alliances. The biotechnology industry, once perceived as clusters of small, entrepreneurial companies with the mission of curing cancers and embodying many ideals of corporate America, is becoming recognized as big business, including, for example, commercial agriculture. Marketplace criticisms of commercial life science — such as pricing, aggressive direct-to-consumer (DTC) marketing, and limited labeling information — threaten to counterbalance public and political support for life science research, to the extent that many supporters of research are demanding a quid pro quo in the form of pricing controls.

These public pressures are rising at a time when the industry is about to shift from a period of record profitability to one of inability to meet sales forecasts despite extraordinary R&D expenditures. In addition to the factioning of consumer markets as addressed above, some 50 major patents held by pharmaceutical companies will expire by 2005, and each pharmaceutical company must produce 45 new drugs annually just to meet standard shareholder expectations and maintain market shares (7 percent annual sales growth) (69). "This will be a formidable task. The top seven pharmaceutical companies produce 45 drugs per year combined. Accelerating this development pace will move from being a competitive advantage to a competitive necessity" (69).

To remain competitive, pharmaceutical and biotechnology companies now are investing approximately $39 billion annually on R&D, $21.1 billion in the United States and $14.1 billion in the EU (70). The U.S.-based companies have more than doubled their R&D expenditure since 1990, when they spent $8.42 billion collectively (59). Moreover investment in R&D is expected to continue to rise — 11 percent this year alone (70). Development pressures are posing an extraordinary strain on pharmaceutical R&D operations. The focus of these operations is shifting from developing new drug targets to selecting among targets and significantly increasing the pace of advanced product development through better data management, clinical trial design, and regulatory submissions. To cut costs and accelerate the pace of clinical research, many pharmaceutical companies have shifted clinical research from academic medical centers and teaching hospitals to contract research organizations such as Quintiles and Parexel. At this time, Quintiles is reputed to be engaged in more clinical research than any other entity in the world.

The pharmaceutical industry also is attempting to lessen its dependency on outside clinical investigators and to reduce competition for patients. The latter has risen significantly in recent years within an overall increase in the amount of clinical drug development associated with accomplishments in biotechnology. Although patients have become more receptive to experimental treatments, they also are identifying these treatments on the Internet and gaining access to them via compassionate use and equivalent approvals rather than subjecting themselves to clinical trials that hold the risk of receiving placebos (71). In addition to turning to global research entities such as Quintiles, pharmaceutical companies are creating an alternative through simulated data. "Companies are

beginning to create populations of 'software people' designed to behave like the real thing and computer-based organs that can be used to test potential therapies before involving humans.... According to a report released by Pricewaterhouse Coopers LLP, New York, virtual trials will probably reduce the amount of clinical resources required in the short term by 10%" (70).

The regulatory accomplishment of FDAMA also poses some longer-term challenges to the developers of innovative commercial life science products. For example, by lessening the distinction between CBER and CDER, the sponsors of biologics are making themselves susceptible to market competition from generic drug manufacturers faced by traditional drug sponsors. Biologic drug products, meaning those approved under 42 U.S.C §262, are excluded from Title I of the Waxman-Hatch Act (72), meaning that they are not susceptible to competition from generics. But the traditional definition of biologic, as set forth under Section 351(a) of the Public Health Service Act and reiterated under section 123(e) of FDAMA, is extremely broad and is being diluted and blurred (18):

> The review and approval of biotech-derived products within the classic drug domain challenges us to define a true biologic, and in doing so brings entire classes of products closer to the realm of generic challenge....With so many scientific, technical, financial, legal, and political impediments, will there ever be a generic biological process? Yes. Maybe it won't appear as familiar as the ANDA [abbreviated new drug application], but the framework for a generic biologics process is taking shape.

CONCLUSION

Today is a glorious but challenging time for commercial life science. FDA now is being modernized and realigned with public and political enthusiasm for the accomplishments of biotechnology. The combination of genomics and informatics establishes a foundation for seemingly infinite scientific possibilities. In the absence of discernible diminishing returns in the laboratory, time and the limitations of imagination are the only firm checks on biotechnology's potential.

However, the accomplishments and potential of biotechnology to improve human health rest upon commercial and regulatory uncertainty, especially given weaknesses within the U.S. health care system. Although the inadequacies of this system are long-standing, they are being exacerbated by advancements in life science. A generation of breakthrough technology now is entering the market and beginning to dramatically expand the ability to treat. Yet this technology also, by introducing costly treatments where there were previously none and turning once fatal conditions into chronic conditions, is testing the limits of health care finance and increasing payer resistance to new technologies. This paradox is sobering, especially given that the primary opportunity cost is improvements to human health.

Life science is highly susceptible to regulation by its very nature, and the FDA is a political entity. The pharmaceutical industry faces extraordinary pressures, some of which will continue to be redirected toward FDA by industry and the people awaiting industry's products. Similarly FDA, positioned within the Department of Health and Human Services, will experience the tremors of a health care system realigning. How the economic realities of contemporary health care will affect the FDA's role in commercial life science and, in particular, the agency's regulation of biotechnology, is an open question. However, the combination of FDAMA and ongoing accomplishments in biotechnology are certain to provide an ongoing incentive for policy makers and public officials inside and outside the agency to work towards a resolution.

ACKNOWLEDGMENT

Associate Professor of Law, Widener University School of Law. The author wrote this article while a SmithKline-Beecham-sponsored Distinguished Fellow in Law and Genetics, Center for the Study of Law, Science and Technology, Arizona State University, School of Law. The background information relayed in this article summarizes a comprehensive, technical treatment presented in Michael J. Malinowski, *Biotechnology: Law, Business, and Regulation* (Aspen Law and Business/Little Brown, 1999). The author thanks Erica Rose for her thoughtful corrections and suggestions.

BIBLIOGRAPHY

1. Tufts Center for the Study of Drug Development, *Impact Report: Analysis and Insight into Critical Drug Development Issues*, July 1999.
2. Food and Drug Administration Modernization Act of 1997, Pub. L. No. 105-115, 111 Stat. 2296 (codified throughout 21 U.S.C.).
3. *Med. Ad. News* (July), 20–23 (1999).
4. Ernst & Young, *Biotech 1999: Bridging the Gap*, No. 48 (December 1998).
5. *Forbes*, May 31 (1999), 1999 WL 19911831.
6. J.M. Nardone and G. Poste, in M.J. Malinowski, ed., *Biotechnology: Law, Business and Regulation*, Aspen/Little Brown, Boston, MA, 1999.
7. M.J. Malinowski and N. Littlefield, in T. Caulfield and B. Wiliams-Jones, eds., *The Commercialization of Genetic Research: Ethical, Legal and Policy Issues*, Plenum, New York, 1999.
8. H. Miller, *Policy Controversy in Biotechnology: An Insider's View*, R.G. Landes, Austin, TX, 1997, p. 198.
9. 51 Fed. Reg., 23,302-93 (1986).
10. A. Katz-Stone, *Wash. Bus. J.* **P21**, (Special Report) (1998).
11. S. Krimsky and R.P. Wrubel, in *Agricultural Biotechnology and the Environment*, University of Illinois Press, Urbana, IL, 1996, p. 251.
12. Reauthorization of the Prescription Drug User Fee Act of 1992, Pub. L. No. 102-571, 106 Stat. 4491 (1992).
13. K.I. Kaitin and E.M. Healy, Drug Development Trends in the User Fee Era. Paper presented at the Annual Meeting of the American Society for Clinical Pharmacology and Therapeutics, San Antonio, TX, March 19, 1999.

14. M.J. Malinowski, in *Biotechnology: Law, Business and Regulation*, Aspen Law & Business/Little Brown, Boston, MA, 1999, ch. 11.
15. Federal Food, Drug, and Cosmetic Act, as amended by the Safe Medical Devices Act of 1990, Pub. L. No. 101-629, 104 Stat. 4511 (codified in scattered sections of 21 U.S.C. §301, 42 U.S.C. §§263b-263n (1968, as amended), and Medical Device Amendments of 1992, Pub. L. No. 102-300, 106 Stat. 239.
16. Public Health Services Act, 42 U.S.C. §262.
17. 21 CFR 600.3(h) (1994).
18. B. Zeid, *BioPharm.*, WL 14995071, March 1, 1999.
19. 21 CFR pt. 807.
20. 21 CFR 807.81(a)(3)(i).
21. 21 C.F.R. Pt. 803 (post-marketing reports, such as Medical Device Reports, on adverse incidents).
22. Boston Consulting Group, *The Pharmaceutical Industry into Its Second Century: From Serendipity to Strategy* **63**, 64 (January 1999).
23. Clinical Laboratory Improvement Act of 1967 and Clinical Laboratory Improvement Amendments of 1988, Pub. L. No. 100-578, 102 Stat. 2903 (codified at 42 U.S.C. §201 note, 263a, 263a note, 263a).
24. A. Huang, *Food Drug Law J.* **53**, 555, 558, 574 (1998).
25. Task Force on Genetic Testing, National Institutes of Health–Department of Energy, *Final Report: Promoting Safe and Effective Genetic Testing in the United States*, U.S. Government Printing Agency, Washington, DC, September 1997.
26. 62 Fed. Reg., 62,243 (November 21, 1997).
27. 37 Fed. Reg., 16,503 (August 15, 1972).
28. *Linder v. United States*, 268 U.S. 5, 18, 45 S.Ct. 446, 449, 69 L.Ed. 819 (1925).
29. *United States v. Evers*, 453 F. Supp. 1141 (M.D. Ala. 1978).
30. Health Insurance Portability and Accountability Act of 1996, §264, Pub. L. 104-191 (August 21, 1996).
31. U.S. Department of Health and Human Services, *Proposed Standards for Privacy of Individually Identifiable Health Information*, Available at: *www.aspe.hhs.gov/admnsimp/pvcsumm.htm*
32. Isadora B. Stehlin, *How FDA Works to Ensure Vaccine Safety*, Available at: *www.verity.fda.gov/search97*
33. U.S. Food and Drug Administration, *Adverse Event Reporting System*, Available at: *www.fda.gov/cber/vaers/vaers.htm*
34. S. Engel, *Med. Ad. News* 96–108, July (1999).
35. L.K. Altman, *N.Y. Times*, A10, October 23 (1999).
36. R.M. Nerem, in L.V. McIntire and F.B. Rudolph, eds., *Biotechnology: Science, Engineering, and Ethical Challenges for the twenty-first Century*, Joseph Henry Press, Washington, DC, 1996, p. 97.
37. 64 Fed. Reg., 52696 (September 30, 1999).
38. U.S. Food and Drug Administration, *Inside FDA: Center for Food Safety and Applied Nutrition*, Available at: *vm.cfsan.fda.gov*
39. U.S. Department of Agriculture, *Biotechnology: An Information Resource*, Available at: *www.nal.usda.gov/bic/*
40. Federal Insecticide, Fungicide and Rodenticide Act, Pub. L. No. 102-300, 106 Stat. §239, codified as amended, 7 U.S.C. §§136 et seq. (1984).
41. U.S. Food and Drug Administration, *FDA's Policy for Foods Developed by Biotechnology, FDA/CFSAN*, Available at: *http://vm.cfsan.fda.gov/lrd/biopolicy.html*
42. N. Littlefield and N. Hadas, *Food Drug Law J.* **55**, (Spring 2000).
43. *Economist*, 15, June 19 (1999).
44. U.S. Food and Drug Administration, *FDA Announces Public Meetings on Bioengineered Foods*, Available at: *vm.cfsan.fda.gov/lrd/hhbioeng.html*
45. *Economist*, February 5 (2000), 2000 WL 8140683.
46. J. Kulynych, *Food Drug Law J.* **54**, 127–128 (1999).
47. *Drug Topics*, 36, June 7 (1999), 1999 WL 10021762.
48. Thomas P. Wood, MD, President, College of American Pathologists, Correspondence, Dockets Management Branch, Food and Drug Administration (dated September 10, 1998).
49. J.E. Henney, *Food Drug Law J.* **54**, 1–4 (1999).
50. N. Littlefield and M.S. Webb, *Intellectual Property Rights News* 3 (Summer 1998).
51. U.S. Food and Drug Administration, *PDUFA II Five-Year Plan*, Available at: *www.fda.gov/oc/pdufa2/5yrplan/pdufa2.pdf*
52. CDER and CBER, *Draft Guidance for Industry on Formal Meetings with Sponsors and Applicants for PDUFA Products*, FDA, Washington, DC, February 1999.
53. 63 Fed. Reg., 59217 (1998) (codified at 21 C.F.R. pt. 814).
54. 63 Fed. Reg., 68676 (1998) (codified at 21 C.R.F. pt. 312).
55. 63 Fed. Reg., 40858 (1998). 63 Fed. Reg. 40858 (1998) (to be codified at 21 C.F.R. pts. 3, 5, 10, 20, 207, 310, 312, 316, 600, 601, 607, 610, 640, and 660) (proposed July 31, 1998).
56. Orphan Drug Act, Pub. L. No. 97-414, 96 Stat. 2049 (1982), codified as amended at 21 U.S.C. §§360aa–360ee (1994), 26 U.S.C. §45C (Supp. II 1994), 42 U.S.C. 236 (1994).
57. U.S. Food and Drug Administration, *The Orphan Drug Regulations*, Available at: *www.verify.fda.gov*
58. G.A. Pulsinelli, *Santa Clara Computer High Tech. Law J.* **15**, 299, 301, 316–324 (May 1999).
59. D. Grady, *N.Y. Times* November 16 (1999).
60. U.S. Food and Drug Administration, *Requirements of Laws and Regulations Enforced by the U.S. Food and Drug Administration*, Available at: *www.fda.gov/opacom/morechoices/smallbusiness/blubook.htm*
61. *F-D-C Reports, The Pink Sheet*, No. 36 (April 26, 1999).
62. *Washington Legal Foundation v. Friedman*, 13 F. Supp.2d 51 (D.D.C. 1998).
63. *Washington Legal Foundation v. Friedman*, 36 F. Supp.2d 418 (D.D.C. 1999).
64. *Washington Legal Foundation v. Henney*, No. 94-1306 (D.D.C. 1999).
65. *F-D-C Reports, The Pink Sheet*, No. 6 (August 17, 1998).
66. FDA, *Food and Drug Administration Modernization Act of 1997: FDA Plan for Statutory Compliance* 5, FDA, Washington, DC, November 1998.
67. R. Bazell, *Her2: The Making of Herceptin, a Revolutionary Treatment for Breast Cancer*, Random House, New York, 1998.
68. D. Rosenblum, *N.Y. Times*, Sect. 4, November 14, pp. 1, 16 (1999).
69. *Med. Ad. News* (July), 3–13 (1999).
70. A. Humphreys, *Med. Ad. News* **64**(July), 1 (1999).
71. G. Kolata and K. Eichenwald, *N.Y. Times*, A1, 30, October 3 (1999).
72. Waxman-Hatch Act, Title I, Pub. L. No. 98-417, 98 Stat 1585 (1984).

See other entries ANIMAL AND ANIMAL MEDICAL BIOTECHNOLOGY; FEDERAL REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE, FDA, ORPHAN DRUG ACT; GENE THERAPY, LAW AND FDA ROLE IN REGULATION; GENETIC INFORMATION, LEGAL, FDA REGULATION OF GENETIC TESTING; HUMAN SUBJECTS RESEARCH, LAW, FDA RULES; RESEARCH ON ANIMALS, LAW, LEGISLATIVE, AND WELFARE ISSUES IN THE USE OF ANIMALS FOR GENETIC ENGINEERING AND XENOTRANSPLANTATION; TRANSGENIC ANIMALS: AN OVERVIEW.

FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, CONGRESS, EPA

WENDY E. WAGNER
Case Western Reserve University
Cleveland, Ohio

OUTLINE

INTRODUCTION

The Environmental Protection Agency's (EPA's) regulation of biotechnology products provides a fascinating case study of the challenges that this important, but still rapidly evolving technology presents to a large federal bureaucracy. In the absence of specific directions from Congress, EPA has used existing regulatory programs originally intended for conventional chemicals to address the new and often very different regulatory challenges posed by biotechnology. The results of EPA's regulatory effort are often disappointing, both in creating clear and effective rules to guide the biotechnology industry and in ensuring adequate public input into important choices about the future. The specific roles that EPA currently plays both as regulator and promoter of biotechnology and the problems that continue to plague EPA's biotechnology programs are the subject of this article.

THE HISTORY OF EPA'S INVOLVEMENT IN BIOTECHNOLOGY

In many areas of public health and environmental protection, Congress has developed legislation that is quite specific with regard to the methods and targets of regulation of environmental and public health risks (1). Yet Congress still has not provided specific legislative direction on the regulation of biotechnology products, despite the several decades over which the public has expressed concern about biotechnology's long-term safety. EPA's regulation of biotechnology thus does not derive from Congress' grand biotechnology plan. Rather, the regulation comes from authorities EPA and other executive branch authorities have read into existing statutes intended for the related, but still quite different purpose of regulating industrial chemicals in food, drugs, pesticides, and consumer products (2).

Congress' silence may be attributed in part to its general pattern of avoiding public controversies until a crisis forces legislative action (3). But Congress may have also been forestalled by aggressive Executive Branch initiatives to fend off such legislation because of White House concerns that Congress might overreact to the public's fear of biotechnology products and restrict the industry in damaging ways (4). In 1974 the Office of Science and Technology Policy (OSTP)—a White House agency—began developing a framework for coordinating the various federal agencies' regulation of biotechnology, taking as its premise that the products of biotechnology do not demand new statutory attention and that the agencies have sufficient existing statutory authorities to regulate these products (5,6). In 1986 the OSTP published its *Coordinated Framework for Regulation of Biotechnology*, which directs agencies to use existing statutory authorities to regulate biotechnology and "to operate their programs in an integrated and coordinated fashion that together should cover the full range of plants, animals and microorganisms derived by the new genetic engineering techniques" (6). Moreover, "[w]here regulatory oversight or review for a particular product is to be performed by more than one agency, the policy establishes a lead agency,..." (6). In 1992 the OSTP provided still further direction by instructing agencies to utilize a risk–benefit test for determining whether and how to regulatory biotechnology products (7).

Because of its statutory mandates to regulate pesticides and other commercial chemical substances not regulated elsewhere, EPA acts as the lead regulatory agency under the OSTP framework for overseeing the safety of these two biotechnology products (8,9). EPA has not always been an enthusiastic participant in the OSTP's coordinated framework, however; it has periodically expressed concern about this "entirely new area of responsibility" in biotechnology regulation that extends well beyond its traditional regulatory mission of regulating only conventional chemical products and wastes (10). Ironically, in fact, even when EPA has passed biotechnology regulations in accordance with OSTP policies, the regulations have occasionally met White House resistance because of the costs they impose on the important and rapidly growing biotechnology industry (11,12).

In considering EPA's programs, it is important to keep this larger Executive Branch context in mind. Professor William Rodgers has characterized the federal government's regulation of biotechnology as consisting of a

number of "regulatory agencies with truncated responsibilities ... spread across the landscape like a minefield with its full potential untested" (8,p. 503). Even more charitable characterizations of existing regulatory programs cannot avoid recognition of the adverse consequences that emerge from the current legislatively deficient approach to biotechnology. Because its programs are so rudderless, the EPA's regulations seem both slow to develop and lacking in coherence and comprehensiveness.

EPA AS REGULATOR OF BIOTECHNOLOGY

Most of EPA's regulatory authority over biotechnology targets two categories of products: pesticide products, and other products of biotechnology that are not pesticides, drugs, food additives, or strains of plants. In far more limited circumstances, EPA may also have the authority to regulate the release of biotechnology products into specific environmental media. These different types of regulatory programs and authorities are discussed in turn.

Pesticides

In the Federal Insecticide and Rodenticide Act (FIFRA), passed initially in 1947 and amended significantly throughout the 1970s, 1980s, and 1990s (13), Congress tasks the EPA with the daunting responsibility of establishing a testing and approval process for new pesticide products, reevaluating and reregistering existing pesticides, and overseeing the safety of the distribution, sale, and use of pesticides (14). In these responsibilities Congress has instructed EPA to ensure that registered pesticides will not present "unreasonable adverse effects on the environment" (15). Since a number of pesticide products are now developed with biotechnology, such as genetically engineered microbes and genetically altered plants that produce their own pesticidal substances ("plant pesticides"), EPA has included these biotechnological pesticides within its regulatory purview. Expectedly, however, biotechnological pesticides present a host of unique characteristics that raise new, and to some extent more difficult challenges to regulators. In response to these challenges, EPA is gradually creating additional subprograms within its pesticide unit that exclusively address the unique products of biotechnology.

Traditional Pesticide Program. The most significant regulatory task for EPA under FIFRA is the registration of all new pesticides. This registration functions predominantly as a "preclearance regulatory regime where adequate study is supposed to be a precondition to commercial use" (8, p. 506). Accordingly, before marketing a pesticide, a manufacturer of a new pesticide must submit data in support of registration that meet EPA requirements for the particular type of pesticide product. These requirements are less onerous than the pre-approval requirements for new drugs under the Food, Drug, and Cosmetic Act described elsewhere in this encyclopedia, but they are nevertheless resource-intensive and time-consuming (16).

For conventional chemical pesticides, a manufacturer follows a standard testing regimen that usually occurs without EPA oversight. The testing and data requirements are standardized in the Code of Federal Regulations (17). EPA oversight begins only when the requisite large-scale testing commences (on at least 10 acres of land or one surface-acre of water). At this point, the manufacturer must secure an Experimental Use Permit (EUP) from EPA (18).

After the necessary testing has been completed, the manufacturer submits a final registration application to the EPA. Based on the application, EPA makes an ultimate determination to approve, reject, or condition approval of the pesticide based on evidence of its safety. EPA's final decisions can be appealed by adversely affected parties in the federal courts (19). After receiving EPA's approval, the manufacturer generally can commence marketing.

Biotechnology Pesticides. Although conventional chemical pesticides continue to dominate its regulatory agenda, EPA has interpreted FIFRA to provide it with regulatory authority over all pesticides, which also include pesticidal substances produced by living organisms (11). But regulating these biological pesticides, especially those biologics produced using biotechnology, requires more specialized programs that adjust to the different scientific uncertainties, as well as to the different potential benefits and risks posed by a genetically engineered organism. EPA is in the process of developing two specialty programs — one for genetically engineered microbial pesticides and one for plant pesticides.

Genetically Engineered Microbial Pesticides. Microbial pesticides are those that employ microorganisms (e.g., bacteria, fungi, algae, and protozoa) and viruses for pesticidal purposes. Although EPA often allows microbial pesticides to undergo less rigorous testing for FIFRA registration, microbial pesticides that are the product of genetic engineering are determined to require slightly higher regulatory intervention (21) because of concerns that these nonnative microbes could spread into the environment in unexpected and even catastrophic ways (22).

As might be expected, EPA's first hurdle in regulating genetically engineered microbial pesticides is defining what they are in terms that are clear but sufficiently flexible to accommodate the often case-by-case safety issues posed by these organisms. After several years of effort, EPA promulgated a definition in 1994 that defines biotechnological microbial pesticides as those with "pesticidal properties [that] have been imparted or enhanced by the introduction of genetic material that has been deliberately modified" (23). Considerable debate surrounds this definition. Scientists criticize EPA's process-based definition as overinclusive, including within its regulatory reach biotechnology products that are almost riskless but yet which must undergo extended regulatory scrutiny simply because of the process by which they are produced (24). Others, including the EPA, have acknowledged the rough cut of this definition but have concluded after surveying alternatives that it provides the greatest clarity to regulated entities, a value that offsets the costs of overinclusiveness (25). Indeed, a survey of most environmental and public health environmental programs reveals similar, bright-line jurisdictional determinations for what will be included within a regulatory program (26).

The second step of EPA's pesticide program for biotechnology products concerns clarifying the requirements for registration. There appear to be two points of significant divergence in the registration of genetically engineered microbes as compared with EPA's treatment of conventional chemical pesticides. The first difference occurs with EPA's oversight of testing activities. In contrast to EPA's oversight of conventional chemical and natural biological pesticide testing, manufacturers of genetically engineered microbial pesticides must notify the EPA before conducting even small-scale testing (less than 10 acres). Manufacturers of biotechnology microbes can avoid this regulatory oversight only if the testing facility meets prescribed containment conditions or if the engineered pesticides are believed to be low risk (27). Despite criticism for this early regulatory intervention during the testing process (8), EPA has consistently justified it as necessary because of the unusual and unpredictable risks associated with accidental release of genetically engineered microbes into the environment (28).

Manufacturers intending to conduct this testing must provide EPA not only with notification, but with data that will provide EPA sufficient information upon which to judge the application for testing (29). After notification, EPA has 90 days to conduct an assessment to determine whether or how to permit the testing (30). Testing may not commence until after EPA has issued its determination (31). If the manufacturer is dissatisfied with either the conditions for testing or EPA's rejection of their application, it has available both regulatory and ultimately judicial avenues for challenging the agency's determination, although great deference is given to the agency during these layers of review (32).

The second difference from conventional chemical pesticides are the regulatory requirements themselves. Data requirements and registration reviews of genetically engineered microbes are often individually tailored to the risks and benefits presented by a particular product. Individual manufacturers are advised to contact the agency directly, rather than commence testing in accordance with EPA's FIFRA regulations (33). EPA's review of the applications, including its use of risk assessment protocols, also appear to be undertaken on a case-by-case basis (34).

Plant Pesticides. EPA has also been working on a specialty program for plant substances that are genetically engineered (using recombinant DNA techniques, not through traditional plant breeding) to produce more or additional natural toxins within the plant that repel specific pests — termed "plant pesticides" (35). Although the program is not yet final, considerable effort and discussion has been dedicated to determining whether and the extent to which EPA should regulate genetically engineered plant pesticides.

Under its proposed policy, EPA considers the active ingredient of a plant pesticide (which is EPA's target of regulation under FIFRA) to be both the substance produced in the plant having the pesticidal effect as well as the genetic material necessary for the production of the substance (36). This EPA policy has been the subject of scientific controversy since it involves regulating the "inherited traits of plants" that "cannot be separated for regulatory purposes from the plant itself," which is quite different from the traditional external pesticides EPA typically regulates under FIFRA (37). Industry has criticized the program as well, both because it includes within its regulatory reach low- to no-risk plant pesticides, and because it may ultimately impose "substantial costs on plant breeding" and discourage the development and use of genetically engineered crops (34,38).

Because certain genetically eng ineered plant pesticides are considered to be virtually no risk, EPA will likely develop a list of "exempted" plant pesticides in its final plant pesticide policy (25,39). EPA is also considering whether pest resistance management plans should be required as part of the approval process for plant pesticides, although industry maintains that it is already voluntarily dedicating considerable resources into pest resistance management (40). For example, in its approval of Bt corn products (genetically modified corn that contains a bacterium to resist pests), EPA did condition registration on adherence to specified refuge requirements, although the pesticide resistance plans were produced initially by the Bt corn industry (41).

Under EPA's proposed program the nonexempted plant pesticides will still be regulated under the traditional FIFRA process. Differences from EPA's regulation of conventional chemical pesticides may occur, however, with regard to the requirements for testing plant pesticides, data requirements for registration, and conditions for approval (42). In fact, conditions for approval — such as warnings — may be particularly problematic for plant pesticides (43). For example, warning users of a plant pesticide to avoid using the plant within danger areas (e.g., near a water body) may be insufficient to keep the plant from spreading to that area naturally. Thus EPA ultimately may reject applications for some plant pesticides because of the realistic obstacles to restrictive labeling, even though in theory the plant could be used safely if warnings could be provided.

Pesticide Residues

EPA has also been charged with regulating pesticide residues in raw agricultural commodities and processed foods under the Food, Drug, and Cosmetic Act (FDCA) (44). This responsibility generally requires EPA to set tolerances for pesticide residues based on risks posed by human dietary consumption. Thus, unlike FIFRA which requires EPA to consider *all* environmental and human health risks associated with release of a pesticide into the environment, FDCA requires EPA only to consider how much pesticide residue can remain on food products to ensure protection of the public health.

With regard to products of biotechnology, it is expected that the FDCA will pose few if any additional requirements. EPA is developing exemptions for certain plant pesticides because of the extremely low risks presented from consumption (45). These exemptions target various plant pesticides that do not result in significantly different dietary exposures, that are derived from nucleic acids, or that consist of plant virus protein coats (46).

"Toxic Substances" Not Otherwise Regulated

Under the Toxic Substances Control Act (TSCA), EPA is authorized to regulate chemicals substances to ensure their safety. This includes the authority to screen new chemicals and to impose controls on those that pose "unreasonable risk of injury to health or the environment" (47). EPA's TSCA program, like its FIFRA program, has been stretched to include the regulation of biotechnology products. The future uncertainties regarding biotechnology products, coupled with the constraints imposed by the TSCA statute that was passed with conventional chemicals in mind, continue to reveal gaps in EPA's authority and appear to pose impediments to a more coherent and comprehensive regulatory program.

The major features of EPA's TSCA program mirror FIFRA in the sense that EPA must first define those products of biotechnology that fall under the jurisdiction of TSCA, since the statute itself provides very little guidance on the matter. Once the jurisdictional bounds have been defined, EPA then reviews biotechnology products in accordance with its recently conceived and evolving specialty rules.

General. The primary emphasis of TSCA is to require all "new chemical substances" to go through a registration and approval process that evaluates, however cursorily, their safety (48). These registration requirements are considerably less onerous than those of FIFRA or the FDCA. Before marketing a new substance, the manufacturer must submit a pre-manufacture notice (PMN) that provides all available data regarding the health and environmental safety of the substance (49). The most notable contrast with FIFRA and FDCA is that if the data are incomplete, there is *no* affirmative obligation on the manufacturer to conduct the research itself unless EPA specifically requires additional testing: The manufacturer must include in its notice only the health or environmental test data in the manufacturer's possession or control (50). EPA's ability to require additional testing is read to be justified only after EPA is able to show that the substance presents the possibility of an unreasonable risk or that it will be produced in substantial quantities. Therefore EPA's authority to require added testing, as a practical matter, is generally constrained for substances where the scientific uncertainties are particularly high or the safety testing scant (51). Indeed, some have suggested that the design of TSCA may tend to lead to less safety testing rather than more (52).

After filing the PMN with all available information, EPA has 90 days within which to consider the application (53). If it believes that the information leads to a conclusion that the substance might be risky, it can require additional testing or deny the application (54). Perhaps because of the time, resource, and burden of proof and related statutory problems that the EPA encounters in administering TSCA, the denial of an application is a rare event for conventional chemical substances (55). Since the newly emerging products of biotechnology are only now moving through the requirements of TSCA, it is difficult to tell how rigorously TSCA will be applied to this industry.

Specific. Recurrent biotechnology-related problems continue to arise under TSCA that have led EPA to create specialty regulations to streamline the process for products of biotechnology. Most notable are added requirements for research and development, a tailor-made review process for biotechnology microbial products, and bright-line exemptions for small quantity users and biotechnology products that are presumptively low risk. Before detailing these specialty programs and requirements, however, the threshold jurisdictional question must be resolved: Is the manufacturer seeking to market a biotechnology product that EPA considers within the jurisdictional reach of TSCA?

Jurisdictional Reach. TSCA is a "gap-filler" statute that applies largely to new chemical substances in commercial use that are not regulated elsewhere as food additives, drugs, or pesticides (8,56). In accordance with this statutory definition, there are particularly thorny jurisdictional issues that must be resolved before EPA will assert its regulatory authority over a biotechnology product, all of which surround the definition of "chemical substance." In interpreting what constitutes a "chemical substance" (57), EPA has determined that living organisms may be included within the definition. In order to limit the consequences of this generous reading of its regulatory jurisdiction under the statute, EPA currently has restricted its regulatory program to microorganisms "used in conversion of biomass for energy, pollutant degradation, enhanced oil recovery, metal extraction and concentration, and certain non-food and non-pesticidal agricultural applications, such as nitrogen fixation" (58).

EPA's regulatory jurisdiction over products of biotechnology is still further limited by its definition of when a microorganism is genetically engineered — only when it results from the "deliberate combination of genetic material originally isolated from organisms of different taxonomic genera" (59). Like FIFRA, this definition results from the EPA's attempt to develop a clear, yet generally inclusive definition based on perceived risks of current products of biotechnology (60). The TSCA regulatory coverage of only commercial research further constrains EPA's regulation of emerging biotechnologies, although to a much more limited extent (12).

Once within the jurisdiction of TSCA, the manufacturer must make itself aware of additional requirements and exemptions that govern EPA's regulation of biotechnology.

Added Considerations and Requirements. The approval process for genetically engineered microbes on its face appears similar to the process for conventional chemicals. In order to commercialize new genetically engineered microbes, the manufacturer, importer, or processor must submit to EPA the biotechnology equivalent of PMNs, termed "Microbial Commercial Activity Notices" (MCANs) (61). EPA then has 90 days to determine whether the organism "may present an unreasonable risk to human health or the environment" (62,63). But several noteworthy differences in biotechnology products present new questions for this established regulatory program. None of these questions appear to have been anticipated by Congress when it passed TSCA in 1976. Some of these questions also appear to remain unresolved by EPA (11).

First, EPA has focused almost exclusively on genetically engineered microbes as the covered set of biotechnology products. While these appear to be the largest category of biotechnology products not otherwise regulated under other statutes, it is not the only set of products that could in theory fit within the statute. One commentator, for example, has suggested that genetically engineered fish could be included within the jurisdictional reach of TSCA (64). Thus the classification of biotechnology products that EPA will regulate under TSCA remains an open issue for debate and possibly the subject of future regulatory expansions.

A second regulatory problem that arises is similar to that occurring for genetically engineered microbial pesticides — the uncertain risks associated with release into the environment and the resulting perceived need for earlier regulatory intervention during the testing stage (65). Thus, in contrast to conventional chemical substances where "small quantity" research and development is exempted from regulatory oversight, EPA maintains an active oversight role in research and development of genetically engineered microbes and requires manufacturers of these microorganisms to submit abbreviated applications (called "TSCA Experimental Release Applications" or TERAs) 60 days before testing intergeneric organisms in the environment (66). In an effort to minimize the adverse impact of this oversight on low-risk genetically engineered microbes, EPA is also promulgating a rule that will exempt from regulatory oversight certain low-risk biotechnology products used in research and development activities (67).

Third, and relatedly, standard reporting requirements and assessment procedures used for conventional chemicals often fall short of providing helpful guidance in overseeing the safety of genetically engineered microbes. For example, the TSCA section 8(e) reporting requirement that manufacturers report "substantial risks" arising from their chemical substances focuses largely on potential health risks such as cancer, rather than on the release and spreading concerns posed by the use of genetically engineered products (68). A similar complaint has been made regarding the use of EPA's risk assessment protocols that were originally developed for traditional TSCA chemical substances. Although EPA has responded to these discontinuities by providing guidance on the types of information to submit for MCANs (69), and in relying heavily on ecological risk assessment protocols for reviewing these applications (70), these adjustments are far from complete or streamlined (70). As a result, EPA appears to determine reporting requirements for biotechnology projects on a case-by-case basis, with the resulting requirements being much less predictable than the more standardized reporting requirements that apply to conventional chemical products (71). In fact, to date, EPA's sole decision under TSCA to approve the commercialization of a genetically engineered microorganism for release into the environment (RMBPC-2, a soil bacterium that fixates nitrogen for alfalfa plants) involved nine years of data (including five years of test data) on the product, generated some controversy within the EPA's scientific advisory panel, and ultimately attained approval conditioned on limited production and monitoring requirements (72).

EPA has also recognized that in some cases complete exemptions from TSCA oversight are justified for genetically engineered microbes for which the risks can be reliably predicted in advance to be minimal to nonexistent. Thus, with the assistance of several years of deliberations, input from its Science Advisory Board, and the public comment process, EPA has also promulgated a variety of general and research-specific exemptions for certain types of genetically engineered microbes (73).

Miscellaneous Other Statutory Authorities

EPA may also have authority under other statutes to regulate the release of a wide variety of genetically engineered organisms into the environment. Under the Clean Water Act, for example, EPA may be able to justify regulation or restriction of the introduction of bioengineered fish into U.S. rivers and streams through its authority to restrict the discharge of "pollutants" into navigable waters (64,74,75). Both the Resource Conservation and Recovery Act (RCRA) and the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) may provide EPA with the ability to regulate or require cleanup of genetically engineered organisms that can be characterized as "hazardous wastes" or "hazardous substances" respectively (75). Finally, the Endangered Species Act, with its prohibition against impairing the critical habits of endangered species, may also provide mechanisms for regulating the introduction of biologically engineered organisms into these often narrowly bounded environments (64,76). None of these authorities appear to have been used by EPA to restrict the release of genetically engineered organisms, although their potential has been recognized by commentators.

EPA AS PROMOTER OF BIOTECHNOLOGY

Although EPA's relationship to the biotechnology industry is primarily one of regulator, EPA also acts in a much more limited capacity as a promoter of biotechnology through its effort to carve out administrative exemptions to existing regulatory programs and through its small grants program for research. EPA's activities in streamlining regulatory programs and creating exemptions to regulatory oversight when the risks of a biotechnology product are minimal are motivated in part by the Executive Branch's relatively consistent support of biotechnology. The agency's efforts to minimize regulatory impediments to biotechnology innovation also result from its conviction that many genetically engineered organisms — such as plant pesticides — are safer and more effective than conventional chemical products.

EPA's role as a promoter of biotechnology not only consists of its effort to minimize the regulation of biotechnology products where justified, but to encourage research in areas like bioremediation and related fields (77). Grants administered through EPA provide some support for biotechnology research activities (78). Most notable is EPA's continued support of genetically engineered bioremediation technologies for hazardous waste cleanup (79). It seems likely that EPA's encouragement of these developments will only increase over time.

PROBLEMS

Biotechnology regulatory programs that are shoe-horned into regulatory programs originally designed for conventional chemicals run the risk of not only inheriting the existing weaknesses of these programs, but exacerbating the weaknesses by stretching the statutes to include the new and very different regulatory challenges posed by products of biotechnology (80). EPA's regulatory programs appear to confirm these dismal predictions.

Incoherent and Incomplete Regulatory Programs

National policy regarding the regulation of biotechnology remains muddled in the separate and often quite different regulatory programs of the various agencies, such as FDA, EPA, and the U.S. Department of Agriculture (USDA). The resulting national regulatory approach has been condemned by some critics as "develop[ed] in response to jurisdictional arguments that can be made to protect and project agency authority instead of in response to national policy considerations" (2, p. 241). EPA has in fact conceded that some important gaps remain in its ability to fit biotechnology products into existing regulatory programs. Regulatory oversight of genetically engineered wildlife and fish remain problematic with regard to existing statutory authorities (64). Existing gaps in regulatory authority will likely grow only more serious as the biotechnology industry expands. Yet in some circumstances, parties injured by inadequately regulated releases of biotechnology products may be left with disappointing legal remedies precisely because existing federal programs could be read to preempt their common law claims (81).

Perhaps equally serious is the complex and often unpredictable nature of EPA's regulatory programs for the biotechnology industry (82). EPA's regulatory programs are not only confused by the fact that its statutory authority never mentions or even alludes to biotechnology, but also because EPA's programs must comply with the Executive Branch biotechnology policy statement that at times appears quite remote or even at odds with the purported original grant of legislative authority (2). Although EPA has dedicated considerable effort to defining its jurisdiction for biotechnology and developing specialty programs to address some of the more categorically different problems that biotechnology presents to existing programs, considerable regulatory uncertainties remain. For example, to determine what sorts of pre-approval testing requirements apply to biotechnology pesticides, a manufacturer must generally *meet* with EPA to hammer out the case-specific tests. The resultant "uncertainty costs" are of continuing concern to the biotechnology industry (83). In contrast, manufacturers of conventional chemical pesticides can generally determine their testing requirements by reference to the Code of Federal Regulations. To make matters worse, the regulatory review process for biotechnology products also tends to be considerably more unpredictable with regard to its costs, delays, and ultimate outcomes.

The absence of any solid legislative grounding to EPA's biotechnology regulatory programs also subjects these programs to seemingly endless waves of revisions, exemptions, and other regulatory changes without a clear, overarching regulatory plan. Rather than falling ahead of the curve and engaging in important regulatory planning for the inevitable biotechnological future, EPA seems often to be endeavoring to catch up to last year's problems. This "regulatory gap" then produces a damaging "false start" for an industry that invests in innovation only to find that their ultimate marketing is severely restricted (84). The resulting outdated, or at least constantly shifting, nature of EPA's programs likely also takes its toll on public understanding and support (85).

Scientific Uncertainties that Complicate Regulation

EPA's effort to fit biotechnology regulation within existing regulatory structures also exacerbates preexisting weaknesses in these statutory programs (2,75). Both TSCA and FIFRA regulatory programs have been criticized for their unjustified over-reliance on quantitative risk assessment, even with regard to assessing the risks of conventional chemicals (2,63,86). In the highly uncertain area of biotechnology, the obstacles to reliable quantitative risk assessments and cumulative impacts are still more daunting. Because the uncertainties often overwhelm the information that science is able to provide, EPA is left with analytical tools that are sorely inadequate, but upon which Congress, the public, and the courts consistently demand that EPA's regulatory judgments be based (2,86). In order to meet these external expectations and demands, EPA may overstate the scientific grounding of its risk assessments for biotechnology products, a tendency that likely improves the short-term credibility of its regulatory actions, but that over time may undercut public understanding, participation, and even support for some of the agency's underlying policy judgments (2). As one author has argued, "[m]aking quantitative risk assessment the sole basis for biotechnology regulation ... is likely to increase public anxiety, mask important political choices in purportedly neutral, scientific terms and ultimately, fail to consider many of the potential hazards presented by biotechnology" (2, pp. 249–250).

In addition, existing statutory programs appear to assume that added testing will resolve remaining, material uncertainties, an assumption that necessarily tends to overemphasize those risks that can be quantified (2). The statutory policies of FIFRA, for instance, tend to presume that a battery of tests will adequately assess the health and environmental risks of pesticides, an assumption much more appropriate for pesticides posing primarily cancer risks than for biopesticides that may behave in unpredictable ways once released into the environment in large numbers (11,87). Even more inappropriate are the policies of TSCA that require the EPA to demonstrate that a product presents a likelihood of health or environmental risk before requiring the manufacturer to conduct additional testing. Since there is almost no information on how these products will react in the environment (and the manufacturer, absent a command from EPA, is not required to conduct the testing

on products not otherwise covered under FIFRA or the FDCA), the burden placed on EPA undercuts its ability to effectively encourage adequate safety research on these new products (63,88).

Finally, the unfortunate consequences of these uncertainties seem to be multiplied by the current regulatory approach that attempts to ignore or cover them up. Current incremental and ad hoc approaches do not make headway in gaining public confidence. Indeed, growing fears of the safety and acceptability of genetically engineered foods by some of the public may threaten the plant pesticide industry more than tentative and still incomplete regulatory oversight by EPA (89).

CONCLUSION

EPA's federal policy making on the products of biotechnology is diverse and complex, due at least in part to the Executive Branch's historic efforts to fend off legislative intervention. Although EPA seems to be endeavoring mightily toward the incremental creation of subprograms for different types of biotechnological pesticides and products, skepticism remains with regard to whether this administrative approach will ultimately meet the escalating needs of both the industry and the concerned public.

BIBLIOGRAPHY

1. W.H. Rodgers, Jr., *Environmental Law*, vols. 1–5, West Publishing, St. Paul, MN, 1988.
2. P. Mostow, *Pace Environ. Law Rev.* **10**, 227,240–242 (1992).
3. J.P. Dwyer, *Ecol. Law Q.* **17**, 233 (1990).
4. Streamlining Federal Regulation of Biotechnology Products, Press Release from Council on Competitiveness, February 23, 1992.
5. 1992 National Biotechnology Policy Report at E-4, reprinted in *Biotechnol. Law Rep.* **12**, 127 (1993). Several documents were produced as a result of this effort, the most recent being a 1992 Federal Oversight document that clarifies the scope of federal regulation of biotechnology. *Fed. Regist.* **57**, 6753 (1992).
6. *Fed. Regist.* **51**, 23302–23303 (1986).
7. *Fed. Regist.* **57**, 6753 (1992). For a critique of this policy, see Mostow (2).
8. W.H. Rodgers, Jr., *Environmental Law: Pesticides and Toxic Substances*, vol. 3, West Publishing, St. Paul, MN, 1988, §6.12, pp. 504–505, Chart I.
9. *Fed. Regist.* **9**, 50859–50861 (1984).
10. *Hearings Before the Subcommittee on Investigation and Oversight and the Subcommittee on Science, Research and Technology of the House Committee on Science and Technology, 98th Cong., 1st sess.*, pp. 131–132 (1983). (Statement of Dan Clay, Acting Asst. Admin., Office of Pesticides and Toxic Substances, EPA).
11. L. Maher, *J. Environ. Law Litig.* **8**, 133,195 (1993).
12. C.C. Vito, *Me. Law Rev.* **45**, 329,341–346 (1993).
13. W.H. Rodgers, Jr., *Environmental Law: Pesticides and Toxic Substances*, vol. 3, West Publishing, St. Paul, MN, 1998, §5.3, pp. 32–54.
14. 7 U.S.C. §§136, *et seq.*
15. 7 U.S.C. §§136a(c)(5)(D), 136d(b).
16. W.H. Rodgers, Jr., *Environmental Law: Pesticides and Toxic Substances*, vol. 3, West Publishing, St. Paul, MN, 1998, §5.10, pp. 127–145.
17. 40 CFR, pt. 158.
18. 7 U.S.C. §136c; 40 CFR, pt. 172.
19. 7 U.S.C. §136n.
20. 7 U.S.C. §136(u).
21. *Fed. Regist.* **59**, 45605 (1994).
22. *Fed. Regist.* **59**, 45608–45609,45611–45612 (1994).
23. *Fed. Regist.* **59**, 45606–45607,45614 (1994).
24. H.J. Miller, *Science* **266**, 1815 (1994).
25. M.J. Angelo, *U. Fla. J. Law Public Policy* **7**, 257,281 (1996).
26. R.V. Percival, A.S. Miller, C.H. Schroeder, and J.P. Leape, *Environmental Regulation: Law, Science, and Policy*, Little, Brown, New York, 1996, pp. 176–179.
27. *Fed. Regist.* **59**, 45602,45606–45607,45614 (1994).
28. *Fed. Regist.* **49**, 40660 (1984).
29. 40 CFR §172.48.
30. 40 CFR §§172.50, 172.57, 172.59.
31. 40 CFR §172.50(a).
32. 40 CFR §172.10.
33. J.R. Holmstead, *ALI-ABA* **C948**, 709,712 (1994); *Daily Environ. Rep. News (BNA)* **138**, A-7 (1997).
34. *Daily Environ. Rep. News (BNA)* **143**, A-3 (1997).
35. *Fed. Regist.* **59**, 60519 (1994); but see *Chem. Regul. Rep. (BNA)* **23**, 184 (1999).
36. *Fed. Regist.* **59**, 60496 (1994); *Fed. Regist.* **59**, 65500 (1994).
37. *Chem. Regul. Rep. (BNA)* **20**, 657 (1996) (quoting report issued by the Institute of Food Technologists).
38. J.E. Beach, *Food Drug Law J.* **53**, 181,191 (1998); *Chem. Regul. (BNA)* **22**, 1930 (1999).
39. *Fed. Regist.* **59**, 60519,60522,60525–60528 (1994).
40. *Chem. Regul. Rep. (BNA)* **21**, 711,1220 (1997).
41. *Chem. Regul. Rep. (BNA)* **23**, 857 (1999); see also *Daily Environ. Rep. News (BNA)* **161**, A-5 (1999) (reporting on research detecting adverse effects of Bt corn on monarch butterflies). EPA has also sought the assistance of its science advisors in determining appropriate pest resistence plans for particular biotechnology products. *Chem. Regul. Rep. (BNA)* **21**, 1414 (1998).
42. 40 CFR §172.3.
43. *Fed. Regist.* **59**, 60510 (1994).
44. 21 U.S.C. §§346a and 348.
45. *Daily Environ. Rep. News (BNA)* **95**, A-2 (1997).
46. *Fed. Regist.* **59**, 60535,60537,60538 (1994).
47. 15 U.S.C. §§2603(a) and 2605(a).
48. 15 U.S.C. §2604(a).
49. 15 U.S.C. §2604(a)(1).
50. 15 U.S.C. §2604(d)(1).
51. 15 U.S.C. §2603(a); D.J. Hayes and A. Claassen, *ALI-ABA* **C864**, 191,208 (1993).
52. M.L. Lyndon, *Mich. Law Rev.* **87**, 1795,1825–1827 (1989).
53. 15 U.S.C. §2604(a)(1).
54. 15 U.S.C. §§2603(a) (require testing); 15 U.S.C. 2604(e)(1)(A) (block manufacture pending testing); 15 U.S.C. 2605 (ban manufacture).
55. General Accounting Office, *Toxic Substances: EPA's Chemical Testing Program Has Made Little Progress*, U.S. Government Printing Office, Washington, DC, 1990.

56. 15 U.S.C. §2602(2)(B) (defining "chemical substance" in relation to other statutes).
57. 15 U.S.C. §2602(2).
58. *Fed. Regist.* **51**, 23324 (1986).
59. 40 CFR §725.3.
60. *Fed. Regist.* **51**, 23325–23326 (1986). EPA can also use other authorities to reach existing products that are used in significantly new ways under its "significant new use rule" (SNUR). 15 U.S.C. §2604(a)(2).
61. 40 CFR §§725.100–190.
62. 40 CFR §725.170.
63. R. Harlow, *Yale Law J.* **95**, 553,565,570–574 (1986).
64. A.G. Gorski, *Univ. Balt. J. Environ. Law* **3**(1), 24–25 (1993).
65. Jaffe, *Harv. Environ. Law Rev.* **11**, 491,540 (1987).
66. 40 CFR §§725.200–288; *Chem. Regul. Rep. (BNA)* **22**, 7 (1998).
67. *Fed. Regist.* **61**, 63143 (1996); *Daily Environ. Rep. News (BNA)* **13**, G-2 (1997).
68. *Daily Environ. Rep. News (BNA)* **206**, A-5 (1997) (quoting EPA biotech attorney, David Giamporcaro).
69. 40 CFR §§725.155–725.160; *Chem. Regul. Rep. (BNA)* **21**, 245 (1997); *Fed. Regist.* **62**, 17910 (1997).
70. *Daily Environ. Rep. News (BNA)* **169**, A-4 (1997).
71. *Daily Environ. Rep. News (BNA)* **206**, A-5 (1997).
72. *Daily Environ. Rep. News (BNA)* **187**, A-2 (1997); *Chem. Regul. Rep. (BNA)* **18**, 1749 (1995); *Chem. Regul. Rep. (BNA)* **21**, 1610 (1998).
73. 40 CFR Subparts E,F, and G.
74. 33 U.S.C. §§1251–1387 (1994).
75. D.E. Hoffmann, *Drake Law Rev.* **38**, 471,498 (1988). (Considering similar regulations under the Clean Air Act).
76. 16 U.S.C. §§1536 and 1538.
77. *Environ. Rep. (BNA)* **29**, 192 (1998).
78. *Fed. Regist.* **59**, 348,359,401 (1994); *Chem. Regul. Rep. (BNA)* **15**, 1646 (1992).
79. *Chem. Regul. Rep. (BNA)* **20**, 279 (1996); *Daily Environ. Rep. News (BNA)* **124**, d21 (1993).
80. C.R. Manor, *Rutgers Comput. Tech. Law J.* **3**, 409 (1987).
81. C.M. Steen, *Ariz. Law Rev.* **38**, 763,791 (1996).
82. *Chem. Regul. Rep. (BNA)* **18**, 1549 (1995).
83. *Daily Environ. Rep. News (BNA)* **143**, A-3 (1997).
84. S. Goldberg, *Culture Clash: Law and Science in America*, New York University Press, 1994, pp. 84–111.
M. Malinowski and M. O'Rourke, *Yale J. Regul.* **13**, 163 (1996).
85. *Chem. Regul. Rep. (BNA)* **18**, 1549 (1995).
86. W.E. Wagner, *Columbia Law Rev.* **95**, 1613 (1995).
87. *Chem. Regul. Rep. (BNA)* **2**, 566 (1997) (reporting criticisms of EPA's risk assessment protocols for genetically engineered microbes).
88. J.S. Applegate, *Columbia Law Rev.* **261**, 319–330 (1991).
R.A. Chadwick, *Hofstra Law Rev.* **24**, 223 (1995).
89. *Daily Environ. Rep. News (BNA)* **143**, A-3 (1997).

See other entries AGRICULTURAL BIOTECHNOLOGY, LAW, AND EPA REGULATION; see also FEDERAL POLICY MAKING FOR BIOTECHNOLOGY entries.

FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, DEPARTMENT OF ENERGY.

See INFORMING FEDERAL POLICY ON BIOTECHNOLOGY: EXECUTIVE BRANCH, DEPARTMENT OF ENERGY.

FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, ELSI

AUDRA WOLFE
University of Pennsylvania
Philadelphia, Pennsylvania

OUTLINE

INTRODUCTION

The Ethical, Legal, and Social Implications (ELSI) Program, part of the Human Genome Project (HGP), is a unique federal policy and research program that attempts to anticipate potential negative consequences of the very project of which it is a part. Between 3 and 5 percent of the budget for the HGP is dedicated to projects that investigate various ethical, legal, and social aspects of genomic research, genetic testing, and genetic information. The ELSI Program refers to a set of administrative and grant-making bodies in both the National Institutes of Health (NIH) and the Department of Energy (DOE) that make policy recommendations, determine the ELSI agenda, and sponsor conferences and research. As a funding agency, ELSI administers extramural and intramural contracts and grants to historians, philosophers, legal scholars, sociologists, bioethicists, and policy analysts on a wide variety of issues. As an advisory board, ELSI has managed task forces and formulated policy recommendations for topics including genetic testing, genetic discrimination in insurance and employment, and the protection of human subjects in genetic research. As the HGP moves toward the completion of human DNA sequencing, ELSI has more recently emphasized questions of human genetic variation and diversity, clinical and nonclinical integration of genetic information, multicultural interactions with genetic information, and fairness in access to genetic testing and treatments. As the first federal scientific research

program to dedicate portions of its own budget to anticipating and analyzing its potential social impact, ELSI represents a landmark in federal science and bioethics policy.

ORGANIZATIONAL STRUCTURE

Origins of ELSI

The government's commitment to funding research into ethical, legal, and social implications of the HGP began immediately with Dr. James Watson's appointment as director of the Office of Human Genome Research of the National Institutes of Health. On October 1, 1988, at the press conference announcing his appointment, Watson suggested that 3 to 5 percent of the NIH HGP funds should support work on the ethical, legal, and social implications of new knowledge about human genetics. Recognizing that these issues were not necessarily new, Watson noted that they would nonetheless be associated with the genome project and deserved serious attention within the project (1). By the end of October, a Joint NIH-DOE ELSI Working Group had been established by the Program Advisory Committee of the HGP with psychologist Nancy Wexler (former director of the NIH's Huntington's Commission) serving as chair (2). Although initiated by Watson and originally housed in the NIH, the DOE began to co-sponsor ELSI programs in December 1989 under congressional pressure.

The original Working Group core included (in addition to Wexler) Thomas Murray, bioethics; Jonathan Beckwith, molecular genetics; Robert Murray, clinical genetics; Patricia King, law and policy; Victor McKusick, human genetics; and Robert Cook-Deegen, policy. In addition to these Working Group members, the individual programs at NIH and DOE were overseen by separate administrators. Eric Juengst, with a background in philosophy and bioethics, served as chief of the NIH ELSI program from 1990 to 1994; Michael Yesley, with experience in the law, originally conducted the DOE ELSI program (2).

Although Congress had not suggested an ELSI component in the HGP, members of Congress were nevertheless concerned about the possible consequences of new genetic information. Stressing questions of commercialization, genetic discrimination, and the choices influencing decisions to screen populations, both the National Research Council and the Office of Technology Assessment discussed social and ethical implications in their preliminary studies of the HGP. In addition Thomas Murray expressed his concern during the HGP hearings before Congress in 1988. The idea to set aside part of HGP's funds for ethical, legal, and social research, however, was Watson's and was only instituted through the NIH's budget recommendations (1).

Form and Function

While the Working Group, as a subcommittee of HGP, received funds from both NIH and DOE, the ELSI Program has both historically and currently functioned through separate NIH and DOE programs. DOE research funds are provided through an ELSI Program in the Office of Health and Environmental Research. Beginning in 1990, NIH extramural grants were administered through the National Human Genome Research Initiative (NHGRI) by the ELSI Branch (later renamed the ELSI Research Program). In 1995 a NIH Office of Policy Coordination (OPC), housed in the Office of the NIH director, was created to sponsor conferences on and analysis of ELSI-type policy issues. In addition the NHGRI's Division of Intramural Research (DIR) maintains an Office of Genome Ethics and Policy Analysis to explore ethical and legal issues arising from applied genetics. While the majority of ELSI projects have been sponsored by either of the grant programs, various task forces and conferences have been supervised by the Joint NIH-DOE ELSI Working Group.

Though unofficial, a division of labor has existed in some of the issues addressed by the NIH and DOE programs. DOE ELSI projects have been much more prominent in the areas of genetic privacy, public education, and intellectual property issues. The most visible NIH ELSI projects, on the other hand, have involved genetic discrimination and genetic testing issues on diseases ranging from cystic fibrosis to breast cancer. Both DOE and NIH, however, have administered a broad range of topics spanning the issues posed by HGP. Both agencies maintain Web sites with full descriptions of their current programs and grants, publications resulting from ELSI funds, and full-text press releases, task force reports, and workshop reports (3,4).

ELSI IN ACTION: PROJECTS AND POLICY

Topics and Questions

In its initial 1990 NIH grant announcement, ELSI stressed that projects should focus on the possible impacts of disease-related genetic information. While those projects that offered policy solutions would be granted the highest priority, the Working Group expressed a commitment to the traditional methods of the social sciences and humanities, bridging perspectives from morality, ethics, the law, policy, history, and public understanding of science. ELSI grant applicants were encouraged to propose projects in nine topics areas: fairness in insurance, employment, the criminal justice system, education, adoption, and the military; psychological and societal responses to individual genetic information; privacy and confidentiality; genetic counseling, including prenatal and presymptomatic testing, testing in the absence of therapeutic options and population screening versus testing; issues of reproductive choice; medical practice, including standards of care, training, the doctor–patient relationship, and patient education; historical abuses of genetic information, especially eugenics; commercialization, including property and intellectual property rights; and philosophical issues such as the meaning of identity, definitions of health and disease, and questions of determinism and reductionism. Clearly, the working group had established a broad agenda for the ELSI experiment (5).

The Working Group prioritized its topics in two sessions in February and September 1990. Without downplaying the research-driven humanities topics included in the original grant announcement, the working group outlined four areas as high priority for policy in the first five years. Policy recommendations for clinical implementation of new

genetic tests, genetic privacy concerns, genetic discrimination in insurance and employment, and professional and public education were seen as especially pressing issues. Although the Working Group's approach to each of these priorities differed, in each area they went beyond extramural grant administration to coordinate the production of policy options.

Projects

Within a year, funding was underway for 16 projects, including several interdisciplinary conferences. By September 1991 that number had increased to 25 external grants and 10 national conferences. ELSI began with 3 percent of the HGP budget; in 1992 NIH increased its commitment to 5 percent (2). Reflecting ELSI's wide range of topics and goals, these projects ranged from specific policy measures to academic conferences designed to define what might constitute ELSI-type issues.

Cystic Fibrosis Consortium. One of ELSI's first high-profile projects implemented the policy goal of determining appropriate uses of genetic testing. At the same time that the Working Group had been formulating its goals, medical scientists had developed a test for heterozygotes (carriers) of cystic fibrosis (CF). The ELSI Working Group emerged as a voice advocating caution before widespread use of a test of not only unknown reliability but also unexplored psychosocial consequences. What did a subject need to know to use the test's results to his or her advantage? What sorts of counseling measures were necessary? Would the public at large be interested in testing for CF? To help answer some of these questions, the Working Group solicited proposals for clinical trials evaluating social as well as medical aspects of CF testing and then created a coordinated project among several applicants. The American Society for Human Genetics bolstered ELSI's power and legitimacy in CF testing by urging restraint until the results of the trials were made clear. The recognition of the importance of "client-centered criteria" was one of the most significant results of the trials. Any understanding of the "success" of genetic tests must take into account individual families' reactions to and uses of the information. Empowering the patient had become at least a nominal goal of CF trials. ELSI sponsored a similar program for evaluating genetic tests for those at risk of cancer, and these programs have set precedents for similar investigations sponsored by the Heart Institute, the National Institute of Mental Health, and the National Child Health Institute (1,2).

Genetic Testing. By 1994 the possibilities for genetic testing had grown far beyond CF. To the rising concern of the ELSI Working Group, a number of tests of variable quality had been introduced to the public without much discussion of patient reception. It was clear that a more systematic approach to the problem of genetic testing had become necessary. As a first step toward addressing this problem, the Working Group commissioned a study by the National Academy of Science's Institute of Medicine on assessing genetic risks. This report raised special concerns about the lack of treatment — preventive or otherwise — for diagnosed diseases and insufficient genetic expertise among those who developed or performed genetic tests. The Working Group established the Task Force on Genetic Testing in 1994 to further explore these issues and to formulate policy recommendations. The voting portion of the Task Force consisted of 15 members nominated by diverse genetic testing interest groups; several nonvoting liaison members from agencies such as the Department of Health and Human Services (HHS) and the Food and Drug Administration (FDA) completed the committee.

The Task Force began their ambitious agenda by surveying the quality of genetic tests, the quality of the laboratories providing the tests, and the competence of testing personnel. They commissioned a series of background papers on the state of genetic testing and clinical laboratories and conducted a survey of all organizations likely to be involved in genetic testing. As a follow-up to the survey, in-depth interviews were conducted at 29 of the almost 500 organizations either currently conducing genetic testing or considering it in the future. The Task Force met in seven different sessions; halfway through the process, they invited public comment on preliminary conclusions and recommendations. In 1997 the Task Force issued its final report and policy recommendations.

The Final Report established several overarching principles before moving on to their specific studies. Informed consent, consideration of an individual family's values in regard to prenatal or carrier testing, confidentiality, and the prevention of discrimination were considered crucial goals regardless of specific policy recommendations. Policy measures focused on questions of oversight, licensure, and expertise. The Task Force expressed concern that genetic tests did not fall under FDA supervision, so the report strongly encouraged interactions with IRBs and the Office of Protection of Human Subjects to determine appropriate guidelines for conducting clinical trials on genetic tests. Tests should be subject to external peer review to ensure their efficacy and scientific merit. The report further forcefully stated that laboratories conducting genetic tests should have Clinical Laboratory Improvement Amendments (CLIA) certification. To demonstrate expertise in medical genetics, laboratory directors or technical supervisors should have American Board of Medical Genetics certification. To improve health care providers' understanding of genetic testing and genetic disease, the Task Force strongly recommended that medical, public health, and nursing education include more human genetics instruction and suggested that the inclusion of human genetics questions on licensing exams might spur the development of medical residencies in genetics. Finally, the Task Force recommended that measures were required to ensure the continued development of tests for rare genetic diseases as well as those possibly effecting a large portion of the population (6).

To implement all of these proposals, the Task Force advised the creation of a formally chartered advisory committee on genetic testing in the Office of the Secretary of Health and Human Services. Such an office was created in June 1998 by Secretary Donna Shalala. While not all of the recommendations have been implemented, the Task

Force's Final Report has been received as an important document in determining the future of genetic test development, and the creation of the Advisory Committee represents a key step in maintaining a formal discussion on issues raised by genetic testing.

Genetic Discrimination. Especially in ELSI's early years, questions of genetic discrimination generally concerned either employment or insurance. In dealing with matters of employment, the Working Group appealed to the Americans with Disabilities Act (ADA) as a possible source of guidelines. For insurance, the Working Group created a special task force that participated in national debates on the American health care system.

The ADA ensures that employers cannot consider disabilities as a factor in hiring decisions. The Equal Employment Opportunity Commission (EEOC) interprets and enforces the ADA. The possibility of pre-symptomatic genetic testing introduced the potential for a new kind of disability categorization and subsequent discrimination. Would the EEOC offer protection to those denied employment on the basis of genetic test results that suggested that a person was at risk for a disease, or of parenting a child with a disease, in the future? Could heterozygotes for recessive illnesses be considered disabled for the purposes of the ADA?

The Working Group responded to these concerns by submitting policy recommendations to the EEOC in April 1991. While NIH and DOE themselves cannot make recommendations concerning legislation and regulations, their subcommittees can. ELSI's policy statement on ADA was therefore incorporated into a recommendation by the NIH-DOE Joint Committee on the Human Genome. This action — though not immediately effective — demonstrated that ELSI might be able to exercise policy options not available to the NIH itself.

ELSI's recommendations included three main points for strengthening ADA. First, heterozygotes for recessive and X-linked disorders should be considered "impaired" for the purposes of protection by ADA. Second, employment entrance exams administered after a job offer should either be voluntary or limited to job-related physical or mental impairments. This amendment would protect against discrimination on the basis of HIV status as well as genetic disorders. Third, EEOC should consider ways to protect employees from employer access to personal medical information unrelated to specific insurance claims (7). EEOC did not initially incorporate ELSI's recommendations into their interpretation of ADA; after several years of negotiations, however, EEOC altered their guidelines to explicitly offer protection from employment discrimination based on either genetic disease or the results of tests predicting the development of genetic disorders in the future.

In the area of health insurance, ELSI created a special Task Force on Genetic Information and Insurance consisting of two Working Group members, representatives from the health insurance industry, public-interest groups, government officials, and ELSI grantees. Operating on the assumption that the advance of genomic science would blur the distinction between genetic and nongenetic information, the task force concluded that specific recommendations based solely on genetic information offered too narrow a solution. Instead, the task force concluded that only the development of a new system of American insurance — one not calculated on individual risk underwriting — could adequately and equitably address the issues raised by genetic testing. The task force's bold recommendations included seven points. First, health status — past, current, or future — should not be used to deny health care coverage. Second, access to the health care system should be expanded so that all Americans would have access to basic medical care. Third, genetic services, including counseling, testing, and treatment, should be included as basic aspects of medical care. Fourth, individual insurance costs should not be based on information concerning health status — past, present, or future. Fifth and sixth, access to health care or insurance programs should neither be based on employment nor disclosure of genetic or other medical information. Finally, additional steps should be taken to protect individuals until the time that universal health coverage should arrive (8).

The preliminary ELSI recommendations were passed along to the White House Task Force on Health Care Reform in 1992 and were included in the Clinton health care reform bill, the Health Care Security Act of 1993. Having made its recommendations, the Task Force then dissolved, and their recommendations were defeated along with the rest of the Clinton health care plan (1). While ELSI measures to protect individuals from genetic discrimination in insurance largely failed, the Health Insurance Portability and Accountability Act of 1996 now provides some measure of protection by excluding genetic tests, in the absence of disease, from preexisting condition clauses. In other words, a patient does not have a preexisting condition until he or she develops a genetic disease regardless of the indications of a genetic test (9).

Genetic Privacy. ELSI's approach to genetic privacy offered legislation instead of task force recommendations. The DOE's ELSI program commissioned several studies to investigate various philosophical, sociological, and legal aspects of genetic privacy. Concerns ranged from the protection of medical records to an individual subject's control over the fate of his or her own genetic samples. Under the direction of this program, George J. Annas, Leonard H. Glantz, and Patricia A. Roche (collaborating at the Health Law Department of the Boston University School of Public Health) drafted ELSI's first legislative document (9). Their "Genetic Privacy Act" was released in 1994. The central tenet of the legislation was that because of both its predictive value and potential implications for other family members, genetic information warrants a higher level of protection than other medical information.

Though a complex bill, the Genetic Privacy Act stressed two issues. First, under no circumstances, except legal order or forensic investigation, should genetic analysis be conducted without the express written consent of individuals (special provisions were included to deal with pregnant women, embryos, minors, and incompetent

persons). A subject should be informed of the nature of testing or research being conducted, and medical personnel and research scientists should abide by any restrictions a subject might make. Second, the act restricted research access to unidentifiable DNA samples. Any sample that could be traced to an individual patient through a name, address, social security number, health insurance number, or any other identifying information should be excluded from genetic research. The Act defined DNA samples very broadly, including banked blood, tissue samples, cell lines, and saliva as well as samples taken expressly for genetic analysis. Their guidelines placed the dignity of the research subject at the center of biomedical ethics — if a subject refused to permit his or her DNA to be used for research or commercial purposes, researchers would be required to respect his or her wishes regardless of the scientific importance of the sample (10). Not surprisingly, some members of the research community and the pharmaceutical industry have responded to the proposed Act with alarm, charging that such measures impede the process of medical research. Although Oregon passed a bill modeled on the Genetic Privacy Act in 1995, New Jersey Governor Christine Whitman vetoed a similar measure in 1996 under pressure from the pharmaceutical industry. Debates about the merits and drawbacks of such a bill continue at both the state and national level (9).

Public Education. Though ELSI placed high priority on increasing public and professional literacy about genetics, genetic testing, and HGP, the Working Group itself has sponsored few public forums on this matter. The DOE ELSI program, on the other hand, has been quite active in creating educational materials for the general public, teachers and schoolchildren, and professionals who encounter genetic information. Though not directly a part of the ELSI program, DOE's Human Genome Management Information System (HGMIS) has also been essential in disseminating information about HGP to the public through press releases, published documents, and an extensive website.

DOE ELSI education projects vary widely in form, content and audience. General public education programs have included exhibits at the American Museum of Natural History and the San Francisco Exploratorium, a radio program (*The DNA Files*) broadcast over National Public Radio, video documentaries such as *A Question of Genes* and *The Secret of Life* for public television and classroom use, forums to inform minority groups about HGP and to spark discussions about its meanings, and Spanish-language radio programming on HGP. To reach a student audience, DOE ELSI has sponsored training programs in genomic techniques for high school and community college teachers, and a pilot program in Seattle, Washington, allows students to perform DNA sequencing experiments in their classrooms. Workshops, conferences, and seminars have also been conducted for medical professionals and judges who increasingly interact with genetic evidence in the courtroom. Another important program at Cold Spring Harbor educates policy makers, legislators, journalists, and other opinion makers on HGP and human genetics in an attempt to influence "downstream" popular perceptions of HGP (11).

As is clear from this partial list of programs, much ELSI education has focused on increasing scientific literacy. While parties on all sides of HGP tend to agree that education is theoretically important, the question of what constitutes education has sparked much controversy. Will Seattle students who understand how to sequence DNA be more capable of dealing with the social implications of HGP through their exposure to scientific techniques? For the architects of the DOE ELSI program, improving scientific literacy is the first step in preparing the public to deal with genetic issues and is an important part of dispelling damaging misconceptions of the goals and methods of HPG. From another perspective, however, these sorts of educational programs divert attention from the real ELSI concerns arising from the genome project.

ELSI'S CRITICS AND FEDERAL SCIENCE POLICY

Policy or Research? Restructuring ELSI

Although ELSI was not the first federally sponsored bioethics board, it did mark a unique path in science policy. As both its critics and supporters have noted, HGP was the first federal science project to fund self-criticism as well scientific research. ELSI's grant structure, however, was more suited to stimulate discussion across a wide spectrum of scholars than to implement policy decisions. Over the past 10 years, the purpose of ELSI as either a policy board or a research enterprise has been a matter of contention.

As discussed in the previous section, the ELSI Working Group actively participated in policy discussions about genetic testing, privacy, discrimination, and education. To its critics, however, the Working Group's extramural grant program appeared to lack coordination and seemed to bear little relation to actual policy concerns. ELSI's most visible early products, such as the Genetic Privacy Act or the Insurance Task Force Report, were spearheaded by individual grant recipients or particular members of the Working Group rather than the Working Group as a whole. In 1992 the Committee on Governmental Operations issued a report stating that ELSI lacked mechanisms for issuing policy measures and recommended the formation of a formal advisory commission to address policy issues (12). Under further criticisms that ELSI projects tended to address "contextual" issues surrounding the HGP rather than specific issues stemming from genomic work, the new director of NGHRI, Francis Collins, strongly suggested that ELSI focus its work on instrumental projects. In 1996, the Working Group came under review by an 11-member Joint NIH-DOE ELSI Evaluation Committee to determine its future role in the HGP.

The committee of legal scholars, scientists, health care professionals, and one science studies scholar released their final report on the future of ELSI in December 1996. The Evaluation Committee pointed to the increasing pace of genomics research, the turn toward applied genetics, and the increasing integration of genetic research

into medical practice as indications of a crucial need for effective policy mechanisms on ELSI issues. Although the report contained several specific criticisms of the effectiveness of the ELSI Working Group, the crux of the Evaluation Committee's criticisms suggested that ELSI's mix of research programs and policy development was incompatible with efficient policy formulation. While stressing the importance of an independent ELSI grant program for public acceptance of HGP, the committee concluded that the ELSI Working Group, as it was then configured, occupied too Byzantine a niche and pursued too many concurrent goals to ever be effective. To resolve these issues, the Evaluation Committee recommended three changes in the structure of ELSI programs: (*1*) the redesign of the ELSI Working Group as the ELSI Research Evaluation Committee, which would coordinate and follow-up on ELSI grants and prioritize research agendas; (*2*) the creation of a NIH-wide policy office, housed in the Office of the Director, to formulate policy issues and monitor compliance with NIH guidelines on genetic research; (*3*) and the establishment of a federally chartered Advisory Committee on Genetics and Public Policy to be housed in the Office of the Secretary of HHS (13).

While the recommendations of the Evaluation Committee were not implemented exactly as outlined in their report, its assessment of ELSI as an impotent program has had repercussions in ELSI's form and function. The tasks of the Working Group, the grants programs, and additional ELSI programs within NIH and HHS have been realigned to give stronger voices to scientists and policy experts. NIH and DOE now collaborate through an ELSI Research Planning and Evaluation Group (ERPEG) similar to that suggested by the Evaluation Committee. NIH's OPC performs many of the NIH-wide functions advocated by the Evaluation Committee. Finally, the Secretary's Advisory Committee on Genetic Testing, while perhaps addressing a more narrow set of policy questions than those envisioned by the Evaluation Committee, has been established as a high-profile policy entity. As conceived by both the Evaluation Committee and the Task Force on Genetic Testing, this Advisory Committee interacts with other federal agencies such as FDA, the Environmental Protection Agency (EPA), the Department of Justice, and the National Bioethics Advisory Committee when addressing issues of broad public and policy interest. As an Advisory Committee on Genetic Testing, however, this HHS committee cannot be expected to issue policy on the complete spectrum of ELSI issues.

A New Role for Government?

ELSI was not the federal government's first foray into biomedical ethics. Its novelty lie instead in its source of funding in the very project it was meant to critique. Over the past 30 years, however, various government agencies have acted as ethical and social commentators on scientific research. Housed in the Department of Health, Education, and Welfare, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1974–1978) participated in debates over fetal tissue research, set guidelines for experiments involving human subjects, and established the guiding questions of the discipline of bioethics in its landmark Belmont Report (1978). The National Commission was followed by the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (1980–1983), which addressed concerns on genetic screening, genetic counseling, and human gene therapy as well as other issues in bioethics. In 1984 this President's Commission was replaced with the less successful Congressional Biomedical Ethics Board and its expert arm, the congressionally appointed Biomedical Ethics Advisory Committee (BEAC, 1988–1989). While the BEAC attempted to address issues of genetic screening and genetic discrimination in insurance, its potential as a force in federal biomedical ethics policy was lost in a political battle over abortion in 1989 (2). More recently the President established a National Bioethics Advisory Committee (NBAC) with a broad mandate to issue policy recommendations on any bioethical issues arising from research (or clinical applications of such research) on human biology and behavior. When established in 1995, the highest priorities of the NBAC were the protection of the rights of human research subjects and questions of the uses of genetic information. The NBAC, then, offers another policy arena in which ELSI-type issues enter public discussion without the involvement of actual ELSI programs.

While federally chartered bioethics commissions oversee biomedical research from outside the scientific research establishment, a self-critical federal science program could be created in numerous ways. In one configuration, scientists regulate their own work to protect the public from potential harm. The NIH guidelines for working with recombinant DNA provide an example of this sort of program. Responding to public concerns about the safety of genetic engineering techniques, scientists created technical guidelines for permissible levels of risk. Individual proposals for research surpassing a certain level of risk require approval from a scientist-run NIH committee. Many of the scientists involved in HGP had experienced this sort of scientific policing either through reviews of their own recombinant work or participation on a review committee. In a different sort of self-critical science, uniquely represented by ELSI, scientific agencies provide funds for the public discussions of medical ethics and social concerns resulting from scientific advances. While the first evaluates individual research proposals in light of current research guidelines determined by scientists, the second type may allow scientific outsiders to determine what those guidelines should be in the first place. By focusing on the ethical, legal, and social *implications* of genome work, however, the ELSI program has largely reacted to consequences of genome work rather than formulation of research guidelines.

Ethics and the Role of History

From its inception, ELSI has garnered serious criticism. The original NIH guidelines for ELSI programs implied that ELSI would fund projects criticizing approaches and consequences of the HGP. Some scientists questioned

whether this was a wise use of HGP funds. In addition, the provision of funding for an ELSI program within the HGP — and not other federal research projects — necessarily implied that the HGP, by definition, required a higher level of public scrutiny than most federal science projects. Those scientists who saw such fears as unwarranted were especially alarmed at the attention ELSI might bring to the HGP. Many critics of science, on the other hand, wondered whether those receiving ELSI funds would have the intellectual freedom to criticize the HGP. In other words, the very idea of ELSI seemed to represent a conflict of interest. Another criticism was whether introducing such relatively large sums of money into the pool of grants available to historians, sociologists, science policy analysts, and bioethicists might unduly skew research topics away from other pressing issues in contemporary biomedicine. Finally, it has been suggested that ELSI serves nothing better than the interests of science. By encouraging limited criticism, scientists at the HGP are freed from the claim that they have ignored public concerns over the uses of genetic technologies such as those publicly voiced during the recombinant DNA controversy of the 1970s.

The strongest criticism has been a difficult one for ELSI to answer. By focusing on the *implications* of the HGP, ELSI leaves little room for discussions of *whether* the HGP should proceed. Questions of resource allocation in federal science research funding, for example, are not appropriate uses of ELSI funds as originally defined. Furthermore this issue highlights the potential for conflict of interest. Is it realistic to expect scholars working with ELSI funds to recommend a reduction in the HGP's budget? The history of ELSI thus far suggests that scholars working with ELSI funds have been able to maintain a reasonable amount of academic freedom. The extramural grant structure of ELSI projects offers a first level of protection since projects are evaluated by peer review instead of HGP (though now reviewed as well by ERPEG), and the very nature of ELSI projects in exploring the social, cultural, and legal context of HGP has encouraged researchers to question HGP itself. Historians and philosophers have been especially interested in the elements of reductionism inherent in HGP and exploring the limitations of such a methodological approach (1).

Underlying most of these criticisms is a question of the singularity of the problems caused by HGP. Regardless of the medical possibilities or "therapeutic gaps" presented by genetic testing technologies, cultural notions associated with genetics do present special concerns. Justifiably or not, Americans have attributed special powers to genetics — over education, socioeconomic status, or psychology — in controlling individual destiny. This attitude was most clearly represented, of course, in the involuntary sterilization eugenic campaigns of the early twentieth century. XXY screening programs represent a more recent example. The historical precedent for misuses of genetic information not only suggests HGP warrants an extra level of caution above that granted other medical technologies but also imparts a particular importance to understanding the past. Discussions of "past abuses of genetics" almost always arise in discussing the function of HGP. In this context the various ELSI programs have been essential in securing public support for HGP and will remain so throughout the completion of the project. ELSI-type issues, however, will not disappear with the completion of genomic sequencing. It remains to be seen how concerns over the ethical, legal, and social implications of HGP will be handled after the project itself ends and leaves only its "implications."

BIBLIOGRAPHY

1. E.T. Juengst, *Soc. Policy Philos.* **13**(2), 63–95 (1996). Available at NIH ELSI Homepage: *http://www.nhgir.nih.gov/ELSI/*
2. R. Cook-Deegan, *The Gene Wars: Science, Politics, and the Human Genome*, Norton, New York, 1994.
3. DOE ELSI Homepage: *http://www.ornl.gov/hgmis/resource.elsi.html* Maintained by the Human Genome Management Information System. Accessed 12/13/99.
4. NIH ELSI Homepage: *http://www.nhrgi.nih.gov/ELSI/* Maintained by the National Human Genome Research Initiative. Accessed 12/13/99.
5. National Center for Human Genome Research, *NIH Guide Contracts Grants* **19**(4), 12–15 (1990).
6. N. Holtzman and M.S. Watson, eds., *Promoting Safe and Effective Genetic Testing in the United States: Final Report of the Task Force on Genetic Testing*, 1997. Available at: *http://www.nhgri.nih.gov/elsi/tfgt_final* Site maintained by the NHGRI, accessed 12/14/99.
7. M.S. Yesley, *Hum. Genome News* **3**(3), 12–14 (1991).
8. NIH-DOE Working Group on Ethical, Legal, and Social Implications of Human Genome Research, *Genetic Information and Health Insurance: Report of the Task Force on Genetic Information and Insurance* NIH-DOE, Washington, DC, 1993. Available at: *http://www.nhgri.nih.gov:80/About_NHGRI/Der/Elsi/itf.html* Site maintained by the NHGRI. Accessed December 14, 1999.
9. P.R. Reilly, *Hum. Genome News* **8**(3 and 4), 1–3 (1997).
10. G.J. Annas, L.H. Glantz, and P.A. Roche, *The Genetic Privacy Act and Commentary*, 1995. Available at: *http://www.ornl.gov/hgmis/resource/privacy/privacy1.html* Site maintained by the Human Genome Management Information System, Accessed 12/13/99.
11. D. Drell and A. Adamson, *DOE ELSI Program Emphasizes Education, Privacy: A Retrospective, 1900–1999*, 1999. Available at: *http://www.ornl.gov/hgmis/resource/elsiprog.html* Site maintained by the Human Genome Management Information System, Accessed 12/13/99.
12. Committee on Governmental Operations, House of Representatives, U.S. Congress, *Designing Genetic Information Policy: The Need for an Independent Policy Review of the Ethical, Legal, and Social Implications of the Human Genome Project*, U.S. Government Printing Office, Washington, DC, 1992.
13. Committee to Evaluate the Ethical, Legal, and Social Implications Program of the Human Genome Project, *Report of the Joint NIH/DOE Committee to Evaluate the Ethical, Legal, and Social Implications Program of the Human Genome Project*, 1996. Available at: *http://www.nhgri.nih.gov:80/Policy_and_public_affairs/Elsi/elsi_recs.html*

See other entries EDUCATION AND TRAINING, PUBLIC EDUCATION ABOUT GENETIC TECHNOLOGY; see also FEDERAL POLICY MAKING FOR BIOTECHNOLOGY entries.

FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, NATIONAL BIOETHICS ADVISORY COMMISSION

HAROLD T. SHAPIRO
Princeton University
Princeton, New Jersey

OUTLINE

INTRODUCTION

The establishment of national bioethics commissions with rather broad responsibilities to advise the federal government regarding public policy matters in the bioethical arena, possibly by recommending new regulations and/or legislation, is, like the discipline of bioethics itself, a relatively new phenomena. In the United States such commissions have been established for a wide variety of complex reasons, under the auspices of various federal entities, within somewhat different venues and with more or less enthusiasm, resources and authority. Overwhelmingly, however, these commissions have been only advisory in nature. Moreover these commissions, unlike their counterparts abroad, have not been created as standing bodies but have been appointed for relatively short terms, presumably at those moments in time where the federal government felt the need for such assistance. In any case, these commissions have been appointed principally to advise the federal government on its ethical responsibilities within biomedical science and clinical medicine, especially in morally controversial areas where the imperatives of advances in the biomedical sciences and associated clinical applications stand in some tension to what some may feel are important ethical obligations, or where our evolving moral sensibilities require a change in traditional practices. In this respect they are, very broadly speaking, expected to clarify the relationship between the new opportunities that biomedical science research practices and clinical applications continue to offer to human welfare and the limitations that might result on these activities because of other important ethical obligations.

America is a morally pluralistic society (i.e., a society dedicated to incorporating a diverse set of ideas regarding those values that will allow for the greatest human flourishing). As such, there is a special need for some mechanism to articulate a national consensus in bioethical matters where that is possible, or at least to delineate the points of disagreement among us and the common ground (if any) where mutual empathy and understanding are the most that can be expected. This is especially so given the rapid advances in biomedical science that inevitably produce new ethical questions. At times these commissions may have additional roles such as identifying emerging issues, defusing controversy, delaying action, or giving the illusion of acting, endorsing a decision that is already made, reviewing the effectiveness of existing rules, giving the impression of open-mindedness on the part of policy makers, and educating the public and professionals on bioethical issues. Many critics have pointed out that the appointment of broad-based commissions with their inevitably process-orientated approach that often focuses on mid-level principles is not an obvious way to resolve controversial moral issues. Indeed, some claim that such a deeply sociopolitical process is unlikely to generate morally adequate recommendations, since commissions rely a great deal on a certain level of consensus in generating their recommendations.

On the other hand, the experience of recent decades provides strong evidence that under the right conditions such commissions not only can become important agents of change in long-established practices, but they also can enhance public discourse on bioethical issues. While commissions are unlikely to generate new moral theories, they can be effective in bringing clarity to an area of moral confusion and/or controversy, providing a forum for ideas about the appropriate use of new biomedical technologies and in finding sufficient common ground to foster consensus and empathetic understanding among groups with different moral perspectives. That is to say, commissions can serve to identify emerging issues, defuse controversy, and monitor on-going compliance with existing regulations and professional standards.

Advances in science and medicine often call for difficult choices, in which benefits must be weighed against risks or in which individual needs may conflict with social norms or laws, or decisions must be made under conditions of great uncertainty. These dilemmas surround us, sometimes engulfing individuals and families, and often entering the political arena. Examples include the debate over doctor-assisted suicide, the use of fetal tissue for therapeutic transplantation, protection of human subjects in medical research, advances in medically assisted reproduction, and the trend to cut the costs of medical care through rationing of resources. As a result often a new scientific advance simultaneously is considered both a blessing and a challenge to certain existing moral commitments. Underlying the debate is a struggle to understand better how to define the moral status of human life at its various

steps of development, the respect owed to human life, and the relative importance of an individual life when weighed against the importance of many lives.

These issues are at the heart of the burgeoning field of bioethics, once the sole province of philosophers and theologians. Increasingly, however, American society finds itself turning to its government for resolution of thorny ethical issues generated by advances in biomedical science and our evolving moral sensibilities, perhaps an inevitable consequence of living at a time of startling scientific advances and in a pluralistic society where no single religious or other moral authority dominates.

As already noted, the creation of federal commissions and the use of federal funds to deliberate on the nature of the ethical constraints under which biomedical science should proceed arises from the need for some public mechanisms to articulate common values and foster consensus in the face of growing cultural and religious heterogeneity and rapidly advancing science. The establishment of these deliberative bodies also signals the increasing importance of medical and biological technologies in national life and the pressing need for reasonable groups of diverse individuals to consider these issues in a public forum. The work of some of these commissions has been widely used in courts to decide cases, in federal and state legislatures to devise statutes, in the formation of professional standards, and as intellectual and policy landmarks.

HISTORICAL BACKGROUND

Over the past quarter century, government forays into bioethics have had significant impacts on the conduct of biomedical research and the delivery of health care. In the United States, four major bioethics bodies have been established by Congress since 1968: the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (National Commission), the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (President's Commission), the Biomedical Ethics Advisory Committee (BEAC), and the National Bioethics Advisory Commission (NBAC). A fifth federal initiative, the Ethics Advisory Board (EAB) was created in response to a recommendation of the National Commission.

Other public ethics bodies have been ad hoc advisory panels of the Department of Health and Human Services. The Human Fetal Tissue Transplantation Research Panel was formed in 1988 to consider the ethical issues in the potential use of human fetal tissue for the treatment of diseases such as Parkinson's disease. In 1994 the National Institutes of Health Human Embryo Research Panel was formed to consider various areas of research involving the ex utero preimplantation human embryo and to provide advice as to those areas of research that are acceptable for federal funding, warrant additional review, and are unacceptable for federal support.

Over the same period of time, bioethics has been addressed sporadically by the Institute of Medicine, National Academy of Sciences, and the now defunct congressional Office of Technology Assessment. In addition several states have launched bioethics commissions to advise state executives and legislatures on a range of concerns, including surrogate parenting, determination of death, the use of advance directives in medical care, and organ transplantation (1).

More recently, the President established in 1995 NBAC, the first standing committee charged to address bioethical issues since 1983. We consider the work of each of these.

THE NATIONAL COMMISSION FOR THE PROTECTION OF HUMAN SUBJECTS OF BIOMEDICAL AND BEHAVIORAL RESEARCH (THE NATIONAL COMMISSION)

In the 1960s a series of scandals involving research with human subjects (i.e., the inappropriate treatment of human subjects) signaled to Congress that some biomedical and behavioral scientists were not adequately self-policing themselves and that some sort of independent oversight was necessary. Most notable among these scandals were the Tuskegee syphilis trials, the Willowbrook hepatitis experiments, the use of prisoners to test drugs, and whole body radiation experiments sponsored by the Department of Defense. Although the scientific community was initially resistant to a formal system of oversight, fearing that it might unnecessarily stymie important research, the powerful message of events like Tuskegee convinced most investigators that important research might actually be stalled for lack of public support if adequate protections were not in place.

The National Commission was established through Title II of the National Research Act of 1974 (Public Law 93-348). The bill, sponsored by Senator Edward Kennedy, reflected congressional concern about reported abuse of human research participants and the moral status of research using biological materials from aborted fetuses.

Congress created the National Commission as part of the Department of Health and Human Services (DHHS), then Department of Health, Education, and Welfare (DHEW). Eleven members were appointed by the Secretary of DHEW: five scientists, three lawyers, two ethicists, and one person in public affairs. In establishing the National Commission, Congress gave it the task of articulating the ethical principles of protecting human subjects in research. It was instructed to then employ those principles to recommend actions by the federal government. Congress also asked the National Commission to address fetal research, which it did in the four-month period allotted.

The Commission issued its first report, "Research on the Fetus," in May 1975 (2). By late 1975 the National Commission's recommendations had been translated into regulations. This first report presaged many more reports that laid a formal foundation for human subject protection in the United States. Much of the ensuing work of the National Commission focused on broader issues in human subjects protection. In essence, the National Commission codified existing DHEW policies on research with human subjects into formal federal regulations and established the requirement of Institutional Review Board (IRB) review and approval of all federally funded research involving human subjects. The resulting regulations,

45 CFR, established IRB review procedures and detailed the required elements of informed consent. In 1978, based on the National Commission's recommendations, DHEW revised its human subjects regulations and added regulations covering pregnant women, fetuses, in vitro fertilization, and prisoners. In its work on the protection of human subjects in research, the National Commission articulated three basic principles: respect for persons, beneficence, and justice. It laid great emphasis on autonomy, elaborated and extended the notion of informed consent, and recognized the special vulnerability of specific populations (e.g., children, prisoners, those institutionalized as mentally infirm).

The National Commission operated from 1974 until 1978, issuing 10 reports (2–11). As noted, many of these reports were translated, often directly, into the now-familiar federal regulations for research involving human subjects (45 CFR 46). Other aspects of its work, for example on psychosurgery, were largely ignored, and its recommendations regarding research on the "institutionally mentally infirm" were never implemented.

One of the last acts of the National Commission was its recommendation that there be a broad-based on-going federal entity to review controversial areas of research, the Ethics Advisory Board or EAB. DHEW incorporated this recommendation into its regulatory framework (45 CFR 46.204) and established the EAB (1). The National Commission also recommended that a successor body be created, but with broader authority to address issues beyond protection of human participants in research. The time was ripe for such a recommendation, since the nation was again confronted with new ethical issues created by the accelerated development of new biological and medical technologies. At the same time, issues regarding the safety of recombinant DNA were being debated on Capitol Hill, and termination of treatment was rapidly becoming a national issue in the wake of the Karen Ann Quinlan case and other court challenges to medical authority. Finding the work of the National Commission generally useful, Congress concurred that a more general mandate for a national bioethics organization was in order and created the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (the President's Commission).

THE ETHICS ADVISORY BOARD

Following the recommendation of the National Commission, the Ethics Advisory Board (EAB) was established in 1978 as an 11-member board that included lawyers, a theologian, a philosopher, clinicians, researchers, and a member of the public. It operated for two years. Although federal regulations define EAB's purview as research involving the fetus, pregnant women, and in vitro fertilization, the EAB charter grants it a broader role. It was intended as an ongoing, standing body charged to review specific proposed protocols or controversial areas of research.

EAB had some of its marching orders from day one. In 1975 DHEW had announced it would fund no proposal for research on human embryos or on the external fertilization of human eggs unless it was reviewed and approved by an independent ethics advisory board. Specifically, federal regulation required an EAB review prior to funding any research on human in vitro fertilization (45 CFR 46.204d). In vitro fertilization was the first topic addressed by EAB and its 1979 report stipulated several criteria for approval of such experiments (12). However, DHHS never implemented any of the general policy recommendations of the EAB, and DHHS disbanded the EAB in 1980 at the direction of the White House's Office of Science and Technology Policy (OSTP). As a result the Board never approved a single proposal before its dissolution, and the moratorium on human in vitro research remained in effect until lifted by President Clinton in 1993. Before it closed its doors, the EAB accomplished a few other tasks. It recommended granting a waiver to permit fetoscopy to diagnose hemoglobinopathies and handled several issues related to Freedom of Information Act inquiries in DHEW.

The budget for the EAB was diverted to the President's Commission in 1979. Since that time, few efforts have been made to reestablish the Board (1). In fact, in 1988 DHHS proposed reestablishment of the EAB and published a proposed charter for a new EAB (53 FR 35232), which expanded membership to 21 individuals. The revised charter was never signed by President Ronald Reagan, and no efforts have been made to revitalize the EAB despite recommendations to do so by various bioethics organizations.

THE PRESIDENT'S COMMISSION

The President's Commission was established by Section III of Public Law 95-622. The enabling legislation specified several tasks, as it had for the National Commission, but it also gave the President's Commission the authority to undertake studies at the request of the President or upon its own initiative. The new Commission was elevated to independent presidential status, in comparison with the National Commission which had been housed within DHHS. The range of the President's Commission work was broadened to encompass activities of the entire federal government and was extended beyond human subjects research to include medical practice.

The eleven members of the President's Commission, drawn from specific areas of expertise by law, were sworn in by President Jimmy Carter in January 1980. It operated until March 1983 and issued eleven reports including a summary of its work (13–23). As a matter of explicit policy, the Commission made few specific recommendations, instead producing consensus reports (24,25).

A focus of much of the President's Commission work concerned protection of human subjects in research, including health research regulations and compensation for research injuries (14,15,20). In 1980 the President's Commission began its congressionally mandated investigation into the adequacy and uniformity of the federal laws governing human subjects research, a topic on which it was supposed to report every two years. In its first biennial report, Protecting Human Subjects (14), the Commission

reported that it was "satisfied that the basic regulations of the Department (DHHS) were adequate if not above improvement." Its first recommendation was that all federal agencies adopt the HHS regulations set forth in 45 CFR 46. The Commission then focused its attention on determining uniformity among federal agencies as measured by the extent to which their rules confirmed the basic regulations of DHHS. Its second biennial report on this issue, Implementing Human Research Regulations (20), recommended that a program of routine site visits to Institutional Review Boards (IRBs) be implemented by relevant federal agencies on a coordinated basis and that all relevant agencies keep a record of the IRBs subject to their jurisdiction. The aim of the report was to increase the adequacy and uniformity of the implementation of existing regulations. In 1991, 16 federal agencies and departments eventually adopted a single set of regulations, known as the "Common Rule."

Its report Defining Death (13) became the foundation for statutory changes adopted in many states. This report helped to formulate and explain the Uniform Determination of Death Act. The Commission also confronted controversies about termination of treatment at the end of life in its reports on making health care decisions (16), and more specifically in deciding to forgo life-sustaining treatment (19). This report, undertaken at the Commission's own initiative, addressed highly contentious issues that continue to be debated in the courts and some legislatures, particularly in Oregon. The Commission also directly confronted the arguments for and against the use of life-sustaining treatments (26). The Commission encouraged patient- and family-centered decision making, and made recommendations regarding appointment of surrogate decision makers. It also concluded that nutrition and hydration were not fundamentally different from other medical treatments, a source of great controversy in several nationally prominent cases (e.g., Karen Ann Quinlan, Nancy Cruzan). This conclusion immersed the Commission in a controversy that led some Senate conservatives to argue that even federal bioethics committees could not be trusted on matters of great social import (26).

The President's Commission also took on the difficult issues of equitable access to health care. Its report, Securing Access to Health Care (22), was the only report that drew a dissenting vote from a commissioner. For a variety of reasons it has been more highly criticized than other Commission reports, perhaps because of the complexity and intensity of the issue, and because of its importance to broader sets of interests (27–30).

The President's Commission also addressed issues not yet fully debated at the national level concerning applications of human genetics research. It issued reports on genetic screening and counseling (21) and on human gene therapy (17). The report "Splicing Life" emphasized the distinction between genetically altering somatic cells, which would not lead to inherited changes, and germ cells (sperm, egg cells, and their precursors), which would induce inherited changes. This distinction illustrated that there could be some forms of gene therapy that would not be morally different from other forms of treatment. The Commission suggested that there be in place policies of research protocol review and ongoing means to evaluate new developments in this area of research. The Commission recommended that the National Institutes of Health review gene therapy through its Recombinant DNA Advisory Committee (RAC). In response, a RAC working group on human gene therapy was established and drafted the "Points to Consider in the Design and Submission of Human Somatic Cell Gene Therapy Protocols" (31,32). In the end, the report served to broaden the discussion about a controversial area of research and kept open the possibility for some forms of gene therapy to be considered in a political climate that was quickly moving toward unnecessarily restrictive legislation.

The Commission's term expired at the end of 1982. Extension of its term was debated in the Senate where conservative interests argued that bioethics should be brought under more direct congressional scrutiny. Ironically, a conclusion of "Splicing Life" would lay the groundwork for the establishment of two subsequent federal efforts, the BEAC and later NBAC. The report noted that there was a need for public debate, which could be mediated by an ad hoc commission on genetics or by a standing federal bioethics commission. Then-Congressman Albert Gore subsequently introduced legislation to create a President's Commission on Human Genetic Engineering, favoring permanent oversight of advances in human genetics and reproduction. This became the seed for legislation to create BEAC with a broader mandate than human genetics, as then-Congressman Gore became convinced that a broader mandate would be more useful (33). During this same period there were several alternate congressional proposals including initiatives to extend the life of the President's Commission, to give the Institute of Medicine (IOM) a mandate to do studies in bioethics, and to have the congressional Office of Technology Assessment (OTA) also do such work (26).

THE BIOMEDICAL ETHICS ADVISORY COMMITTEE

Congress took bioethics into its own hands in 1985 when it passed the Health Research Extension Act (Public Law 99-158), despite President Ronald Reagan's veto of the measure. The BEAC was a 14-member group whose multidisciplinary membership was appointed by the Biomedical Ethics Board (BEB), comprised of 12 Members of Congress, three each from the majority and minority parties of the House and Senate. BEAC was to be directly responsible for carrying out the studies of topics in biomedical ethics mandated by legislation or specified by the congressional Board. It took almost a year for the party leaders of the House and Senate to appoint the 12 members of the congressional Board, which then took on the responsibility of appointing the 14 members of BEAC, the operational arm. The appointment process took nearly two and a half years and resulted in deepened mistrust among members of the congressional Board, particularly around issues concerning abortion.

BEAC was required to prepare at least three reports on specified topics, as well as to provide annual reports. The first mandated report, on implications of human

genetic engineering, stemmed from Representative Gore's bill proposing an extension of the President's Commission (H.R. 98-2788). The deadline for the second report, on fetal research, expired before BEAC members were appointed. The fetal research mandate was reinstated in the Omnibus Health Extension Act of 1988 (Public Law 100-607) with the deadline delayed until November 1990. The third mandate stemmed from Senator William Armstrong's proposed amendment to the 1988 AIDS bill. BEAC finally met in September 1988, less than a week before its authorization expired (33). Congressional haggling over appointment to BEAC of anti-abortion members essentially closed BEAC down before it ever began its work; the office was closed at the end of September 1989.

THE NIH HUMAN FETAL TISSUE TRANSPLANTATION RESEARCH PANEL

Another effort to address bioethics at the national level, this time on an ad hoc basis, began in May 1988, when Health and Human Services Assistant Secretary Windom initiated a moratorium on the use of fetal tissue in transplantation research funded by the federal government. Fetal tissue had long been used in research, and the National Commission had previously developed the guidelines that were incorporated into federal regulations for use of fetuses in research (45 CFR 46). When scientists proposed using fetal tissue for neural grafting as an experimental treatment for Parkinson's disease, questions were raised as to whether the existing guidelines adequately covered therapeutic intent.

Lacking an Ethics Advisory Board, DHHS directed the National Institutes of Health (NIH) to convene a panel to advise DHHS about the ethical implications of such research, specifically whether the moral issues surrounding the source of such tissue (elective abortions) could ethically be separated from the therapeutic use to which such tissue is put. The issues proved to be complex and divisive (34). The panel heard testimony from disease groups, researchers, and those opposed to the research and voted on a set of specific recommendations. A majority were in favor of permitting such research as long as three conditions were met in addition to IRB approval, namely (*1*) the decision to donate tissue was kept separate from and made only after the decision to abort, (*2*) the process for abortion was not altered in any way, and (*3*) the informed consent of both parents was obtained in cases when the fathers could be contacted. The majority of the panel argued that they did not have to directly engage in questions about the morality of abortion, since the practice was legal (34). Two separate dissenting statements argued that the abortion issue should not be sidestepped. The report was approved by the Advisory Committee to the Director, NIH, urging acceptance of the recommendations (35). No action was ever taken on the panel's report.

THE NIH HUMAN EMBRYO RESEARCH PANEL

In 1994 research use of fetal tissues would again be the focus of bioethics debates, this time ex utero preimplantation human embryos produced by in vitro fertilization or other sources. Because there was no EAB, NIH was sitting on several protocols that could not be funded because they involved the research use of early-stage human embryos. The status of these protocols, in terms of need for EAB review, was unclear because they concerned only research that involves extracorporeal human embryos or parthenogenetically activated oocytes. Research involving in utero human embryos, or fetuses, was not at issue, since guidelines for such research were already embodied in federal laws and regulations governing human subjects research. Research involving human germ-line gene modification also was not within the Panel's scope. In addition therapeutic human fetal tissue transplantation research, the topic taken up by this Panel's predecessor, was also not a part of the Panel's mandate because guidelines were already in place to govern such research (36).

In 1994 the NIH Director appointed the Human Embryo Research Panel to consider various areas of research involving the ex utero preimplantation human embryo and to provide advice as to those areas that (*1*) are acceptable for federal funding, (*2*) warrant additional review, and (*3*) are unacceptable for federal support. For those areas of research considered acceptable for federal funding, the Panel recommended specific guidelines for the review and conduct of this research. In addition to developing guidance for research deemed acceptable, the Panel addressed issues surrounding the sources of gametes and embryos for research, transfer of embryos to a uterus, parthenogenesis, and systems for review and oversight.

One of the most difficult issues the Panel had to consider was whether it is ethically permissible to fertilize donated oocytes expressly for research purposes or whether researchers should be restricted to the use of embryos remaining from infertility treatments that are donated by women or couples. The Panel concluded that studies that require the fertilization of oocytes are needed to answer crucial questions in reproductive medicine, and that it would therefore not be wise to prohibit altogether the fertilization and study of oocytes for research purposes. It concluded that the use of oocytes fertilized expressly for research should be allowed only under two stringent conditions: when the research by its very nature cannot otherwise be validly conducted or when a compelling case can be made that it is necessary for the validity of a study that is potentially of outstanding scientific and therapeutic value. One member of the Panel dissented from the Panel conclusion that under this condition oocytes may be fertilized expressly for research purposes.

It was this one of several recommendations that brought yet another federal bioethics panel into great controversy. Before the Panel had even finished presenting its findings to the NIH Director and his Advisory Panel, President William Clinton issued an Executive Order prohibiting the use of federal funds for research in which oocytes were fertilized expressly for research purposes. A subsequent congressional ban extended the prohibition to include any research that involves exposing embryos to risk of destruction for nontherapeutic research (P.L. 104-91 and P.L. 104-208).

BIOETHICAL ANALYSES IN OTHER FEDERAL SETTINGS

During the past three decades, bioethics debates have found a place in several federal settings besides focused national commissions and panels. For example, the OTA was established in 1972 as an analytical arm of Congress, to anticipate how science and technology would raise issues for policy makers, and to advise Congress on federal policies affecting science and technology development. It rendered technical advice about how to promote or regulate science and technology, and gave early warning about the impacts of emerging technologies. Eliminated by Congress in 1995, OTA contributed markedly to bioethics debates in its 24-year history, especially during the interludes when there was no standing federal commission. Its 1983 report on genetic testing in the workplace (37) explicitly incorporated bioethics analysis. A 1984 report on "Human Gene Therapy" (38) addressed the ethical issues of the use of recombinant DNA techniques in therapeutic treatment. A succession of reports included chapters on ethical considerations or had extensive discussion of ethical issues (37–50). Although bioethics was becoming an important component of the OTA analyses, OTA was not a bioethics commission and held a much broader mandate from Congress.

Matters of bioethics at the IOM, National Academy of Sciences (NAS), grew out of concerns with medical practice and public health. In its early years IOM fostered a small bioethics program, issuing a 1974 report, "The Ethics of Health Care," an inspection of the ethical underpinnings of medical practice (51). Since then IOM has continued to incorporate sections or chapters on bioethics into many reports. In 1994 IOM completed a study of issues in genetic testing (52) and in 1995 completed a systematic review of past and ongoing bioethics commissions at the federal, state, and international levels (53).

The National Human Genome Research Institute at the NIH and a parallel program under the Department of Energy (DOE) since their inception have devoted a fraction of their genome research budgets to analysis of the social and ethical implications. The Ethical, Legal, and Social Implications (ELSI) Program, established in 1990, is a grant-making and policy-making body within the NIH. It is currently the largest federal supporter of bioethics research, with an annual budget of approximately $7 million. Over the years ELSI research projects have focused on a wide range of issues including discrimination in insurance and employment based on genetic information, when and how new genetic tests should be integrated into mainstream health care services, informed consent in genetic research protocols, and public and professional education about genetics research and bioethics. An ELSI Working Group has an advisory role in overseeing the research portfolio of NIH and the DOE and helps formulate the requests for grant proposals and program announcements. Its mandate also includes the formulation of policy options. It has issued policy statements on the need for pilot studies of cystic fibrosis screening (54), on protection from genetic discrimination under the Americans with Disabilities Act (55), and on genetic discrimination and breast cancer.

PRESIDENT'S ADVISORY COMMITTEE ON HUMAN RADIATION EXPERIMENTS

In response to allegations that the United States had performed human radiation experiments and exposed unwitting human subjects to dangerous levels of radiation during the cold war, President Clinton in 1994 established the Advisory Committee on Human Radiation Experiments (ACHRE). The committee was charged with identifying the ethical and scientific standards by which experiments conducted by the federal government between 1944 and 1974 should be judged, and whether those experiments met those standards. Furthermore President Clinton asked the committee to consider whether there were identifiable medical or scientific purposes for the experiments, whether there was follow-up care for the subjects, and whether the experiments met the pre-1974 or current standards for informed consent and other ethical principles for research involving human subjects.

The fourteen member committee, comprised of one citizen, two lawyers, two ethicists, four scientists, three physicians, one statistician, and one professor of humanities, delivered its report a year and a half later, in October 1995. In accordance with its mandate, the report starts with an overview of the ethics of human subjects research from 1944 to 1974, which details the ethical principles adopted by the DOE and the then Department of Health, Education and Welfare, as well as the then Atomic Energy Commission. As additional background, common practices involving humans subjects research in medical research are detailed. The second part of the report focuses on specific experiments, such as plutonium injections and total-body irradiation, examining in depth the protocols used, the level of consent obtained by the researchers, as well as the risks and benefits to which the subjects were exposed. The third part then addresses the current picture in human subjects research, focusing on the existing federal system of human subjects protection. As part of its efforts to gain a complete picture of current practices, ACHRE also conducted an independent study of current research protocols as well as the perceptions of subjects involved in research.

In October 1995 ACHRE reported its findings and recommendations, finding "evidence of serious deficiencies in some parts of the current system for the protection of rights and interests of human subjects." In particular, following a survey of research proposal documents, the ACHRE concluded that these materials provided insufficient evidence with which to make judgments about the voluntariness of the subjects' participation and about the justification for involving individual specific subjects in the research. Review of consent forms also evidenced deficiencies, according to the committee, and patient-subjects interviewed in a separate study seemed to be confused about the difference between research and therapy (56). In addition there were several intentional releases of radioactive substances into the environment without the knowledge or consent of the surrounding community and hundreds of uranium miners died as a result of exposure to radon and other radioactive materials at levels in excess of those known to be hazardous, even though they were being monitored by the federal government.

While ACHRE found that most of the research conducted between 1944 and 1974 involved only minimal risk, they also identified experiments that violated accepted norms of informed consent and placed subjects at risk for cancer and other illnesses without appropriate consent processes. The committee also concluded that although protections mandated by federal regulations for human subjects research were by and large in place, these rules were adequate only when applied to healthy and independent human subjects.

The committee recommended that the government personally apologize to subjects who had involuntarily been exposed to substantial radiation and, where appropriate, offer them financial compensation. In addition the committee took note of both deficiencies in current protections for children and the mentally ill who were serving as subjects in medical experiments and with research subjects' perceptions that research is primarily therapeutic. Finally, ACHRE made a series of recommendations focused on removing current difficulties in interpreting current federal regulations, and, where appropriate, expanding them.

ESTABLISHMENT OF THE NATIONAL BIOETHICS ADVISORY COMMISSION

In the fall of 1993, the White House Office of Science and Technology Policy (OSTP) was approached by NIH, the DOE, and other agencies to consider establishing a standing expert commission on bioethics. The proposal stemmed in part from a congressional request that NIH and DOE establish an advisory committee on genetic privacy. OSTP, however, expressed a need for a high-level group to serve as a shared resource to address a broad set of ethical issues, including genetic privacy, and to complement specialized committees and boards already supported by the various mission agencies.

In August 1994 OSTP published in the Federal Register a draft charter for a National Bioethics Advisory Commission (NBAC). The resulting NBAC charter reflects public comments received as well as bipartisan input from Congress. NBAC provides advice and makes recommendations to the National Science and Technology Council and to other appropriate government entities on relevant bioethical issues.

In addition to chartering NBAC, the President also charged the executive branch agencies that conduct, support, or regulate research involving human subjects to review their policies and procedures for protection of research subjects. This directive was a response to the recommendations contained in the report of the President's Advisory Committee on Human Radiation Experiments, which had just concluded its review of protections (or lack thereof) of U.S. citizens exposed to radiation experiments several decades ago, including American soldiers who were purposely exposed to radiation during atmospheric nuclear tests. Three issues raised by the Advisory Committee on Human Radiation Experiments provided some of the impetus to create NBAC: (*1*) the need for a continuing public forum on the interpretation and application of ethics rules and principles for the conduct of human subject research, (*2*) the need to maintain consistency in ethical standards for human subjects research across the 19 federal agencies and departments that support such efforts, and (*3*) the need to review the current Institutional Review Board (IRB) system.

According to the President's Executive Order, federal agencies were required to report the results of their review to NBAC, which was to pursue, as its first priority, protection of the rights and welfare of human research subjects. The charter also requires that NBAC consider "issues in the management and use of genetic information, including but not limited to human gene patenting." The Commission also may consider additional issues suggested by executive branch agencies, Congress, and the public, or that originate within the Commission itself. NBAC is not a regulatory committee and does not review or approve individual projects. Rather, it defines and identifies broad overarching principles to govern the ethical conduct of research. The 18 members of NBAC are presidentially appointed and represent science, medicine, law, ethics, theology, and public policy. The Commission held its inaugural meeting on October 4, 1996.

In February 1997 the work of the Commission was diverted toward an unexpected development. Within days of the published report of the apparently successful cloning of a sheep using a technique called somatic cell nuclear transfer, President Clinton instituted a ban on federal funding for research related to cloning of human beings. In addition the President asked NBAC to address within 90 days the ethical and legal issues that surround the subject of cloning human beings. This provided an opportunity for initiating a thoughtful analysis of the many dimensions of the issue, including a careful consideration of the potential risks and benefits. It also presented an occasion to review the current legal status of cloning and the potential constitutional challenges that might be raised if new legislation were enacted to restrict the creation of a child through somatic cell nuclear transfer. The Commission quickly commissioned eight papers on the scientific, legal, ethical, religious, and policy aspects of the prospect of human cloning and met five times over the following 90 days. It delivered its report, "Cloning Human Beings," to the President at a White House ceremony on June 9, 1997.

The second topic undertaken by the Commission was the review of the federal agency reports on human subjects protections required by the Executive Order creating NBAC. This is an ongoing project with a comprehensive report due in the fiscal year 1999–2000.

The second topic also concerned an aspect of the protection of the rights and welfare of human research subjects; namely how ethically acceptable research can be conducted with human subjects who suffer from mental disorders that may affect their decision-making capacity. The Commission's report, which was delivered to the President in January 1999, made a number of recommendations to strengthen regulations in this area.

A third topic was an inquiry into the appropriate use of human biological materials, particularly those materials that have been collected in tissue banks over the last century. The analysis was carried out with a focus on the fact that developments in biomedical technologies

now enabled investigators to gather much more personal information about those who donated these samples than had ever been anticipated when most of these samples were collected. The Commission therefore addressed the increasing concern that the use of genetic information found in these materials might infringe upon an individual's privacy, and if misused could result in discrimination. In particular, NBAC assessed the adequacy in this new context of existing federal regulations for the protection of human subjects that are incorporated in the so-called Common Rule. The report with its recommendations was delivered to the President in July 1999.

Late in 1998 the President again made a special request to the Commission as a result of the announcement that scientists had been able to isolate and culture human embryonic stem cells. This immediately raised a number of long-debated ethical issues, since the only current sources of these materials were early embryos or fetal tissue. The President asked the Commission to advise on how to best take advantage of the great promise of these materials while also giving consideration to a broad range of the ethical issues involved. The Commission delivered its report and recommendations to the President in August 1999.

To date, NBAC has scheduled the release of two further reports, both in fiscal 1999–2000. The first of these will take up the ethical issues involved in biomedical research protocols involving sponsors and/or investigators and/or research sites within a number of different countries that may not all share the same ethical concerns in the biomedical area. They may differ, for example, on how human research subjects should be protected, or the conditions under which drug trials ought to be allowed to proceed. The second report will be the comprehensive assessment of the current efforts of federal departments and agencies and their grantees to implement existing regulations.

BIBLIOGRAPHY

1. U.S. Congress, Office of Technology Assessment, *Biomedical Ethics in U.S. Public Policy*, OTA-BP-BBS-105, U.S. Government Printing Office, Washington, DC, 1993.
2. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Research on the Fetus*, U.S. Government Printing Office, Washington, DC, 1975.
3. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Research Involving Prisoners*, U.S. Government Printing Office, Washington, DC, 1976.
4. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Research Involving Children*, U.S. Government Printing Office, Washington, DC, 1977.
5. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Psychosurgery*, U.S. Government Printing Office, Washington, DC, 1977.
6. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Disclosure of Research Information under the Freedom of Information Act*, U.S. Government Printing Office, Washington, DC, 1977.
7. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Special Study: Implications of Advances in Biomedical and Behavioral Research*, U.S. Government Printing Office, Washington, DC, 1978.
8. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Research Involving Those Institutionalized as Mentally Infirm*, U.S. Government Printing Office, Washington, DC, 1978.
9. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Institutional Review Boards*, U.S. Government Printing Office, Washington, DC, 1978.
10. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Ethical Guidelines for the Delivery of Health Services by DHEW*, U.S. Government Printing Office, Washington, DC, 1978.
11. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education, and Welfare, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, U.S. Government Printing Office, Washington, DC, 1978.
12. Ethics Advisory Board, U.S. Department of Health, Education, and Welfare, *Report and Conclusions: Support of Research Involving Human In vitro Fertilization and Embryo Transfer*, U.S. Government Printing Office, Washington, DC, 1979.
13. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Defining Death*, U.S. Government Printing Office, Washington, DC, 1981.
14. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Protecting Human Subjects*, U.S. Government Printing Office, Washington, DC, 1981.
15. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Whistleblowing in Biomedical Research*, U.S. Government Printing Office, Washington, DC, 1982.
16. President's Commission for the Study of Ethical Problems in Medicine and Behavioral Research, *Making Health Care Decisions* (with Vols. 2 and 3 appendixes), U.S. Government Printing Office, Washington, DC, 1982.
17. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life*, U.S. Government Printing Office, Washington, DC, 1982.
18. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Compensating Research Injury*, U.S. Government Printing Office, Washington, DC, 1982.
19. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Deciding to Forgo Life-Sustaining Treatment*, U.S. Government Printing Office, Washington, DC, 1983.
20. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Implementing Human Research Regulations*, U.S. Government Printing Office, Washington, DC, 1983.
21. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Screening and Counseling for Genetic Conditions*, U.S. Government Printing Office, Washington, DC, 1983.

22. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Securing Access to Health Care*, U.S. Government Printing Office, Washington, DC, 1983.
23. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Summing Up*, U.S. Government Printing Office, Washington, DC, 1983.
24. M.B. Abram and S.M. Wolf, *N. Engl. J. Med.* **310**(10), 627–632 (1984).
25. A.J. Weisbard, *Ethics* **97**, 776–785 (1987).
26. K.E. Hanna, R.M. Cook-Deegan, and R.Y. Nishimi, *Politics Life Sci.* **12**(2), 205–219 (1993).
27. J.D. Arras, *Cardozo Law Rev.* **6**(2), 321–346 (1984).
28. R. Bayer, *Cardozo Law Rev.* **6**(2), 303–320 (1984).
29. N. Daniels, *Just Health Care*, Cambridge University Press, New York, 1985, pp. 81–83.
30. N. Daniels, *Am I My Parents Keeper? An Essay on Justice between the Young and the Old*, Oxford University Press, New York, 1988.
31. J.C. Fletcher, *Hum. Gene Ther.* **1**(1), 55–68 (1990).
32. T.H. Murray, *Hum. Gene Ther.* **1**(1), 49–54 (1990).
33. R.M. Cook-Deegan, *The Gene Wars*, Norton, New York, 1994.
34. J. Childress, in K.E. Hanna, ed., *Biomedical Politics*, National Academy Press, Washington, DC, 1991, pp. 215–248.
35. Advisory Committee to the Director, *Human Fetal Tissue Transplantation Research*, National Institutes of Health, Bethesda, MD, 1988.
36. National Institutes of Health (NIH), *Report of the Human Embryo Research Panel*, NIH, Bethesda, MD, 1994.
37. U.S. Congress, Office of Technology Assessment, *The Role of Genetic Testing in the Prevention of Occupational Disease*, OTA-BA-194, U.S. Government Printing Office, Washington, DC, 1983.
38. U.S. Congress, Office of Technology Assessment, *Human Gene Therapy — A Background Paper*, OTA-BP-BA-32, U.S. Government Printing Office, Washington, DC, 1984.
39. U.S. Congress, Office of Technology Assessment, *Impacts of Neuroscience*, OTA-BP-BA-24, U.S. Government Printing Office, Washington, DC, 1984.
40. U.S. Congress, Office of Technology Assessment, *Reproductive Health Hazards in the Workplace*, OTA-BA-266, U.S. Government Printing Office, Washington, DC, 1985; reprinted by Lippincott, Philadelphia, PA.
41. U.S. Congress, Office of Technology Assessment, *Alternatives to Animal Use in Research, Testing, and Education*, OTA-BA-273, U.S. Government Printing Office, Washington, DC, 1986.
42. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology, 1: Ownership of Human Tissues and Cells — Special Report*. OTA-BA-337, U.S. Government Printing Office, Washington, DC, 1987; reprinted by Lippincott, Philadelphia, PA.
43. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology, 2: Public Perceptions of Biotechnology — Background Paper*, OTA-BP-BA-350, U.S. Government Printing Office, Washington, DC, 1987.
44. U.S. Congress, Office of Technology Assessment, *Life-Sustaining Technologies and the Elderly*, OTA-BA-306, U.S. Government Printing Office, Washington, DC, 1987; reprinted by Lippincott, Philadelphia, PA.
45. U.S. Congress, Office of Technology Assessment, *Losing a Million Minds: Confronting the Tragedy of Alzheimer's Disease and Other Dementias*, OTA-BA-323, U.S. Government Printing Office, Washington, DC, 1987; also available through Lippincott, Philadelphia, PA and Human Sciences Press, New York.
46. U.S. Congress, Office of Technology Assessment, *Infertility: Medical and Social Choices*, OTA-BA-358, U.S. Government Printing Office, Washington, DC, 1988.
47. U.S. Congress, Office of Technology Assessment, *Mapping Our Genes — Genome Projects: How Big? How Fast?* OTA-BA-373, U.S. Government Printing Office, Washington, DC, 1988; reprinted by Johns Hopkins University Press, Baltimore, MD.
48. U.S. Congress, Office of Technology Assessment, *Biotechnology 4. U.S. Investment in Biotechnology*, OTA-BA-360, U.S. Government Printing Office, Washington, DC, 1988.
49. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology 5. Patenting Life*, OTA-BA-370, U.S. Government Printing Office, Washington, DC, 1989; reprinted by Dekker, New York.
50. U.S. Congress, Office of Technology Assessment, *Neural Grafting: Repairing the Brain and Spinal Cord*, OTA-BA-462, U.S. Government Printing Office, Washington, DC, 1990.
51. L.R. Tancredi, ed., *Ethics of Health Care. Papers of the Conference on Health Care and Changing Values, November 27–29, 1973*, National Academy Press, Washington, DC, 1974.
52. Institute of Medicine, *Assessing Genetic Risks*, National Academy Press, Washington, DC, 1994.
53. Institute of Medicine, *Society's Choices: Social and Ethical Decision Making in Biomedicine*, National Academy Press, Washington, DC, 1995.
54. Working Group on Ethical, Legal, and Social Issues in Human Genome Research, *Workshop on the Introduction of New Genetic Tests*, National Institutes of Health and Department of Energy, Washington, DC, 1990.
55. Joint Working Group on Ethical, Legal, and Social Issues, *Genetic Discrimination and the Americas with Disabilities Act*, Submitted to the Equal Employment Opportunity Commission, Washington, DC, 1991.
56. Advisory Committee on Human Radiation Experiments, *Final Report*, U.S. Government Printing Office, Washington, DC, 1995.

See other entries CLONING, ETHICS; CLONING, OVERVIEW OF HUMAN CLONING; EDUCATION AND TRAINING, PUBLIC EDUCATION ABOUT GENETIC TECHNOLOGY; see also FEDERAL POLICY MAKING FOR BIOTECHNOLOGY entries.

FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, NATIONAL HUMAN GENOME RESEARCH INSTITUTE

KATHI E. HANNA
Science and Health Policy Consultant
Prince Frederick, Maryland

OUTLINE

INTRODUCTION

The Human Genome Project of the National Institutes of Health (NIH) and the Department of Energy (DOE) was initiated in fiscal year 1988 as a line item in the federal budget to map and sequence the entire complement of genetic information in the human genome. The project, the first major, federally funded biology initiative, was originally scheduled to take 15 years at a cost of approximately $3 billion. As of early 2000 the project is ahead of schedule and under cost.

The project is headquarted in the National Human Genome Research Institute (NHGRI), originally called the National Center for Human Genome Research (NCHGR). NHGRI is one of 24 institutes, centers, or divisions that make up NIH, the federal government's primary agency for the support of biomedical research. The collective research components of NIH make up the largest biomedical research facility in the world. NIH is part of the U.S. Department of Health and Human Services (DHHS).

Although much of the genome research effort takes place in the United States, the Human Genome Project is a worldwide scientific effort with the goal of analyzing the structure of human DNA by determining the location of the estimated 100,000 genes in the human genome and identifying the sequence of its 3 billion base pairs. The four major goals of the Human Genome Project include (*1*) mapping and sequencing the human genome; (*2*) mapping and sequencing the DNA of model organisms; (*3*) computerized data collection, storage, and handling of this information; and (*4*) examining and addressing related ethical, legal, and social implications of such a research effort. The information generated by the Human Genome Project will be a major resource for the areas of basic and applied biomedical and behavioral research in the twenty-first century.

The ambitious nature of the scientific goals of the project was itself a source of initial controversy. Simultaneously hailed as the search for the biological "holy grail" (1), and big science at its worst (2–4), the Human Genome Project is unprecedented in many ways. Besides being "big biology," the research alliance between NIH and DOE was also a unique first (5–8), as was the allocation of 3 percent of the research budget for the study of ethical, legal, and social implications of the application of knowledge gained from the mapping and sequencing research enterprise. Never before had the federal government rushed headlong into such an ambitious research program, while at the same time supporting efforts that would raise questions about the wisdom, pace, and potential social consequences of its actions.

Although these ethical and social concerns were not new when the Human Genome Project was first conceived — they were previously raised in concert with early genetic diagnostic capabilities such as sickle cell carrier screening and the use of prenatal diagnosis for selective abortion — the debate about the Human Genome Project brought many of these issues to the surface once again because of the scale and magnitude of the mapping effort. Whereas ethical, legal, and social concerns raised by the application of genetic technology to human health were previously addressed on a case-by-case basis, the accelerated pace of new discoveries from the Human Genome Project was likely to exponentially increase the volume and intensity of concerns, rendering a casuistic approach dangerously obsolete. The genome project is leading to genetic tests that will be faster, cheaper, more accurate, and more applicable to a multitude of diseases. Thus it was believed by the first leadership of the project that a more broad-stroked policy approach was required.

This scientific effort is already producing information that is leading to the detection and diagnosis of genetic disorders. The long-range goal, however, is to go beyond diagnostics and to provide improved treatment, prevention, and ultimately cures. The interim phase, the phase in which gene detection is possible but understanding of gene function is limited (and therefore treatment is unavailable), has been thought by some to be the period in which the most significant deleterious social, ethical, and legal consequences might arise, particularly concerning the use and potential abuse of such information in terms of discrimination, stigmatization, and potential medical harm (9,10).

An Early Commitment to Ethical Analyses

James D. Watson, co-discoverer of the molecular structure of DNA and an early proponent of a federal effort to map the human genome, recognized the need to confront the policy issues raised by the possible applications of genetic information early in the project. He reiterated his commitment to addressing concerns about the risks posed by misuse of genetic information at a press conference in October 1988 announcing his appointment as the first head of the NIH Office of Human Genome Research:

> Some very real dilemmas exist already about the privacy of DNA. The problems are with us now, independent of the genome program, but they will be associated with it. We should devote real money to discussing these issues. People are afraid of genetic knowledge instead of seeing it as an opportunity (11).

Watson felt that the NIH program should spend some of its genome money on pursuing the social, legal, and ethical issues raised by rapid advances in genetic knowledge. His

advocacy led to the creation of the Ethical, Legal, and Social Implications (ELSI) Program, a grant-making and policy-making body within NIH. In recent years NHGRI has committed 5 percent of its annual research budget to the ELSI program. The DOE Office of Energy Research, NHGRI's partner in the Human Genome Project, also reserves a portion of its funding for ethical and legal research and education.

Never before had a leading scientist taken such a strong position regarding the need to commit federal funds (which otherwise might have gone to funding research) to the study of the ethical implications of research. Watson continued to defend his surprising and somewhat controversial proposal as the debate about federal support for the project went on (11). Because concerns about the social and ethical implications of genetic research were not new in Washington — and, in fact, the subject of several congressional hearings as well as the work of the National Academy of Sciences (NAS), the congressional Office of Technology Assessment (OTA), and the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (President's Commission) (12,13) — some argue that Watson was wise to take the bull by the horns and preempt any attempt by policy makers to prematurely and perhaps unnecessarily inhibit progress through overzealous regulation or legislation.

The fact that an icon in modern American molecular genetics would argue so strongly for public funding for social studies of science was welcome news to some observers and suspect to others, who viewed the diversion of funds from science to social research as, at best, an "unavoidable political tax" that the shrewd Watson was willing to pay to accomplish scientific goals. An ethics tax, like any tax, is not without controversy. While it is encouraging that the Human Genome Project has an ethics component, the value of such an organization in affecting health care decisions, research agendas, and policy remains to be seen, even today. And some initial observers were downright skeptical. In the words of Judith Swazey:

> ELSI — an imagistically unfortunate acronym — certainly is being taken seriously by the social scientists, ethicists, lawyers, and assorted other scholars, who have seldom had such financial largesse available to them, and their studies should yield a body of interesting and in some case practically useful findings and recommendations. But in both the short term and the long run, the significance of the ELSI component will be greatly diminished if the concerns that generated it, and its work and results, are seen by scientists and clinicians as politically necessary but basically irrelevant appendages to the "real work" of the Genome Project (14).

ELSI is not the only component of the NHGRI involved in policy issues. The director's office provides overall leadership to the institute, sets policies, and develops scientific, fiscal and management strategies for both the extramural and intramural programs. The office oversees intramural, collaborative, and field research to study human genetic diseases and formulates research goals and long-range plans to accomplish the mission of the Human Genome Project, including the study of the ethical, legal, and social implications of genome research. The office coordinates the NIH human genome program with those of other federal and private agencies and with other international programs, and fosters support for international meetings, workshops, and other activities to promote efficient international coordination and data exchange.

ETHICAL, LEGAL AND SOCIAL IMPLICATIONS PROGRAM

The planners of the U.S. Human Genome Project recognized that the information gained from mapping and sequencing the human genome would have profound implications for individuals, families, and society. When the original Genome Center was approved and funded by Congress in 1990, its Advisory Committee and NIH and DOE staff had already developed a five-year scientific plan (15,16). Part of the plan addressed ethical, legal, and social considerations, specifically: (*1*) develop programs addressing the understanding of the ethical, legal, and social implications of the human genome project, and (*2*) identify and define the major issues and develop initial policy options to address them (15,16).

The Genome Center Advisory Committee, in its initial deliberations, decided to spin off working groups to address specific areas of the project, such as genetic testing and insurance discrimination and genetic testing and policies related to disability. Federal rules concerning working groups are intended to make them temporary. Advisory Committee members must chair each working group; the other members are actually ad hoc technical consultants, serving at the pleasure of the director of NHGRI.

Development of the ELSI Agenda

The first ELSI Working Group met in 1989 to define and develop a plan of activities. The Working Group operationalized its mission by agreeing to the following activities (16):

- Stimulate research on issues through grant making
- Refine the research agenda through workshops, commissioned papers, and invited lectures
- Solicit public input through town meetings and public testimony
- Support the development of educational materials
- Encourage international collaboration in this area.

Thus, at the operational level, the ELSI Working Group developed realistic and practical goals following the model of data gathering and dissemination. In a sense, its early mission was to study what policy makers and the public should study. In terms of policy making, the group developed the following objectives (16):

- Clarify the ethical, legal, and social consequences of mapping and sequencing the human genome through a program of targeted research
- Develop policy options at professional, institutional, governmental, and societal levels to ensure that

genetic information is used to maximize the benefit to individuals and society
- Improve understanding of the issues and policy options through educational initiatives at public, professional, and policy-making levels
- Stimulate public discussion of the issues and policy options.

In addition specific topics were recommended for research support, including fairness in the use of genetic information, the impact of knowledge of genetic variation on the individual, privacy and confidentiality, the impact of the Human Genome Project on genetic counseling, reproductive decisions influenced by genetic information, issues raised by the introduction of genetics into mainstream medical practice, uses and misuses of genetics in the past and the relevance to the present, and commercialization of the products of the Human Genome Project (16). At the time this agenda was set, much policy research had already been conducted or was underway on some of the topics, such as the use of genetic information by employers (17,18), in the criminal justice system (19,20), commercialization (21–23) and genetic testing when no therapy is available. One wonders whether the Working Group found existing work to be so inconclusive as to warrant repeat attention. Nevertheless, the development of a laundry list for topics to be addressed by future grantees is an expansive, if inefficient, method for setting priorities.

Eventually three sets of issues were identified as particularly important initial considerations: (*1*) privacy of genetic information, (*2*) safety and efficacy of new genetic testing options, and (*3*) fairness in the use of genetic information (16). While all critical issues, the issues were narrowly confined to what could be considered a civil liberties orientation. The narrow agenda was likely due to the lack of diversity in perspectives and membership among the Working Group membership, which was largely constituted of academicians and policy researchers (24). Were the membership of the Working Group more diverse, other equally important issues might have been placed on the agenda, such as the effects of commercial interests on the research agenda, intellectual property rights, conflicts of interest for genome scientists, and quality assurance and control beyond issues of safety and efficacy.

Beyond setting a research agenda, the ELSI program was initially assigned the broad goal to "develop the safeguards required as new genetic information is put to practical purposes" (15). The language of the ELSI documents is filled with ambitious verbiage such as "develop sound policy recommendations that will govern the confidentiality of genetic test results, insure equal access to adequate education and counseling for patients, establish minimum qualifications for clinicians, assure quality control for genetic tests, establish guidelines for genetic testing programs, and define ethical and legal responsibilities of clinicians who perform tests" (15).

The basic flaw in the original design of the ELSI program was that it failed to consider the fact that it has no authority to affect policy and no clear route for communicating the information it gathers to the policy arena. These flaws would be corrected later, when the Genome Center received NIH Institute status, and through the creation of its policy and ethics office, which could approach thorny issues with a more pragmatic and policy-oriented perspective.

In the Beginning: Policy Making through Extramural Research

What distinguishes ELSI from other national ethics bodies is its mandate to administer a grants program. The ELSI Research Program, established in 1990, is responsible for funding and managing research grants and education projects that examine ELSI issues at institutions throughout the United States. It also supports workshops, research consortia, and policy conferences related to funded research and education projects. The ELSI grants programs solicit proposals through program announcements (PAs) and requests for applications (RFAs). At NIH, the Center for Scientific Review (previously the Division of Research Grants) reviews all grants applications and assigns them to the appropriate study sections for peer review for scientific merit. The multidisciplinary review groups consist of bioethicists, educators, genetic counselors, lawyers, theologians, philosophers, psychologists, and geneticists.

Although the development of PAs and RFAs is an iterative process that ideally accounts with addressing critical needs identified by the sponsoring agency and the scientific community, the peer-review method of selecting grants cannot guarantee proper and appropriate attention is being given to important issues. It is not the best way to set a policy agenda because the only citizens with access to the process are those schooled in an academic or professional discipline and capable of responding to the requirements of grant writing. In many ways it is a reductionist process that runs the risk of ignoring the most pressing policy issues. Academicians are not representative of society and can be dangerously naive when it comes to public policy. On the other hand, setting a policy agenda through a bottom-up approach provides the potential for more long-term analytical approaches to issues that might otherwise be subjected to political winds (24,25).

ELSI grantees are hardly representative of the general population or a broad array of disciplines. They are all specialists in genetics and ethics having written numerous publications on the topics they propose to study. It is a small universe that directs and benefits from the ELSI grants program.

On the other hand, the ELSI program is principally designed to support academic research, and this it does well. In fact, one of the major products of the ELSI program has been articles published by the investigators it has funded. As of 1996 the extramural research effort had funded over 125 research and education projects and related activities. These projects have resulted in the publication of over 150 journal articles and books, the development of education programs aimed at health professionals and the general public, and the establishment of policy recommendations on issues ranging from the use of genetic tests to preventing discrimination based on genetic information (26).

Surely such productivity enhances the scholarly writings in the field of bioethics, and is consistent with the traditional output of federally funded research, but is it not yet clear whether these efforts have reached the public in the most effective manner possible. The majority of the writings that have arisen from the grant funds appear in peer-reviewed journals and the academic press, hardly accessible to most policy makers and much of the public.

Until recently there was no mechanism for ensuring that the results of these scholarly pursuits made their way back to the policy arena unless one relied, in the words of one grantee's abstract, on absorption of the facts by "a general audience of intelligent readers." This lack of feedback from its extramural program into the policy process was perhaps the most important barrier to the early ELSI program's efforts as a policy-making body.

Formation of the NHGRI Offices of Policy Coordination and Genome Ethics

Recognizing the need to analyze and coordinate policy issues within NHGRI and the broader community, NHGRI established the Office of Policy Coordination in the Office of the Director in 1995 (intramural program). This Office provides information and analysis on ELSI policy and legislative issues and sponsors workshops and conferences, which assist in the development of policy options and recommendations related to ELSI issues. In addition, the Office has established collaborative relationships with a number of other NIH Institutes and has formed cooperative relationships with a number of other federal agencies, such as DOE, the Centers for Disease Control and Prevention (CDC), the Health Resources Services Administration (HRSA), the National Science Foundation (NSF), and the Food and Drug Administration (FDA).

In 1996 NHGRI established the Office of Genome Ethics (OGE) in the Division of Intramural Research to assist genome researchers in the NHGRI intramural program in identifying and addressing ethical issues arising from genome research.

The addition of these two offices moved NHGRI to a new level in policy making. Staff with the authority to act on behalf of the Institute are able to enter into policy discussions with congressional committees, other agencies, the scientific community, health advocacy groups, and the public.

SPECIFIC AREAS IN WHICH NHGRI HAS CONTRIBUTED TO POLICY DEBATES

The ELSI program and its bureaucratic counterparts in policy (mentioned above) have contributed to policy debates through support of extramural research, publications, and more recently, participation in legislative debates.

Privacy and Fairness in the Use and Interpretation of Genetic Information

As of 1996, the ELSI program had funded 31 research projects designed to examine the use (or misuse) and interpretation (or misinterpretation) of genetic information, including use by insurers.

Current projects examine the suitability of using genetic technologies for forensic and other law enforcement purposes; the need for standards; the development and application of supporting technology and instrumentation; the current understanding of the statistics and population genetics required in the interpretation of the data; and the social, legal, and ethical issues surrounding the use of these technologies (26).

Another project is developing opportunities for the public to comment on emerging genetic technologies. This project seeks input from lay and professional communities about how each feels about genetic technologies. The investigators are working with these communities to formulate model laws, institutional policies, professional standards of practice, and approaches to clinical decision making.

A current ELSI-funded study is designed to examine the assumptions made in medicine about health, normality, disease causation, and disease susceptibility, and another study is designed to explore the meaning of human genetics in popular culture. Gaining insight into both medical and popular ideas about the interpretation and understanding of genetic information can help in understanding the impact of this information on health care decisions, human relationships, and social policies.

The ELSI Working Group undertook one intensive effort to influence policy directly, rather than through discussion and grant making. As far back as 1990, members of the ELSI Working Group and the ELSI program staff had interactions with the U.S. Equal Employment Opportunity Commission (EEOC), expressing concerns about the lack of employment protections in place for individuals who might be identified as having genetic predisposition to disease. There were fears that although individuals with disabilities would be protected from discrimination under the Americans with Disabilities Act (ADA), individuals who were suspected or known to have a genetic predisposition to develop a disability or have children with disabilities would not be protected. As a result of discussions with the ELSI Working Group, ELSI program staff, and ELSI grantees, and following Commissioners' deliberations, the EEOC provided guidance on March 15, 1995, that clarifies that protection under ADA extends to individuals who may be discriminated against in employment decisions based on genetic information (27).

The ELSI Working Group had long been concerned about the fair use of genetic information, particularly as it relates to health insurance. In response to this concern, the ELSI Working Group spun off its first task force in 1991 on genetic information and health insurance. The Task Force released its recommendations in 1993, which included among others, recommendations that would prohibit the use of genetic information in denying or limiting health care coverage or services and ensure universal access to and participation in a program of basic health services, including genetic health care services (10).

In 1995 the ELSI Working Group developed and published the following recommendations for state and federal policy makers to protect against genetic discrimination (10):

- Insurance providers should be prohibited from using genetic information, or an individual's request for genetic services, to deny or limit any coverage or establish eligibility, continuation, enrollment or contribution requirements.
- Insurance providers should be prohibited from establishing differential rates or premium payments based on genetic information, or an individual's request for genetic services.
- Insurance providers should be prohibited from requesting or requiring collection or disclosure of genetic information. Insurance providers and other holders of genetic information should be prohibited from releasing genetic information without prior written authorization of the individual. Written authorization should be required for each disclosure and include to whom the disclosure would be made.

Genetic discrimination was also a priority for the National Action Plan on Breast Cancer (NAPBC), a public–private partnership established to address the research, education, and policy issues in breast cancer. Building on their shared concerns, the ELSI Working Group and the NAPBC co-sponsored a workshop on July 11, 1995, to address the issue of genetic discrimination and health insurance. Based on the information presented at the workshop and subsequent discussions, the ELSI Working Group, NAPBC, and NHGRI staff developed and published recommendations designed to protect against genetic discrimination.

The group recommended that employment organizations should be prohibited from using genetic information to affect the hiring of an individual or to affect the terms, conditions, privileges, benefits, or termination of employment unless the employment organization can prove this information is job related and consistent with business necessity. In addition employment organizations should be prohibited from requesting or requiring collection or disclosure of genetic information prior to a conditional offer of employment, and under all other circumstances employment organizations should be prohibited from requesting or requiring collection or disclosure of genetic information unless the employment organization can prove this information is job related and consistent with business necessity, or otherwise mandated by law. Furthermore written informed consent should be required for each request, collection, or disclosure. Employment organizations should be restricted from access to genetic information contained in medical records released by individuals as a condition of employment, in claims filed for reimbursement of health care costs, and other sources. Employment organizations should be prohibited from releasing genetic information without prior written authorization of the individual. Written authorization should be required for each disclosure and include to whom the disclosure will be made. Violators of these provisions should be subject to strong enforcement mechanisms, including a private right of action (29).

Since the publication of these recommendations, several bills have been introduced in state legislatures and the U.S. Congress to address the issue of genetic discrimination in health insurance The NAPBC/ELSI Working Group recommendations were considered during the development of a number of these bills. They were also taken into consideration in the deliberations about broader health insurance reform (26).

Clinical Integration of New Genetic Technologies

NHGRI has funded nearly 50 research projects to examine the impact of integrating genetic technologies into health care practice, to establish a better understanding of the current state of knowledge by health professionals, and to develop recommendations about how best to improve knowledge and incorporate these technologies into health care practice (26). The focus of work in this area has been on basic research in clinical ethical issues, professional issues and standards, and applied research designed to examine the impact of genetic testing and counseling. The ELSI program has sponsored two special initiatives in this area. The first, in 1991, was a Request for Applications (RFA) that solicited applications to study issues surrounding genetic testing and counseling for cystic fibrosis mutations. The second RFA, released in 1994, was designed to stimulate the study of issues surrounding genetic testing and counseling for heritable risk of breast, ovarian, and colon cancer.

In 1989 the cystic fibrosis gene was discovered, and mutations, which resulted in disease, were identified. Shortly after the discovery, concerns were expressed that there would likely be an increasing demand for such testing and that inadequate numbers of health professionals were prepared to provide such testing. Further concerns were expressed that not enough was known about how such testing could best be carried out safely and with appropriate pre-test education and post-test counseling. As a result NHGRI, along with the National Institute of Child Health and Human Development, and the National Institute of Nursing Research released an RFA, "Studies of Testing and Counseling for Cystic Fibrosis Mutations," to solicit applications in order to help establish practices that improve professional interpretation and patient understanding of CF testing, and more generally to examine alternative approaches to genetics education, testing, and counseling.

As a result of this initiative, eight research projects were funded. The findings of these studies suggested that interest in testing for cystic fibrosis mutations was much lower than had been expected in the general population. Investigators also discovered that although there was limited knowledge about genetics and cystic fibrosis in the population, it was possible to develop a variety of satisfactory alternative education strategies (e.g., videos and brochures) about testing. Furthermore the investigators saw no evidence of undue anxiety in most individuals tested (30,31).

A second major initiative was undertaken in 1994 by NHGRI in anticipation of the discovery of a number of cancer predisposing genes. This initiative (also an RFA) was co-sponsored by the National Cancer Institute, the National Institute of Mental Health and the National Institute of Nursing Research (32). It solicited applications for studies designed to examine the psychosocial and

clinical impact of using gene-based diagnostic tests in families with heritable forms of breast, ovarian, and colon cancer to identify those individuals who have an increased risk of developing cancer. Knowledge and attitudes about genetic testing for cancer risks are also being assessed, and information is being gathered to establish clinical protocols for the optimum use of these risk assessment technologies in the future. Once completed, these projects will provide valuable experience-based guidance for genetic testing for cancer susceptibility genes.

In 1996 the ELSI Working Group formed a Task Force on Genetic Testing. The Task Force was charged with examining the current state of genetic testing in the United States and (if needed) making recommendations to ensure the development and delivery of safe and effective genetic tests. This group specifically examined the scientific validation of new genetic tests; laboratory quality assurance; and education, counseling, and delivery of genetic tests. It sought broad participation by federal agencies, professional societies, the private biotechnology industry, insurers and consumers. A set of principles was published in 1997 (33).

Issues Surrounding Genetic Research

The ELSI program has supported research projects aimed at examining ethical, legal, philosophical, and ethnocultural issues surrounding genetics research. Research has been or is being conducted to examine issues surrounding informed consent in genetics research, explore how the research agenda was set for the Human Genome Project, study academe–industry relationships in genetics research, develop a legal research agenda, examine the impact of the Human Genome Project on women, and identify strategies for documenting the history of the Human Genome Project as it occurs. A research project designed to gather information about the status of informed consent for genetics research resulted in the development of recommendations regarding the components of informed consent for genetics research using stored samples (34).

To address concerns about informed consent in genetics research, the NIH Office for Protection from Research Risks (OPRR) and the ELSI program collaborated to convene a workshop to develop guidance for investigators and Institutional Review Boards (IRBs) who were increasingly being asked to approve genetics research protocols. The deliberations of this group resulted in the publication of a chapter on "Human Genetics Research" in the most recent version of OPRR's *IRB Guidebook*, which is distributed to IRBs. This is the first time that guidance on human subjects protections for genetics research has been provided in the *Guidebook* (35).

Stored tissue samples are valuable resources for genetics research. Due to increasing concerns about the adequacy of informed consent and privacy protections when stored tissue samples are used in genetics research, the CDC and the ELSI program supported a meeting to explore these issues. After intensive deliberations, recommendations were developed and published in December 1995 in the *Journal of the American Medical Association* (34). As a direct result of these deliberations, a number of other groups took up this issue, including the American College of Medical Genetics, the American Society of Human Genetics (36), the College of American Pathologists (37), and the National Bioethics Advisory Commission (38).

Public and Professional Education

ELSI supports projects designed to educate health and other professionals about genetics and genetic technologies, to develop formal curriculum materials for kindergarten through college-age students, and to educate consumers and the public about these issues.

For example, one study found that knowledge of genetics and genetic tests among physicians is increasing, but deficiencies in knowledge still exist (39). The study also revealed that primary care physicians are more likely to be directive when providing genetic tests rather than providing options from which patients may choose. Another survey revealed the limited amount of education in genetics of a wide variety of health professionals in university-affiliated programs (40). These health professionals reported that they deal on a daily basis with individuals with genetic disorders and their families and that they participate in providing genetic information and counseling to those families (41). They further recognize the need for more education in this area. This survey also revealed that consumers were more likely to have heard about the Human Genome Project than were health professionals. Information obtained through such surveys has been valuable in the ELSI program's efforts to examine its educational priorities and has led to the designation of health professional education as a high-priority area (25).

Another project was designed to educate state policy makers about the Human Genome Project and increase their knowledge about the social, legal, and ethical issues surrounding the research. Regional meetings were held around the country and a publication was developed and widely distributed to state lawmakers and other interested policy makers. A related project was designed to educate appellate judges and journalists about the Human Genome Project and its implications for the future. During the course of this project, an integrated textbook, a casebook, and a teaching manual appropriate for each group was developed and educational workshops provided.

CONCLUSIONS

A 1996 review of the ELSI program concluded that the establishment of an ELSI program at NHGRI was a "novel departure and an experiment" (25). The initial goals of ELSI, to raise the level of awareness of the ethical, legal, and social issues surrounding genetics research, have not been fully met, although tremendous progress has been made. The establishment of the Offices of Policy Coordination and Genome Ethics within NHGRI has increased the chances of advancing these goals. These offices have the authority and the capability to take the findings of ELSI-supported studies and communicate them to the communities making policy. Their creation has elevated the level of policy discourse about these issues.

Although NHGRI and its ethics and policy programs will continue to play a central role in addressing the ethical issues presented by the Human Genome Project, NHGRI does not stand alone in this effort. The U.S. Congress, other agencies of the federal government, professional organizations, universities and other research institutions, regulatory agencies, and industrial enterprises also have a vested interest in how these issues are debated and resolved. The ELSI program must work with all these parties, building on the experience it has gained and the relationships it has established in the first decade of the program.

BIBLIOGRAPHY

1. Hall, *Smithsonian* February (1990).
2. R. Lewin, *Science* **232**, 1598–1600 (1986a).
3. R. Lewin, *Science* **233**, 620–621 (1986b).
4. J. Walsh and J. Marks, *Nature* **322**, 590 (1990).
5. R.M. Cook-Deegan, in K.E. Hanna, ed., *Biomedical Politics*, National Academy Press, Washington, DC, 1991.
6. National Research Council, *Mapping and Sequencing the Human Genome*, National Academy Press, Washington, DC, 1988.
7. U.S. Congress, Office of Technology Assessment, *Mapping Our Genes — Genome Projects: How Big? How Fast?* No. OTA-BA-373. U.S. Government Printing Office, Washington, DC, 1988.
8. L. Roberts, *Science* **245**, 1438–1440 (1989).
9. N.A. Holtzman, *Proceed with Caution: Predicting Genetic Risks in the Recombinant DNA Era*, Johns Hopkins University Press, Baltimore, MD, 1989.
10. K.L. Hudson et al., *Science* **270**, 391–393 (1995).
11. J.D. Watson, NIH Press Conference, Videotape, Available at: National Reference Center for Bioethics Literature, Georgetown University, Washington, DC, 1988.
12. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life*, U.S. Government Printing Office, Washington, DC, 1982.
13. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Screening and Counseling for Genetic Conditions*, U.S. Government Printing Office, Washington, DC, 1983.
14. J. Swazey, in G.J. Annas and S. Elias, eds., *Gene Mapping: Using Law and Ethics as Guides*, Oxford University Press, New York, 1992.
15. National Center for Human Genome Research, *Annual Report I — FY 1990*, National Institutes of Health, Bethesda, MD, 1990.
16. U.S. Department of Health and Human Services and U.S. Department of Energy, *Understanding Our Genetic Inheritance; The U.S. Human Genome Project: The First Five Years FY 1991–1995*, No. 90-1580, National Institutes of Health, Bethesda, MD, 1990.
17. U.S. Congress, Office of Technology Assessment, *The Role of Genetic Testing in the Prevention of Occupational Disease*, No. OTA-BA-194, U.S. Government Printing Office, Washington, DC, 1983.
18. U.S. Congress, Office of Technology Assessment, *Genetic Monitoring and Screening in the Workplace*, No. OTA-BA-455. U.S. Government Printing Office, Washington, DC, 1991.
19. U.S. Congress, Office Technology Assessment, *Genetic Witness: Forensic Uses of DNA Tests*, U.S. Government Printing Office, Washington, DC, 1990.
20. National Research Council, Committee on DNA Typing, *DNA Technology in Forensic Science*, National Academy Press, Washington, DC, 1992.
21. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology, 1: Ownership of Human Tissues and Cells — Special Report*, No. OTA-BA-337, U.S. Government Printing Office, Washington, DC, 1987.
22. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology, 4. US Investment in Biotechnology*, No. OTA-BA-360, U.S. Government Printing Office, Washington, DC, 1988.
23. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology. 5. Patenting Life*, No. OTA-BA-370, U.S. Government Printing Office, Washington, DC, 1989.
24. K.E. Hanna, in R.E. Bulger et al., eds., *Society's Choices: Social and Ethical Decision Making in Biomedicine*, National Academy Press, Washington, DC, 1995.
25. K.E. Hanna, R.M. Cook-Deegan, and R. Nishimi, *Politics Life Sci.* **12**(2), 205–219 (1993).
26. Review of the Ethical, Legal and Social Implications Research Program and Related Activities (1990–1995), Available on *www.nhgri.nih.gov*
27. American Council of Life Insurance and Health Insurance Association of America, *Report of the ACLI-HIAA Task Force on Genetic Testing*, American Council of Life Insurance, Washington, DC, 1991.
28. NIH-DOE Working Group on Ethical, Legal, and Social Implications of Human Genome Research, *Genetic Information and Health Insurance: Report of the Task Force on Genetic Information and Insurance*, No. 93-3686, National Institutes of Health, Bethesda, MD, 1993.
29. K. Rothenberg et al., *Science* **275** (1997).
30. S. Loader et al., *Am. J. Hum. Genet.* **59**, 234–247 (1996).
31. P.T. Rowley, S. Loader, and J.C. Levenkron, *Gene. Test.* **1**(1), 53–59 (1997).
32. National Human Genome Research Institute, National Institute of Mental Health and National Institute of Nursing Research, *NIH Guide to Grants and Contracts*. **25**(13), April 26 (1996).
33. NIH-DOE Working Group on Ethical, Legal, and Social Implications of Human Genome Research, *Promoting Safe and Effective Genetic Testing in the United States: Final Report of the Task Force on Genetic Testing*, National Institutes of Health, Bethesda, MD, September 1997.
34. E.W. Clayton et al., *J. Am. Med. Assoc.* **274**(22), 1786–1792 (1995).
35. Office of Protection from Research Risks, *Protecting Human Research Subjects: Institutional Review Board Guidebook*, U.S. Government Printing Office, Washington, DC, 1993.
36. American College of Medical Genetics Storage of Genetic Materials Committee, *Am. J. Hum. Gen.* **57**, 1499–1500 (1995).
37. American Society for Investigative Pathology Position Statement, *Balancing Research Progress and Informed Consent*, October 1995.
38. National Bioethics Advisory Commission, *Research Involving Human Biological Materials: Ethical Issues and Policy Guidance*, U.S. Government Printing Office, Rockville, MD, 1999.
39. F.S. Collins, *J. Am. Med. Assoc.* **278**(15), 1285–1286 (1997).

40. E.V. Lapham and J.O. Weiss, in *Human Genome Education Model Project*, Georgetown University, Washington, DC, 1996.
41. D. Eunpu and J. Weiss, *J. Genet. Couns.* **2**, 93–113 (1993).

See other entries FEDERAL POLICY MAKING FOR BIOTECHNOLOGY.

FEDERAL POLICY MAKING FOR BIOTECHNOLOGY EXECUTIVE BRANCH OFFICE OF MANAGEMENT AND BUDGET

GILBERT S. OMENN
University of Michigan
Ann Arbor, Michigan

OUTLINE

INTRODUCTION

The Office of Management and Budget (OMB) is the nerve center of all Executive Branch agencies. OMB, with the President's directive and/or concurrence, issues budget guidance for the upcoming year, monitors budget commitments and outlays in the current year, and negotiates with agency heads the agency budget and policy and programmatic initiatives. Major disagreements are taken to the President by Cabinet secretaries or agency heads and the OMB director for resolution. OMB staff also coordinates all agency responses to congressional inquiries by reviewing the substance of these written documents and helps to anticipate the kinds of questions Senators, Congress members, and their staffs might raise. The aim of the annual Budget Request to Congress, and of all the attendant negotiations and testimony, is to try to put the President's stamp on the direction and priorities of the departments and agencies.

The budget process involves overlapping work on three fiscal years. During 1999, for example, the Congress enacted, with the President's signature, 13 appropriations bills covering the various departments' and agencies' budgets for fiscal year 2000 (FY00), beginning October 1, 1999. Sometimes Congress is unable to complete that process on time, leading to Continuing Budget Resolutions to keep the government functioning and able to pay its bills until final appropriations are enacted. Meanwhile the agencies function between October 1 of the preceding year and September 30 of the current year under the budget for the current fiscal year. OMB monitors the budget authority committed and the outlays actually made during the current fiscal year. In the example of the budget process during 1999, the agencies were already deeply engaged in budget preparations for fiscal year 2001 to be submitted to OMB in the fall of 1999 and by the President to Congress in February 2000.

The OMB began as the Bureau of the Budget in 1921, pursuant to an Act of Congress. In 1969 industrialist Roy Ash led a commission that recommended to President Nixon substantial changes in OMB, creating a management division and renaming the Bureau the Office Management and Budget. Considerable emphasis was laid on management improvement goals.

In *The Prune Book: The 100 Toughest Management and Policy-making Jobs in Washington*, the first of an invaluable series of books from the nonpartisan Council for Excellence in Government, the deputy director position in OMB was highlighted as focused on the management side of OMB. Government wide efforts to enhance the administrative operations of the Executive Branch drew more attention during the 1980s due to rapidly growing budget deficits and became a high-profile activity of Vice-President Gore during the Clinton administration.

The legendary OMB official Paul H. O'Neill, who rose from entry budget examiner in the Kennedy administration to deputy director of OMB in the Ford administration, told Prune Book author John Trattner that "OMB is a wonderful place to be. Every important issue of government goes through it" (1, p. 75) The title "Prune Book" is a knockoff of the title of the quadrennial "Plum Book," listing the approximately 4000 federal government positions to which the newly elected or re-elected president can make political appointments. Among the plums, the prunes require really able, well-prepared appointees!

The work of OMB was managed in the 1980s by program associate directors responsible for the areas of economics and government (e.g., Departments of the Treasury, Commerce, Justice); human resources, veterans, and labor (Departments of Health and Human Services, Veterans Affairs, Labor, Education, Social Security Administration); national security and international affairs (Departments of Defense and State, Agency for International Development, CIA, and other security agencies); and natural resources, energy, and science (Departments of Energy, Interior, Agriculture, National Science Foundation, NASA, Environmental Protection Agency). Subsequently, in the Clinton administration, the OMB directorate was divided to address Social Security and social welfare agencies in one division and health care financing and health research agencies in the other division. Before the 1970 reorganization, there were no appointed program associate directors in the Bureau of the Budget; the only political appointee was the director.

Another section of OMB, called the Office of Information and Regulatory Affairs, has grown substantially since the late 1970s and early 1980s. Efforts to characterize the full impact on the economy of federal regulations, especially health, safety, and environmental regulations, included concepts of a "regulatory impact," the costs and benefits to the economy and to society of certain major regulations; a "regulatory budget," the idea of limiting the aggregate of such costs or net costs in any year to some politically chosen maximum; "cost-effectiveness

analysis," comparing alternative paths to achieve similar protection of health or the environment; and "cost–benefit analysis," putting all health, environmental, and social benefits into dollar terms, then estimating the costs of compliance and of goods forgone, and calculating the ratio or net benefit. The Reagan administration and, later, the Republican Congress of 1995 to 1996 put a high priority on regulatory reform, which was interpreted by environmentalists and many Democrats as a dismantling of the regulatory agencies and regulatory protections so hard-won, rather than just a drive for more cost-effective, efficient regulatory priorities and programs, which all administrations have aimed to achieve. Currently agencies and OMB are working to interpret and implement the Government Performance and Results Act of 1993.

A certain tension exists in OMB. Only the several top officials, including the program associate directors, are political appointees who change with each administration. The career staff represent the ongoing collective institutional history of the Executive Branch. They have a long tradition of pride in their professionalism. Former OMB political appointees commonly describe the career staff as "the best in government." In recent administrations, the career staff have been tapped for important budget and administrative assistant secretary positions in various agencies by shrewd agency heads and Cabinet secretaries. Outsiders, like congressional appropriations staff members, have described the internal process at OMB as requiring for a program associate director position "someone who is very politically sensitive and knows how to deal with the Hill [Congress] and the agencies, or someone who is very analytic. The best is a combination of both. You have a pretty high-powered staff throwing a lot of information and data at you and trying to push you into a position usually against the agencies, and you have to be able to look at that critically" (1, p. 81).

OMB ROLE IN BIOTECHNOLOGY

OMB has central responsibility for the review of budget requests from all agencies conducting or funding biotechnology research, including the National Institutes of Health (NIH), the National Science Foundation (NSF), the Department of Agriculture (USDA), and the Environmental Protection Agency (EPA). Similarly OMB can become involved with agencies whose policy interests are in the development of this sector of the economy (Department of Commerce); in safety for workers, the public, and the environment (Occupational Safety and Health Administration in the Department of Labor, Food and Drug Administration, Environmental Protection Agency); in potential for military applications (DOD, CIA and other security agencies, Department of State); and in international cooperation in all aspects of R&D, applications, and regulation (EPA, Justice, State, NIH, NSF, Defense, Agriculture, AID, Commerce). OMB also has responsibility for extensive vetting within the administration and coordination of views and statements before Executive Branch officials testify before congressional committees or submit written responses to such committees.

As long as the agency activities are continuing on a generally approved course consistent with administration policy and there is good cooperation among agencies without conflicts or turf battles requiring mediation by OMB, any topical area like biotechnology gets little attention in a given year. Periodically, however, a department, agency, or the President may decide to highlight an agency initiative or research and development (R&D), as has recently happened with biotechnology. Alternatively, external events — like congressional fights over fetal research, the reported cloning of the sheep Dolly in Scotland, business sector demands for greater access to government biotechnology research, disputes over what is patentable, or reports of biological warfare agents in other countries — bring this topic to the fore.

In addition OMB often cooperates with other Executive Office agencies, such as the Office of Science and Technology Policy, the Office of Environmental Policy, and the Domestic Policy staff, in reviews of R&D initiatives and performance and coordination of R&D agencies. Thus OMB was an active participant in the late 1970s as recombinant DNA research and biotechnology start-up companies were emerging, and again in the 1980s as biotechnology faced challenges from environmentalists and the Congress about potential hazards of environmental, agricultural, medical, and chemical industry applications. OMB staff participated in an initiative that led to a 1986 report from the White House Domestic Policy Council's Working Group on Biotechnology; that Working Group was co-chaired by Bernadine Healy, then associate director of the Office of Science and Technology Policy (OSTP) and later director of the NIH, and David Kingsbury, associate director of the NSF. The Working Group report laid out a plan for coordination of agency roles and for a balance in the regulation of biotechnology research and the stimulation of product development. Biotechnology applications were again highlighted by the Clinton administration's National Science and Technology Council, a broad interagency effort coordinated by the OSTP, in 1993 and 1994.

A major and continuing investment is the National Human Genome Project, the effort to sequence the human genome. Its scale has attracted the interest of the OMB and the President, and there is sufficient turf competition to require OMB mediation. Discourse in the scientific community during 1985 and 1986 led to first a competitive and then a cooperative initiative by both NIH and the Department of Energy's (DOE) Office of Energy Research and National Laboratories with line items in the FY88 budget. Interestingly the budget request for FY88 from the Department of Health and Human Services (DHHS) for the NIH effort on the Human Genome Project was $25 million. According to OMB sources, it was OMB that increased the request to the nice, round figure of $100 million suitable for emphasis in the President's budget request. For excellent scientific reasons, the mandate was expanded to sequence genomes of other organisms — both infectious agents and "model organisms," like yeast and earthworms and mice, for studying underlying biology and comparing and inferring gene relationships and evolutionary changes with humans.

The NIH program matured into a distinct NIH organ, the National Human Genome Research Institute, in the mid-1990s. Both the NIH and DOE programs have included components of special funding on ethical, legal, and social implications of the new genetics, research and conferences that generate lots of policy issues requiring the administration's attention and OMB coordination.

BUDGET INITIATIVE AND ANALYSIS

Fiscal year 1991 stands out as the year in which the President's Budget Request to Congress included a major section on R&D, titled "Expanding the Human Frontier." Another section dealt with "Improving Productivity and the Quality of Life through Biotechnology" (2, pp. 59–63). Detailed presentation of that budget document is revealing about the role of OMB and of the government more generally in this important area of R&D and social policy.

Using data from 1987 published by the Congressional Office of Technology Assessment in 1988, the document included an introductory figure showing that the sources of U.S. biotechnology investment for 1987 were 59.4 percent federal government, 38.3 percent industry, and 2.3 percent state governments. Of course, "biotechnology" had to be defined and be put in context. Thus the document begins, "Biotechnology is an ancient practice that includes such familiar applications as the use of yeast in baking bread and cultures in making cheese. Recent breakthroughs in biotechnology, such as recombinant DNA techniques, cell fusion, and gene therapy, offer unprecedented opportunities for improving the nation's productivity, health, and well-being. Uncertainties in the returns on biotechnology investment, however, stemming from market barriers and unnecessary regulation, have retarded progress. Increasing Federal investment in basic biotechnology research will spur further advances, as will initiatives that improve the payoffs on investments" (3, pp. 72–75).

The characterization of regulation, an editorial insertion, reflects policy views of the Bush administration. It was then felt that the long struggle from the beginning of the recombinant DNA era had overcome fears of magical and mystical powers of the technology, as capable of creating wholly new living things with unpredictable properties. This was 14 years after science advisers in 1977 had explained to senior staff in EPA, OMB, and the Domestic Policy office what recombinant DNA techniques were. Many of these individuals were attorneys, as were some of the congressional staff addressing these issues in light of the scary, though wise, decision of the scientific community to issue a moratorium on recombinant DNA research three years earlier at the famed Asilomar Conference, in order to develop biological containment techniques. The 1977 discussions helped explain the use of standard techniques of preparation of DNA and proteins, and actions of enzymes; use of chromatography or electrophoresis to separate and identify molecules; special and predictable roles of the remarkable restriction and ligase enzymes; and specific experiments and proposed experiments. A particularly memorable event was the November 8, 1977, hearing of the Senate Committee on Science and Technology, chaired by Senator Adlai Stevenson of Illinois, at which prominent scientists, NIH Director Donald Frederickson, and Presidential Science Adviser Frank Press testified (4). In their exuberance about the potential for recombinant DNA applications and their disdain for the emerging NIH Guidelines for Recombinant DNA Research, some testifying scientists nearly brought the wrath of the Congress down upon the research community in the form of prohibitive regulatory requirements. Such a result was averted through cooperative review of the NIH process and the scientific methods themselves (5).

The FY91 Budget claimed that "Advances in biotechnology hold much promise. They can help improve the availability and quality of the food supply; prevent, identify, and cure disease; and reduce the hazards of industrial waste. Cell fusion, the merging of the genetic material of two cells of different species, can accelerate the selective breeding process for producing hardier and more fruitful crops and livestock. Gene therapy, replacing defective genetic material with normal DNA, may enable doctors to attack directly the source of major diseases, including cancer. In drugs, foods, agriculture, waste management, and energy, biotechnological advances offer the possibility of improvements that will make a real difference in people's lives. In this sense, biotechnology is an "enabling" technology: we may be able to make products safer or more cheaply, and we may be able to produce goods that we could not produce at all using traditional methods" (2, p. 59).

Clearly, the Bush administration, led by the OMB Director, Richard Darman, was hitching itself to the biotechnology revolution. "Biotechnology is a classic case of investing for the future. U.S. industry is spending at least $2 billion a year on biotechnology research and development, even though sales of products manufactured using biotechnology only reached the $1 billion mark for the first time in 1989. It is clear that the private sector believes the return on this investment will be great. The budget reflects a similar belief for the Federal investment" (2, p. 60).

The FY91 Budget proposed $8.6 billion, an increase of $218 million (6 percent) over the FY90 Budget, for biotechnology research and development. As all readers might imagine, the "budget" reported for such R&D depends mightily on the criteria for inclusion or exclusion. In fact it is common for experienced officials who must respond to requests from policy makers, Congress, or journalists for the amount of spending on a given need or opportunity to decide first whether it is in one's policy interest to generate a "big number," indicating a lot going on, or a "small number," indicating a need for much more investment, given the presumed opportunities. A smaller number offers the attractive feature of making any increment a bigger percentage increase, of course.

The estimate was built up through a host of not necessarily consistent definitions in 12 specified agencies: the DHHS (primarily the NIH, but also the Food and Drug Administration, FDA); the DOE, Commerce (including the National Institute for Standards and Technology, NIST, and the National Oceanic and Atmospheric Administration, NOAA), Defense, Veterans Affairs, and

Agriculture; the NSF, and the EPA. In fact NIH has struggled for decades with the definitions devised for characterization of research as "basic" or "applied" for governmentwide summaries. A close reading of the FY91 narrative several years later reveals some inconsistencies even in the document. There was no tabulation of the sources, by agency, of the $8.6 billion total. Only two agencies' budgets were highlighted. NIH was described as having a proposed increase of $280 million (more than the $218 million total for all agencies combined, noted above), of which $48 million was in the Human Genome Project. The National Research Initiative for Agriculture was described as having a doubling of the competitive research grants program, from $48 to $100 million, of which half was expected to be related to plant and animal biotechnology research. The description of DOE roles had no budget information.

The Budget did describe the emergence of 400 start-up firms and diversification of an estimated 200 existing companies into biotechnology and 200 companies supporting biotechnology with materials, instruments, equipment, and services. Data from 1987, probably the OTA report (6) (see the discussion above) were cited as showing $110 million in biotechnology-related expenditures by states, of which 38 claimed to be active in such investments in a 1986 survey through centers of excellence, university initiatives, incubator facilities for new firms, or grants for research projects.

On the management side, the administration highlighted the use of Cooperative Research and Development Agreements (CRADAs) by the NIH. The DOE and its national laboratories similarly utilized this mechanism to stimulate collaborations with industry and spin-offs of federal research into companies. Federal scientists were permitted and encouraged to begin filing patent applications. The Budget expanded funding for the Patent and Trademark Office (PTO) of the Department of Commerce, which in 1988 had instituted a 13 point plan to accelerate the review and award of biotechnology patents. PTO joined forces with the biotechnology industry to create a Biotechnology Institute to enhance training and technical expertise in the Patent Examining Corps. The FY91 Budget's process also introduced the concept of a system of user fees at FDA to provide staffing and other resources that would speed up the review of drugs and medical devices, including products that use techniques developed through biotechnology.

With regard to regulation, the FY91 Budget document emphasized the Coordinated Framework for Regulation of Biotechnology, the 1986 report from the Domestic Policy Council published in the Federal Register and mentioned above, and the ongoing Biotechnology Science Coordinating Committee, under the aegis of the Council on Competitiveness. Citing scientific assessments from the National Academy of Sciences and the Congressional Office of Technology Assessment, the Framework concluded that regulation should focus on the characteristics and risks of an organism or product, not on the process by which it was produced (i.e., by recombinant DNA, or cell fusion, or other biotechnology methods versus chemical synthesis). At the agency level the FDA concluded that no new procedures are required for genetically engineered products. In contrast, the USDA developed a new rule to review genetically engineered organisms that could possibly pose a high risk as a plant pest, and the EPA developed new regulations to tailor its rules and review procedures for microorganisms proposed as pest control agents or as bioremediation agents for soil and water contamination.

BUDGET FOLLOW-ON TO THE BUDGET INITIATIVE

The Budget for fiscal year 1992 contained a similar section (3) entitled "Expanding the Human Frontier through Biotechnology." Some of the budget numbers are surprising, revealing inconsistencies year-to-year in the Budget analysis. This time there was a tabulation of agency budgets for "biotechnology R&D," as shown in Table 1.

As shown in the table the narrower definition did not change the percentage increase, but it certainly did make the total effort seem smaller than did the figures highlighted in the FY91 document. OMB staff confirmed that the narrative was thematic, not tied analytically to the actual Budget. OSTP and OMB worked together, and with NIH and the Office of the Assistant Secretary for Health, DHHS, to seek standardized definitions and made big changes from year to year. As documented here, these shifts revealed one of the fundamental problems with biotechnology as a budget priority: The research was so hard to define even for descriptive purposes within and across agencies that no common understanding emerged and the budget effort was abandoned.

Otherwise, the narrative closely tracked the FY91 document, with highlights of achievements and trends in the diverse technologies being applied in human health, agriculture, foods and animal husbandry, and bioremediation and waste management for the environment. University/government/industry cooperation and investment in research training were emphasized, as was the ongoing influence of the 1986 Coordinated Framework for Federal Regulation of Biotechnology.

Current OMB senior staff have confirmed that no similar Budget highlights have focused on biotechnology

Table 1. Agency Budgets for Biotechnology

	Budget Authority in ($millions)	
Department or Agency	FY91 enacted	FY92 proposed
DHHS	$3296	$3557
USDA	119	139
DOE	110	140
NSF	130	132
DOD	118	123
VA, EPA, and other agencies	17	17
Total, all agencies	3788	4107 (+8%)
In addition:		
Directly-related activities	1663	1810
Broader science-based activities	1998	2144
Scale-up activities	25	32
Grand total (not given)	7474	8093 (+8%)

and no across-the-government estimates of budgets for biotechnology R&D have been generated by OMB since 1992. (OSTP did make an estimate of $4.3 billion in FY94.) In part, this observation reflects emphasis on other initiatives, and a kind of stay-the-course investment in the Human Genome Project and related areas. More important, it reflects the fact that biotechnology has become embedded in all areas of cell and molecular biology research and applications to human and animal diseases and industrial and environmental needs.

BIBLIOGRAPHY

1. J.H. Trattner, *The Prune Book: The 100 Toughest Management and Policy-Making Jobs in Washington, DC*, Madison Books, the Center for Excellence in Government, Washington, DC, 1988; 6th ed., 2000.
2. Office of the President of the United States, *The Budget for Fiscal Year 1991*, Office of Management and Budget, Government Printing Office, Washington, DC, 1990.
3. Office of the President of the United States, *The Budget for Fiscal Year 1992*, Office of Management and Budget, Government Printing Office, Washington, DC, 1991.
4. F. Press, *Science* **211**, 139–145 (1981).
5. G.S. Omenn, A.N. Schechter, R.F. Goldberger, and A. Dean, eds., *The Impact of Protein Chemistry on the Biomedical Sciences, an International Symposium in Honor of Christian B. Anfinsen*, Academic Press, New York, 1984.
6. Office of Technology Assessment, *U.S. Investment in Technology*, vol. 2, Government Printing Office, Washington, DC, July, 1988.

See other entries FEDERAL POLICY MAKING FOR BIOTECHNOLOGY.

FEDERAL POLICY MAKING FOR BIOTECHNOLOGY, EXECUTIVE BRANCH, RAC

LEROY WALTERS
Kennedy Institute of Ethics, Georgetown University
Washington, District of Columbia

OUTLINE

EARLY HISTORY AND WORK OF RAC, 1974 TO 1983 AND 1984 TO 1990

The Recombinant DNA Advisory Committee (RAC) of the National Institutes of Health (NIH) has had a long and distinguished history. RAC was established in fall 1974, shortly before the Asilomar meeting on research with recombinant DNA. Credit for the creation of the advisory committee and for the idea of devising guidelines for the safe conduct of recombinant DNA research is shared by a committee of scientists chaired by Paul Berg, who suggested the approach, and by NIH Director Donald S. Fredrickson, who implemented the plan and shepherded the field through its early years (1,2, pp. 80–96, 154–263; 3, pp. 272–274; 4,5). The committee met for the first time in February 1975, immediately after the Asilomar meeting (2, pp. 99–153; 3, pp. 274–278). From that moment until the early 1980s the RAC set the safety standards for all recombinant DNA research being conducted in the United States. These standards became known as the NIH "Guidelines for Research Involving Recombinant DNA Molecules" (6). The NIH Guidelines were adopted, in whole or in part, by many other industrialized countries.

In the early years most recombinant DNA research was funded by NIH and the National Science Foundation (NSF), so academic researchers had little choice but to follow the Guidelines. However, private companies also voluntarily complied with the RAC's Guidelines, in part to avoid regulation by their states or municipalities. While Congress considered numerous bills that would have regulated recombinant DNA research, especially in 1977, in the end the Congress deferred to the NIH and the RAC (2, pp. 312–337).

Historical research has demonstrated quite conclusively that a major violation of the Guidelines occurred in January 1977. A plasmid that had not yet been certified by NIH (pBR322) was used in the laboratory of Herbert Boyer of the University of California at San Francisco (UCSF) in an experiment that successfully cloned the rat insulin gene (7, pp. 112–180). NIH was informed of the violation in March 1977, but because of upcoming congressional hearings and pending legislation, the incident was kept quiet (7, p. 136). There is disagreement among witnesses about whether the prohibited clones were completely destroyed during March 1977; it is at least possible that the rat insulin genes were removed from the clones and re-used in an experiment with a certified vector in April (7, pp. 137–139, 167–173). In May 1977 the UCSF researchers announced their success in cloning the rat insulin gene at a press conference, and their paper documenting the experiment appeared in the June 17th issue of *Science*. However, a September 30th article by Nicholas Wade in *Science* raised questions about the possible use of an uncertified vector by the UCSF researchers (8). This article led to an intense confrontation between Senators Adlai E. Stevenson III and Harrison Schmitt, on the one hand, and researchers Herbert Boyer and William Rutter, on the other, at a November 1977 hearing (7, pp. 169–173). Despite the criticism directed at the researchers by the senators, no legislation followed.

The first major revision of the Guidelines occurred in December 1978. The Secretary of Health, Education,

and Welfare, Joseph Califano, became deeply involved in the revision process. While the substantive provisions of the Guidelines were being substantially relaxed, Califano expanded the RAC to 25 members, appointed several additional nonscientist members, and insisted that NIH promptly prepare a plan to assess the risks of recombinant DNA research to human health and the environment (2, pp. 247–248).

By late 1978 and 1979 it was becoming clear that most kinds of laboratory research with recombinant DNA were safe for both laboratory workers and the environment. New issues arose, however, such as the use of recombinant DNA techniques for large-scale production of human insulin (2, p. 355; 9) and the deliberate release of recombinant DNA into the environment, for example, to lower the temperature at which strawberry plants freeze. These new technologies, while initially overseen by the RAC, gradually moved to the appropriate regulatory agencies, the Food and Drug Administration (FDA) and the Environmental Protection Agency (EPA).

As NIH and the RAC ceded oversight authority regarding these areas of biotechnology to other agencies in the early 1980s, it appeared that RAC's advisory role would gradually diminish and might, in fact, disappear. Novel host-vector systems were seldom submitted to RAC, and most problems of physical and biological containment seemed to have been solved. Through a strange and perhaps fortuitous quirk of history, the research base for a new technique called "human gene therapy" was gradually being developed in the early 1980s. This technique had certain continuities with the laboratory research that had been RAC's central focus in the 1970s. From one perspective, gene therapy was the introduction of recombinant DNA (or products derived from recombinant DNA) into human beings. However, gene-therapy research was clearly a hybrid field. On the one hand, it was highly technical and required the expertise of molecular biologists, virologists, and human geneticists. On the other hand, gene-therapy research was human-subjects research, which was governed by its own set of rules and which was, at least in its broad outlines, quite comprehensible to laypeople.

Two events in 1980 had sparked public interest in the topic of human gene therapy — as well as in genetic engineering more broadly (10,11; 12, pp. 146–150). The first was a letter sent in June to President Carter from leaders of the Jewish, Catholic, and Protestant religious communities, expressing concern about the potentially deleterious effects of genetic engineering (13, pp. 95–96). A few months later the *Los Angeles Times* broke the story of Martin Cline's unapproved attempts to perform gene therapy on two patients afflicted with β-thalassemia, one patient in Israel and the other in Italy (14, 15, pp. 189–267). From these two pivotal events one can draw a direct line to subsequent ethical and public-policy discussions of gene therapy in the United States.

In 1982 a report by a presidential commission on bioethics, *Splicing Life* (13), and a congressional hearing on human genetic engineering (16) framed the major ethical issues in gene-therapy research. In response to those hearings, NIH and RAC began in 1983 to consider whether the committee should volunteer to review gene-therapy research protocols on a case-by-case basis. Over the course of a year NIH and RAC moved step-by-step toward accepting the oversight of gene-therapy research, in part because its other work was essentially finished and in part because no other agency or committee was prepared at that time to review this emerging field of research. A working group on human gene therapy was established during the summer of 1984 as a subcommittee to RAC, and this working group began developing Guidelines for gene-therapy research, which became known as the "Points to Consider," in the fall of that year (10,11). Once again, Congress deferred to the Executive Branch and to its public advisory committee, RAC. It did not pass legislation regulating gene-therapy research, nor did it establish a presidential advisory committee on the Human Applications of Genetic Engineering, as recommended by Congressman Albert Gore, Jr., in H.R. 2788 (April 27, 1983). The Congressional Office of Technology Assessment also published a report in late 1984, *Human Gene Therapy: A Background Paper* (17) that seemed to accept the merits of the approach being taken by NIH and RAC.

What were the central ethical questions to be asked about any proposed gene-therapy research protocol? The many questions asked in the RAC's Guidelines — the Points to Consider document — can be reduced to four rather simple and straightforward questions:

1. What are the potential harms and benefits of the research to the research subjects who will participate in a planned study?
2. How will these potential harms and benefits be communicated to prospective research subjects so that they can make voluntary and informed decisions about whether to participate in the research?
3. How will the selection among potential research subjects be made in a fair and equitable way, especially in cases where more people want to participate than can be enrolled in a study?
4. How will the privacy of research subjects be protected and the confidentiality of their medical information preserved?

If it is possible to develop guidelines for an emerging field of biomedical research too early, RAC and its working group did so. Working group members hurried to finish polishing the Points to Consider document in the spring and summer of 1985, then had to wait for almost two years for even a "preclinical" gene-therapy protocol. In the summer and fall of 1988, the first gene-marking study was reviewed and approved by the working group (now called the Human Gene Therapy Subcommittee) and the parent committee, RAC. Finally, in 1990, two gene-therapy studies were reviewed and approved. On September 14, 1990, the first officially sanctioned gene-therapy study began when W. French Anderson, R. Michael Blaese, and their colleagues administered genetically modified T-cells to a four-year-old girl named Ashanti DeSilva (15, pp. 13–52, 326–349; 18, pp. 227–240).

In its guideline-writing efforts and its review of the earliest preclinical and clinical protocols, the RAC was

supported by a series of excellent NIH staff people in an office called the Office of Recombinant DNA Activities (ORDA). The professionalism of this staff, its commitment to the public health and the protection of human subjects, and the long tenure of many of its members all contributed significantly to any success that RAC has had in its oversight responsibilities over the years.

PARALLEL EFFORTS BY NIH AND FDA, 1991 TO 1995

Gene-therapy research gradually expanded during the early 1990s. The number of new gene-therapy and gene-marking protocols submitted year by year to RAC can be summarized as follows:

1990: 2	1993: 31
1991: 9	1994: 33
1992: 23	1995: 44

An important innovation in the monitoring of clinical protocols adopted by RAC deserves special mention. Pursuant to a suggestion by the late Brigid Leventhal, a pediatric cancer researcher from Johns Hopkins University, RAC created a system that asked researchers to submit for the annual reports on adverse events and biological (as opposed to clinical) efficacy in ongoing gene-therapy studies (19). This safety monitoring activity of RAC eventually became known as "data management" (20,21). Dr. Leventhal presented the initial data-monitoring report to RAC in December 1992 and covered 40 patients (24).

The data management activities of the years 1992 through 1994 laid the groundwork for one of RAC's finest achievements in its oversight of gene-therapy research. In preparation for the June 1995 RAC meeting, RAC members and the ORDA staff undertook a comprehensive review of gene-therapy and gene-marking studies that had been reviewed and approved to date. This review, which was published in *Human Gene Therapy* on September 10, 1996 (23), revealed that during the first four years of intensive gene-therapy research there were hints of benefit in several studies but that in no case had a patient been cured of his or her disease by this new experimental approach.

As RAC continued to review clinical protocols in the mid-1990s, several widely acknowledged problems in the conduct and oversight of clinical research began to become apparent for gene-therapy research as well. The first cluster of problems surrounded the role and work of local Institutional Review Boards (IRBs). IRBs frequently lacked the necessary expertise to evaluate the technical aspects of HGT protocols. In addition, there were sometimes conflicts of interest within the local institution, especially when gene-therapy programs were considered to be showpieces for academic medical centers. Further IRBs' work involved primarily the front-end approval of paper protocols, with little monitoring of the actual conduct of the trials. RAC encountered a second set of problems as it examined the consent forms submitted with gene-therapy research protocols. Consent forms were often incomplete, omitting important information about the proposed research. In addition the forms were often misleading, especially in their descriptions of Phase I studies. The questions of what the local institution would do in case a research subject were injured and who would pay for experimental procedures were frequently evaded by means of artful circumlocution. Retrospective analysis also revealed that consent forms were sometimes not updated to report, for example, toxicities encountered during the course of a trial.

In the early 1990s FDA had also greatly enhanced its capability to review Investigational New Drug (IND) applications that employed gene-therapy techniques (24,25). FDA officials and reviewers regularly attended RAC meetings and increasingly participated in RAC discussions. Researchers began to note differences in the kinds of information being sought by RAC and FDA, and some researchers also complained that they had to jump over two regulatory hurdles rather than one.

In response to these complaints and similar complaints by some AIDS activists and biotechnology companies, NIH and FDA sought, in 1994, to work out a system for dual submission of protocols and coordinated review. In retrospect, it seems quite clear that this well-meaning effort did not go far enough and that serious differences in emphasis and approach remained between NIH and its advisory committee, RAC, on the one hand, and FDA, on the other. The two agencies also failed to agree on how to develop a data-management system for gene-therapy research.

VERMA COMMITTEE REPORT AND ORKIN-MOTULSKY COMMITTEE REPORT, SEPTEMBER 1995 AND DECEMBER 1995

In September 1995 a committee chaired by Inder Verma submitted recommendations to NIH Director Harold Varmus regarding the appropriate role of RAC in the review of gene-therapy research. The committee concluded that the RAC had an important ongoing role in the review of such research but recommended that RAC publicly review only research protocols that raise novel questions, for example, protocols that employ a new vector or seek to treat a new disease. For all other protocols, those that do not raise novel questions, the Verma Committee recommended that the review be conducted solely by FDA (26,27).

Three months later, in December 1995, a committee chaired by Stuart Orkin and Arno Motulsky delivered a somber verdict on the first five years of publicly reviewed and approved gene-therapy research: Not a single study had demonstrated clinical benefit to patients from gene therapy alone. The committee recommended that more attention be paid to the infrastructure for gene-therapy research, including the development of better vectors and of a better understanding of human immunology (28,29).

EIGHTEEN MONTHS OF UNCERTAINTY, MAY 1996 TO OCTOBER 1997

In May 1996 NIH Director Harold Varmus announced his intention to abolish the RAC in a speech delivered in Hilton Head, South Carolina (30). This proposal was formulated more precisely in a *Federal Register* notice

published in July 1996 (31). Over the next year and a quarter RAC's future role was debated by academic people, patient advocacy groups, biotechnology companies, several members of Congress, and RAC members themselves. Two general revisions of the original plan were published in the *Federal Register*, the first in November 1996 and the second in February 1997 (32,33). Finally, on October 31, 1997, a new oversight system for gene-therapy research was formally announced in the *Federal Register* (34). According to this final plan, RAC and NIH would no longer approve or disapprove gene-therapy research protocols. Instead, the RAC would discuss protocols that raised novel issues and make suggestions to the authors of the protocols. It was understood by all that RAC discussions would also inform FDA reviewers in their confidential negotiations with the sponsors of gene-therapy research who had submitted the same protocols as part of the IND review process.

There are five other features of the October 1997 plan that are especially worthy of note. First, the Office of Recombinant DNA Activities accepted responsibility for developing a data-management system to assist RAC in its review of adverse events and its annual audit of gene-therapy research. Second, gene-therapy researchers had a clearly articulated duty to inform ORDA and RAC of any changes in RAC-reviewed protocols that occurred between the time of RAC review and time that the researchers received permission from FDA to proceed with their proposed research (under an IND). Third, gene-therapy researchers also had a clearly articulated duty immediately to report to ORDA the occurrence of "any serious adverse event" in a gene-therapy research protocol. Fourth, researchers were required to submit annual data reports to ORDA for inclusion in the data-management system and analysis by RAC. Finally, ORDA and RAC would plan Gene Therapy Policy Conferences to look at broad themes surrounding gene-therapy research, both in the present and in the future.

FROM OCTOBER 1997 TO THE PRESENT: HOW IS THE NEW SYSTEM WORKING?

There is some good news to report from the early years of the new oversight arrangement. The Gene Therapy Policy Conferences have been highly successful in promoting interdisciplinary discussion of three important topics: genetic enhancement, in utero gene therapy, and the use of lentiviruses as vectors. RAC members continue to be deeply committed to their public roles and have been quite forthright in expressing concern about being asked to treat adverse-event reports as proprietary information. Similarly the staff people at ORDA (recently made a part of the NIH Office of Biotechnology Activities, OBA), have devoted long hours to fulfilling the roles assigned to them under the October 1997 agreement.

However, in September 1999 the world of gene-therapy research was shaken to its foundations by the unexpected death of an eighteen-year-old patient, Jesse Gelsinger. Mr. Gelsinger, whose ornithine transcarbamylase (OTC) deficiency was under reasonable control through a combination of drugs and diet, was a participant in a study being conducted at the University of Pennsylvania. He died four days after receiving an infusion of an adenoviral vector and the OTC gene (35,36). Months of intensive review and a full day of discussion at the December 1999 RAC meeting failed to clarify the precise cause of Mr. Gelsinger's death.

In December 1999 inspectors from FDA charged that the researchers conducting the OTC deficiency protocol in which Mr. Gelsinger died had violated several FDA regulations (37,38). Further inspections led to a publicly released set of "investigational observations" and FDA's placing a clinical hold on all gene-therapy protocols currently being conducted by the University of Pennsylvania research group (39,40). The Penn researchers replied to the observations in February 2000 (41,42), and a final resolution of this tragic incident is awaited in the near future.

In response to the death of Mr. Gelsinger, the NIH Office of Biotechnology Activities (OBA) and FDA's Center for Biologics Evaluation and Review (CBER) initiated an intensive joint review of all gene-therapy protocols using adenoviral vectors. This review, conducted in the last three months of 1999, sought to gather and analyze all adverse events that had occurred in gene-therapy studies using adenoviral vectors — approximately 25 percent of all U.S. protocols. In addition, OBA reminded gene-therapy researchers whose protocols had been registered with NIH of their duty to report all serious events to NIH. At the December 1999 RAC meeting, a working group on adenoviral vectors reported a pervasive lack of standardization in the characterization and use of such vectors (43).

Taken together, the events and discussions that occurred from September 1999 through February 2000 revealed that there are serious problems in the current oversight system for gene-therapy research in the United States. First, an online database for the data-management system, discussed and planned for since 1994, is still not available in the year 2000 (44). Initially delays occurred because of FDA's 1995 decision not to collaborate in the development of the database. In recent years ORDA has not had sufficient staff or resources to complete the development of the database.

Second, many gene-therapy researchers who are covered by the NIH "Guidelines for Research Involving Recombinant DNA Molecules" have neglected to file immediate reports with ORDA of serious adverse events that have occurred in the trials they are conducting. According to a December 21, 1999, letter from former NIH Director Harold Varmus, to Congressman Henry Waxman, only 39 (or 5.6 percent) of 691 serious adverse events in gene-therapy research using adenoviral vectors had been reported to ORDA before October 1999, when NIH and FDA began their vigorous joint effort to gather and analyze those events (45).

Third, the lack of coordination between NIH and RAC, on the one hand, and FDA, on the other, has continued in certain arenas. The two parent agencies have had different histories and sometimes reflect those histories in divergent approaches to the same question. Important issues remain unclarified, for example, whether RAC is advisory to

FDA or not. Some modes of FDA-NIH cooperation that could have been initiated by October 1997, at the latest, commenced only in late 1999, in response to a crisis. In December 1999 FDA changed its standard operating procedure on two important points and began providing weekly summaries to OBA of amendments to gene-therapy research protocols and adverse-event reports that it has received during the preceding week.

CONCLUSION

At the beginning of 2000 both RAC and the field of gene-therapy research face uncertain futures. It is still the case that no published report has demonstrated clear efficacy for gene-therapy procedures alone — that is, without adjunctive therapy. Now one relatively healthy patient has died in a gene-therapy trial, and several additional patients seem to have experienced laboratory or clinical toxicities. The RAC is frequently cited as a model for the oversight of emerging biomedical technologies, for example, the field of xenotransplantation research. However, several important questions about the RAC's role vis-à-vis gene therapy and the role of future committees vis-à-vis emerging biomedical technologies remain to be resolved. Among these questions, some of the most important are the following:

- What should be the oversight body's relationship to the major federal funding agency for biomedical research — NIH?
- How should the oversight body relate to the major federal regulatory agency for the approval and licensing of new biotechnology products — FDA?
- What should be the relationship of the oversight body to the major federal regulatory office for the protection of human subjects — now called the Office for Human Research Protections?
- Should different guidelines apply to human-subjects research conducted with public support and research conducted with private funds?
- What would be the best mechanism for independent safety monitoring of ongoing clinical trials in important new fields of biomedical research?
- And how can the consent process be structured, and consent forms be written, in a way that discloses all pertinent information to prospective subjects, and does so in a setting where they can freely decide whether to become volunteers in the war on disease?

ACKNOWLEDGMENT

I wish to thank my colleague, Linda A. Powell, for helpful editorial assistance.

BIBLIOGRAPHY

1. P. Berg et al., *Science* **185**, 3034 (1974).
2. S. Krimsky, *Genetic Alchemy: The Social History of the Recombinant DNA Controversy*, MIT Press, Cambridge, MA, 1982.
3. D.S. Fredrickson, in K.E. Hanna, ed., *Biomedical Politics*, National Academy Press, Washington, DC, 1991, pp. 258–307.
4. J. Lear, *Recombinant DNA: The Untold Story*, Crown Publishers, New York, 1978.
5. N. Wade, *The Ultimate Experiment: Man-Made Evolution*, Walker and Company, New York, 1977.
6. Office of Recombinant DNA Activities, *Recombinant DNA Research: Documents Relating to "NIH Guidelines for Research Involving Recombinant DNA Molecules,"* vols. 1–20, National Institutes of Health, Bethesda, MD, August 1976–December 1995.
7. S.S. Hall, *Invisible Frontier: The Race to Synthesize a Human Gene*, Atlantic Monthly Press, New York, 1987.
8. N. Wade, *Science* **197**, 1342–1345 (1977).
9. *Fed. Regist.* **45**, 24968–24971 (1980).
10. L. Walters, *Hum. Gene Ther.* **2**, 115–122 (1991).
11. E. Milewski, *Recomb. DNA Techn. Bull.* **8**, 176–180 (1985).
12. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997.
13. United States, President's Commission for the Study of Ethical Problems in Health Care and Biomedical and Behavioral Research, *Splicing Life*, U.S. Government Printing Office, Washington, DC, November 1982.
14. P. Jacobs, *L. A. Times*, October 8, 1, 26 (1980).
15. L. Thompson, *Correcting the Code: Inventing the Genetic Cure for the Human Body*, Simon & Schuster, New York, 1994.
16. United States, Congress, House, Committee on Science and Technology, Subcommittee on Investigations and Oversight, *Human Genetic Engineering*, 97th Congr., 2nd Sess., November 16–18, 1982.
17. United States, Congress, Office of Technology Assessment, *Human Gene Therapy: A Background Paper*, OTA, Washington, DC, December 1984.
18. J. Lyon and P. Gorner, *Altered Fates: Gene Therapy and the Retooling of Human Life*, Norton, New York, 1995.
19. Minutes of the Human Gene Therapy Subcommittee Meeting, July 29–30, 1991, *Recomb. DNA Res.* **14**, 907, 910 (1994).
20. Minutes of the Human Gene Therapy Subcommittee Meeting, November 21–22, 1991, *Recomb. DNA Res.* **15**, 289–293 (1994).
21. Minutes of the RAC Meeting, June 1–2, 1992, *Recomb. DNA Res.* **15**, 707–708 (1994).
22. Minutes of the RAC Meeting, December 3–4, 1992, *Recomb. DNA Res.* **16**, 582–585 (1994).
23. G. Ross et al., *Hum. Gene Ther.* **7**, 1781–1790 (1996).
24. Center for Biologics Evaluation and Research, *Points to Consider in Human Somatic Cell Therapy and Gene Therapy*, FDA, Rockville, MD, 1991.
25. D.A. Kessler et al., *N. Engl. J. Med.* **329**, 1169–1173 (1993).
26. Ad Hoc Review Committee for the Recombinant DNA Advisory Committee (Inder Verma, Chair), "Executive Summary of Findings and Recommendations," September 8, 1995, Unpublished document submitted to NIH Director Harold Varmus.
27. E. Marshall, *Science* **267**, 1588 (1995).
28. Office of the Director, National Institutes of Health, "Report and Recommendations of the Panel to Assess the NIH Investment in Research on Gene Therapy," Stuart H. Orkin and Arno G. Motulsky, Co-Chairs, December 7, 1995. Unpublished report.
29. E. Marshall, *Science* **270**, 1751 (1995).

30. E. Marshall, *Science* **272**, 945 (1996).
31. *Fed. Regist.* **61**, 35774 (1996).
32. *Fed. Regist.* **61**, 59725 (1996).
33. *Fed. Regist.* **62**, 7108 (1997).
34. *Fed. Regist.* **62**, 59032 (1997).
35. R. Weiss and D. Nelson, *Washington Post*, September 29, A1, A21 (1999).
36. N. Wade, *N.Y. Times*, September 29, A20 (1999).
37. R. Weiss and D. Nelson, *Washington Post*, December 8, A1, A10 (1999).
38. S.G. Stolberg, *N.Y. Times*, December 9, A22 (1999).
39. R. Weiss and D. Nelson, *Washington Post*, January 22, A1, A12 (2000).
40. S.G. Stolberg, *N.Y. Times*, January 22, A1, A10 (2000).
41. R. Weiss and D. Nelson, *Washington Post*, February 15, A3 (2000).
42. S.G. Stolberg, *N.Y. Times*, February 15, A16 (2000).
43. P. Smaglik, *Nature* **402**, 707 (1999).
44. Minutes of the RAC Meeting, June 8–9, 1995, *Hum. Gene Ther.* **7**, 417–422 (1996).
45. H. Varmus to H. Waxman, Letter dated December 21, 1999, p. 7.

See other FEDERAL POLICY MAKING FOR BIOTECHNOLOGY entries.

FEDERAL REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE, FDA, ORPHAN DRUG ACT

ROBERT A. BOHRER
California Western School of Law
San Diego, California

OUTLINE

The Orphan Drug Act was enacted to provide sufficient additional incentives to spur the development of therapeutics for smaller patient populations (1). One of the key incentives is market exclusivity for orphan drugs. The Federal Drug Administration (FDA) may not approve a second application for the same drug for the same orphan indication for seven years. The incentives available under the Orphan Drug Act have proved attractive to biotechnology companies, since several of the first and most important products of biotechnology, including human growth hormone and erythropoetin, were approved as orphan drugs. Both of these products were the cause of controversy over the interpretation of the "same drug" market exclusivity. FDA's regulations in response to those early problems have failed to eliminate the "same drug" interpretation problem, as can be seen from the most recent biotechnology orphan drug approvals of two forms of interferon-β for multiple sclerosis.

OVERVIEW OF THE ORPHAN DRUG ACT

Orphan Drug Act's Purpose and Incentives

A drug can be an "orphan drug" if it is for a rare disease or condition that affects fewer than 200,000 patients in the United States or for which there is no reasonable expectation of recovering the cost of developing the drug from sales in the United States (§360bb). Ten to 20 million Americans (about 9 percent of the population) suffer from more than 5000 rare diseases and that number will likely increase as the Human Genome Project uncovers more genetic causes of human diseases.

The purpose of the Orphan Drug Act (ODA) is to provide incentives for the research and development of new drugs for these smaller patient populations. ODA is intended to assist drug companies with two of the key drug development constraints, the cost and duration of the FDA approval process and the issue of intellectual property protection. ODA can reduce the development costs of orphan drugs before FDA approval by facilitating and expediting the FDA review process, and it can increase the financial returns from the development of orphan drugs after approval by providing additional market exclusivity. The development costs required to obtain FDA approval are reduced through a number of ODA's provisions. FDA provides help to pharmaceutical companies regarding the FDA's drug approval process, and FDA advice on the planning of clinical trials may be of particular value to small biotechnology companies with little prior experience in drug development. The Internal Revenue Code (2) provides tax breaks for expenses related to orphan drug development, and the FDA may help fund the clinical testing necessary for approval of an orphan drug (§360ee). In addition the Orphan Products Board coordinates the federal agencies involved in drug research and regulation. After approval, the intellectual property

protection provisions of ODA (§360cc) provide a seven-year term of exclusive marketing rights for the drug in the orphan disease population to increase the financial returns for the orphan drug sponsor.

For manufacturers the market exclusivity provision is likely to be the most important incentive offered by ODA because of the potential for very large profits during the period of exclusivity. ODA market protection is narrow, since only the use of that drug for treating the designated rare disease is protected. However, if the drug is not approved for any other medical indication, then orphan drug exclusivity is intended to be more or less as effective as patent protection. A second pharmaceutical manufacturer may seek FDA approval of a different drug for the same disease (or the same orphan drug for different orphan diseases or non-orphan diseases), but the sponsor of a subsequent drug for the same disease bears the burden of proof to demonstrate that its drug is different.

In placing the burden of distinguishing its drug on the second orphan drug sponsor, ODA parallels patent law, an older and more comprehensive body of law for providing exclusive rights as an incentive for innovation. Every patent applicant must disclose all prior relevant inventions and public knowledge ("prior art") to the Patent Office and explain how her invention differs from the prior art to a significant degree (the difference, to be significant, must be nonobvious) (3). The simple premise for market exclusivity, whether under ODA or through a patent, is that awarding a monopoly to an innovative product is generally economically justified (despite the monopoly output restrictions and the correspondingly higher prices) when the investment in innovation would be unlikely without market protection. Indeed, when enacting the ODA, Congress determined that the pharmaceutical industry needed these special economic incentives to undertake research and development for diseases affecting fewer people.

Procedures of the Orphan Drug Act

Qualification for orphan drug benefits is a two-step process: (*1*) designation and (*2*) drug approval. After drug approval the pharmaceutical manufacturer obtains a product license to sell the drug with ODA's seven-year market exclusivity.

Designation. A pharmaceutical manufacturer (sponsor) seeks orphan status designation for a drug by (*1*) certifying that the product is for a rare condition, (*2*) providing a scientific rationale for using the drug for that rare condition, and (*3*) providing supporting epidemiological data. The designation process gives the sponsor an early opportunity to interact with the FDA and learn of any significant issues that might arise later in the course of development and approval. The required scientific rationale for the orphan drug's usefulness need only be an explanation of the sponsor's hypothesis and some experimental evidence from animal model or laboratory studies. However, a sponsor may seek orphan drug designation at any point in the research and development process, even after the final stages of human clinical testing, before submitting an application for marketing approval.

ODA does not limit the number of drugs that may be designated for a particular rare disease. If a first orphan drug has obtained market approval, however, then FDA must not approve an application for designation by a second sponsor of the same drug until seven years have passed. FDA can grant orphan drug designation status to new versions of an already marketed drug, but the second, similar drug will not be approved unless FDA determines that the second applicant's is clinically superior to the already marketed drug.

Approval. While any sponsor of an orphan drug may receive the development-phase benefits of ODA, only the first manufacturer to receive full FDA drug approval receives the exclusive marketing rights for any one drug. Although the FDA is liberal in awarding orphan drug designation, the standard for approval is consistently high.

The FDA drug approval process consists of preclinical studies supporting the safety and possible efficacy of the drug in animals followed by three phases of clinical investigation. The pharmaceutical manufacturer must submit extensive scientific and medical data including chemical, pharmacological, and clinical studies. The marketing approval applications for both drugs and biologics must show that the products are "safe and effective" for their intended use.

Two major statutory exceptions to the seven-year orphan drug market exclusivity are (*1*) where the market exclusivity holder cannot provide sufficient quantities of the drug to patients who need it and (*2*) where the market exclusivity holder consents to subsequent approvals. The FDA must, before making a finding of nonavailability, give the market exclusivity holder notice and the opportunity to comment. Orphan drugs have generally spent less time in development than non-orphan drugs. This may be because an orphan drug application usually involves fewer patients and fewer clinical trials (which is inevitable for diseases with fewer sufferers). Orphan drugs are often the only available treatment for the rare disease or condition, which is why orphan drug protection is necessary to make the research and development commercially viable. Furthermore an orphan drug sponsor can seek FDA approval to allow patients access to the drug even before marketing approval, either through a Treatment IND (4) or an orphan drug open protocol (§360dd).

Orphan Drug Market Exclusivity in Biotechnology Business Strategy

One of the greatest challenges for the emerging biotechnology company and its legal counsel is to integrate its intellectual property and regulatory strategies with its financial plan. After discovering a new compound that may have beneficial medical properties, a company typically applies for a patent, since patent protection precludes other companies from selling or obtaining the patent on the compound. However, there are many compounds that have entered into the public domain by sitting on the shelf for a number of years after a first attempt to prove it effective against a particular disease. AZT was originally tried

against cancer and abandoned for that use. AZT was later given orphan drug exclusivity because at the time of its designation less than 200,000 persons in the United States were identified as having progressed from (HIV) positive status to (AIDS). It has brought its sponsor billions of dollars in revenues as an AIDS drug.

Congress had expected that ODA would be primarily used by sponsors of orphan drugs that did not qualify for product, or composition of matter, patents. Indeed, one of the most powerful and widely used applications of biotechnology is to enable the production of commercially viable quantities of otherwise rare or very difficult to produce compounds, the prior knowledge of which may preclude composition of matter patent protection. Recombinant Human Growth Hormone, the subject of the first ODA dispute, was precisely such a drug. However, orphan drug exclusivity can still be of significant value even where a compound is also the subject of a patent. The statutory patent term of 20 years from the filing date of the patent application is running during the more than seven years it takes to obtain perform the required preclinical and clinical tests required to obtain FDA approval for the average new drug application. The Drug Price Competition and Patent Term Restoration Act of 1984 (5) can partially restore a portion of the years spent in clinical development. However, the patent term remaining after FDA approval can be less than seven years, raising the value of the orphan drug period of market exclusivity, which does not begin until FDA approval.

ODA thus provides incentives for biotechnology companies in the intersecting concerns of intellectual property protection and rapid drug approval. In a biopharmaceutical development strategy, it is important to get into the marketplace sooner, with a longer patent term remaining, even to the extent of influencing, if not dictating, the choice of initial target indication for a biotechnology company's lead compound. As long as the time period necessary for gaining product approval effectively consumes a substantial portion of a drug's patent-life or biotechnology companies use genetic engineering to produce recombinant versions of known proteins, ODA will continue to be a central part of many biotechnology companies product development strategy.

Effect of the Orphan Drug Act on Drug Innovation

Pharmaceutical manufacturers have long argued that ODA procedures do not provide sufficient certainty to guide their research and development investments. One source of such uncertainty is the potential of competition among sponsors of the same or similar orphan drug. Although two sponsors of the same drug may receive orphan drug designation during the development phase, the seven-year marketing exclusivity is awarded solely to the first company to achieve market approval for the drug. Thus a company may be the first to conceive of an orphan drug program, such as interferon-β in the treatment for multiple sclerosis, invest resources into preclinical and clinical research, and then lose that investment if another company is the first to receive drug approval for that orphan drug. Applications for orphan designation, which are usually made when clinical trials are ready to begin, are kept secret until the FDA approves the designation and because companies need not apply for designation until late in the development process, a company may not realize that it is in a race with another company for orphan drug approval until it has completed its preclinical development and readied its lead compound for human trials. At that point a company may well feel pressure from its investors to reach the next milestone, with the competition continuing and the stakes growing ever higher. The initial uncertainty and the potential for an expensive, winner-take-all race can substantially undermine the incentives of the orphan drug.

Although such uncertainty undermines the goal of encouraging innovation, the uncertainty of a race to develop a product may be unavoidable. Of course, the risk of being beaten to the marketplace is quite different than the risk of being first to market only to find that winning the race provided no real victory at all. That may result when what were initially believed to be minor differences between two versions of a drug result in the orphan drug approval of both (as was the case with different forms of interferon-β, i.e., Avonex and Betaseron) for multiple sclerosis. The ODA has been successful in bring many orphan drugs to market, with more new drugs undergoing research and development despite the uncertainty of the scope of its market exclusivity. Nevertheless, the new issues raised by the FDA's recent decision to approve Avonex without closely examining its degree of identity to Betaseron may substantially impair the future effectiveness of ODA.

Small biotechnology companies have developed most orphan drugs. For small biotechnology companies, demonstrating to investors the ability to successfully develop any product is a significant corporate milestone. Also diseases affecting 200,000 Americans are "rare" under the law but may be sizable, even hugely profitable, markets for small companies. The orphan drug designation, based on U.S. disease populations, does not account for the additional potential profits on international sales, a key target for both large pharmaceutical companies and biotechnology companies. Furthermore drugs developed for a rare disorder may also work on more common diseases. For example, the biotechnology company Genentech originally developed human growth hormone (hGH) to treat children with hypopituitary dwarfism, but hGH is also useful in treating other growth deficiencies "off-label." Such extremely successful orphans, like human growth hormone and erythropoietin (EPO), are "blockbuster drugs," and both were the subject of major battles over the scope of ODA protection.

Critics of ODA, spurred by such instances of enormously profitable orphan drugs, are concerned that market exclusivity leads to higher prices that limit access to the drug. Critics also assert that profitable orphan drugs violate the spirit behind ODA, apparently believing that only minimal commercial incentives are justified. These criticisms led, in October 1990, to an amendment to ODA, passed by both houses of Congress, to permit simultaneous licensing of the same orphan drug for the same rare disease under some circumstances. President George Bush "pocket-vetoed" the amendment, however,

because he believed that the bill would weaken "the marketing incentives provided by the Act." It would be difficult to legislatively fine-tune the financial returns available under ODA to provide just enough to encourage development without going so low as to negatively impact the number of future orphan drugs. Unless the important unmet needs of orphan disease suffers can be met in some other way, the potential lucrative incentives of the ODA will continue to be essential to providing new therapies for those patients.

SCIENTIFIC CONTEXT OF THE PROBLEM OF ORPHAN DRUG MARKET EXCLUSIVITY

Since the 1985 to 1986 battle over two versions of recombinant Human Growth Hormone, FDA has been repeatedly confronted with a continuing problem with ODA, the "same or different" problem. If FDA considers the two structurally very similar drug variants to be "different," then FDA approves both drugs. When "same or different" is thus narrowly construed and an orphan drug's market exclusivity thereby narrowed, the incentives to develop such orphan drugs are substantially diminished by the risk that a second manufacturer could enter the market with a similar, almost copycat, variation. This is especially problematic for biotechnology companies, because while the research and development costs to bring a drug to market are high, it is relatively easy to make minor changes in genes and proteins that are unlikely to have pharmacological significance. On the other hand, some seemingly minor sequence changes may result in significant differences in protein activity.

The problem for FDA in defining sameness under ODA is to distinguish between significant and insignificant changes in a way that provides as much guidance and clarity as possible for companies considering the development of orphan drugs, while preserving the incentives to develop orphan drugs. A very clear, very narrow definition would substantially reduce the potential profit of an orphan drug, while a very clear, very broad definition might well deprive patients of the benefit of improved versions of a drug. Prior to the development of recombinant proteins by the biotechnology industry, when smaller, simpler chemical structures provided the basis for most drugs, two drugs were considered the same if they have the same active moiety. This strict structural approach worked well because, for small molecules produced by chemical syntheses, even slight changes in the chemical structure of the active moiety are reasonably likely to result in significant pharmacological differences. For large biological molecules, however, slight structural variations often do not result in pharmacological differences. With the growth of the biotechnology industry, an increasingly large percentage of designated orphan drugs are proteins or other large biological molecules.

All of the biotechnology orphan drugs have been proteins and glycoproteins. Proteins consist of strings of up to several hundred amino acids. It is usually possible to substitute similar amino acids for other amino acids in the protein without any noticeable change in biological activity. The problem of insignificant variation is even greater for glycoproteins. Glycoproteins are proteins with attached carbohydrate (sugars or saccharides) groups. The saccharide portion is attached to the protein enzymatically after the amino acid primary structure is produced, and depends on the particular glycosylation enzymes of the producing cell, rather than being genetically determined like the primary amino acid structure of the protein. Thus different species of animals add different saccharides to the same primary-structure protein. Even different cell lines of the same species and different cells in the same body can add different saccharides to the same primary-structure protein. Thus, when recombinant proteins are made by inserting the same gene sequence in different organisms, such as *Escherichia coli*, yeast, and Chinese hamster ovary (CHO) cells, the resulting proteins will certainly be glycosylated differently (*E. coli* and other bacteria do not glycosylate at all) and may nevertheless have the same biological activity. Whether or not such differences in glycosylation will result in differences in biological activity is not currently predictable.

For a company pursuing the very costly development of a gene-based, protein-based, or glycoprotein-based drug, the problem of predicting intellectual property protection against variant competitor drugs is a major concern.

HISTORY OF "SAME DRUG" UNDER THE ORPHAN DRUG ACT

Human Growth Hormone

Almost from the start ODA was plagued by controversy over whether two competing products were the same or were sufficiently different that the second product could be approved for the same indication. The ODA definition of a "different" drug was the issue in litigation over the orphan drug designation for human growth hormone (6). Human growth hormone is a protein secreted by the human pituitary gland that can strongly affect the growth rate and adult height of children. Administration of this hormone increases growth in children with hypopituitary dwarfism, for which indication Genentech received FDA approval. The protein has been known for some time, and was originally purified from the pituitary glands of human cadavers, a source that ultimately proved to be prone to serious contamination.

Genentech developed, with the assistance of orphan designation funding, a genetically engineered human growth hormone, and was granted market exclusivity for this hormone (Protropin). Patent protection was unavailable on the product itself because the natural protein was previously known. Eight months later FDA granted Eli Lilly & Co. orphan drug market exclusivity for a human growth hormone that differed by only a single amino acid. Genentech's Protropin had an additional amino acid that Lilly's human growth hormone (Humatrope) lacked. From a medical and clinical standpoint, there was no difference in safety and efficacy between Genentech's Protropin and Lilly's Humatrope.

Genentech first filed a citizen petition with FDA claiming that Lilly's drug was, for the purposes of ODA, the same as Protropin and therefore ineligible for marketing approval. When FDA granted Lilly approval and market

exclusivity for Humatrope, Genentech went to court to ask for judicial review of FDA's action.

The U.S. District Court for the District of Columbia held that the two human growth hormones were not the "same" drug under ODA, but also explicitly declined to provide any "universal rule for determining whether two drugs are different.... That responsibility is statutorily imposed on the FDA. Until the FDA endeavors to meet that obligation, the courts will be forced to make case-by-case determinations based on the broad polices embodied in the Act." (6, p. 306).

Erythropoietin

The protein erythropoietin was also the subject of a "same or different" orphan drug controversy, in an astronomically high-stakes race to clone the gene for the human hormone erythropoietin. Erythropoietin is a glycoprotein that stimulates red blood cell production, which is useful in the treatment of anemia. While the protein was known, no method of making commercially practicable quantities was available prior to the application of genetic engineering. Erythropoietin remains the key product in Amgen's success as a biotechnology company, accounting for over a billion dollars a year in annual revenues.

In 1989 Amgen received orphan drug marketing exclusivity for its erythropoietin product Epogen, for the treatment of the chronic anemia associated with end-stage renal disease, a "rare disease." Seeking access to the market, Chugai and its marketing partner Genetics Institute tried to obtain orphan drug status for their erythropoeitin product, Marogen, arguing that because the Marogen glycosylation pattern differed from Amgen's Epogen, Marogen should not be blocked from the market. Eventually FDA denied the Chugai/Genetics Institute application. The battle over erythropoietin continued, however, until Chugai and Genetics Institute lost a patent dispute with Amgen. Amgen gained exclusive control of the protein by virtue of their patent on the gene sequence used to produce the protein (7).

FDA Orphan Drug Regulations

The problems for orphan drug developers raised by the growth hormone and erythropoetin cases caused FDA to issue regulations under ODA to refine the definition of "same" drug by defining "different." FDA orphan drug regulations tried "to ensure both that improved therapies will always be marketable and that orphan drug exclusivity does not preclude significant improvements in treating rare diseases" (8). To be considered different under ODA, small molecule drugs cannot have the same "active moiety." Large biological drugs cannot contain the "same principal molecular structure" (9). FDA orphan drug regulations provide a "presumption of sameness" even when differences occur in protein structure. FDA believed this broader protection was consistent with congressional intent.

When it proposed its ODA regulations, FDA also considered a definition of same that would have held similar macromolecules to be the same unless the structural differences could be reasonably expected to have pharmacological significance. However, FDA realized that the regulations did not and could not specify the kinds of structural differences likely to be related to differences in pharmacological activity and therefore the regulation would create a new area of uncertainty that might also have undermined the incentives of ODA.

The final regulation, adopted in 1993, clearly rejected the notion that a macromolecule's chemical structure should be the ultimate determinant of whether or not a second orphan drug application was for the same drug as a prior approved drug. Rather, FDA decided that a second sponsor should always be able to overcome the presumption of sameness for macromolecules with the same principal structure by demonstrating significant clinical differences. "With regard to macromolecular drugs, clinical superiority by itself will render a subsequent drug different." This "clinical differences" standard was based on the principle that the market exclusivity should not create a barrier to needed patient therapies. Thus clinical superiority by itself will always lead to approval of a subsequent drug despite its substantial similarity to a prior approved orphan drug. In its ODA regulations, FDA essentially changed the focus from the underlying question of same and different to the question of under what circumstances the FDA will determine that a second drug is clinically superior to the first.

The FDA regulations define a "clinically superior" drug as one that "is shown to provide a significant therapeutic advantage over and above that provided by an approved orphan drug." Therapeutic advantage, or clinical superiority, can be show one of three ways: (*1*) greater effectiveness, (*2*) greater safety, or (*3*) in "unusual cases," demonstration that the drug makes a major contribution to patient care (10). To demonstrate greater effectiveness, the same kind of evidence is needed as that generally required to support a comparative effectiveness claim for two different drugs, that is, an improvement as assessed by the drug's "effect on a clinically meaningful endpoint in adequate and well controlled clinical trials." To support a claim of superior safety, the company seeking approval of the second product must establish that its product provides "greater safety in a substantial portion of the target populations, for example, by the elimination of an ingredient or contaminant that is associated with relatively frequent adverse effects." FDA interpretation is that even "a small demonstrated ... diminution in adverse reactions may be sufficient to allow a finding of clinical superiority" (11). Finally, a second drug can be considered "clinically superior" if it makes some other "major contribution to patient care." FDA intends this to be "a narrow category." such as the development of an oral dosage form for which there had been only a parental form (12).

The FDA's decision to rely on clinical superiority as the ultimate determinant of whether a second orphan drug was different than a prior, similar drug, left the FDA with considerable discretion to decide when head-to-head comparisons of the two drugs would be required to prove clinical superiority. One comment in response to FDA orphan drug regulations suggested that as proof of clinical superiority, FDA should always require a demonstration in

head-to-head, double-blind studies, just as for comparative efficacy claims. FDA acknowledged the value of head-to-head studies but chose nevertheless to retain discretion to make a finding of clinical superiority on the basis of other well-designed studies that it finds demonstrate a benefit on a clinically significant endpoint.

To summarize the FDA's regulations on the same and different problem, the general principal is that structurally similar drugs will be considered the same unless the second drug is shown to be clinically superior to the first. Where the issue is efficacy, the situation is quite clear: The only way to show that drug B is more effective than drug A is to directly compare their performance on an important efficacy endpoint. Where the difference is safety, or adverse effects, or "contribution to patient care," the guidelines become significantly less clear, as became obvious in the most recent case involving the variant forms of Interferon-β. If the safety of two drugs is judged on a case-by-case basis, without head-to-head data (because the second drug has a different dosage schedule or route of administration), a second drug sponsor can quite literally go to school on the first drug's data in an effort to produce fewer adverse effects.

The first sponsor of a drug for a difficult clinical indication such as multiple sclerosis would ordinarily design clinical trials to demonstrate safety and efficacy at the maximum tolerated dose. The general course of development for new compounds follows that pattern because, in most cases, a higher dose is more likely to have a statistically significant impact on the primary clinical endpoints. Following approval, clinicians often find that lower doses are effective in actual practice (13). However, a second sponsor would then have every incentive to do its studies at a lower dose or different dosing regimen, in an effort to look for efficacy under conditions that decreased adverse reactions. In defining the Orphan Drug Act's "such drug" protection in terms of clinical superiority, the FDA shifted the focus from molecular structure to clinical significance without substantially reducing the real uncertainty that Orphan Drug sponsors face.

BETASERON AND AVONEX: THE ORPHAN DRUG ACT'S LATEST CONTROVERSY

The FDA's decision that interferon-β produced in CHO cells and interferon-β produced in *E. coli* are sufficiently different to approve both for the treatment of multiple sclerosis once again raises substantial questions about ODA's market exclusivity incentive. Berlex Laboratories, Inc. had received FDA approval to market an *E. coli*-produced drug, interferon-β-1b, trade-named Betaseron, as an orphan drug for the treatment of relapsing-remitting multiple sclerosis. Berlex felt that FDA clearly erred when it decided that Biogen's drug, interferon-β-1a, trade-named Avonex, was sufficiently different to receive approval for the same patient population. Accordingly Berlex filed a suit against FDA seeking to have its determination reversed. On October 7, 1996, the U.S. District Court for the District of Columbia dismissed the lawsuit, (*Berlex Laboratories, Inc. v. FDA*), (14). The court held FDA acted lawfully when it determined that Avonex was "clinically superior" to Berlex's interferon-β-1b drug Betaseron and that "FDA's determination that Avonex is safe, pure and potent is amply supported by the record."

The question raised by the controversy over variant forms of interferon-β is not the correctness of the District Court's decision to uphold FDA. The District Court's opinion was properly based on considerations of administrative law, particularly the issue of the proper relationship between a court and an agency on a technical matter within the primary competence of the administrative agency. The Court did not attempt to decide whether or not Avonex and Betaseron are sufficiently different that Avonex is entitled to be a second entrant into the market for relapsing-remitting multiple sclerosis. Rather, the Court reviewed the regularity and sufficiency of FDA's administrative process. As a decision about administrative law and judicial deference to agency expertise, the decision in Berlex Labs is unexceptionable. What is needed is an analysis of whether FDA's regulations, in light of the interferon-β controversy, need further clarification to avoid negatively affecting the development of future orphan drugs.

Interferon-β provides an important context for examining FDA's interpretation of ODA because it is in fact a paradigmatic example of the Act's real-world effects. The potential therapeutic benefits of interferon-β had spurred both companies' efforts to produce human interferon-β by recombinant DNA procedures. While the protein is also being fought over in complex patent litigation, the ODA incentives were very likely to have been significant in the decision of the competing sponsors to attempt to prove its value in the treatment of multiple sclerosis (MS). The competing drugs each sell for about $7000 for an annual supply; both are taken by injection. Betaseron is injected under the skin every second day at a dose of 250 micrograms per injection, while Avonex is injected into muscle once a week at a dose of 30 micrograms per injection, a difference that may have played a key role in FDA's approval.

Betaseron is produced from a human interferon-β-1b gene that has been cloned and expressed in the bacterium *E. coli*. The gene and protein sequence of Betaseron varies from that of the natural molecule by one codon and its corresponding amino acid, a difference that is related to the bacterial expression of proteins. Because it is produced in bacteria, Betaseron is not glycosylated but does have an antiviral activity similar to that of native human interferon-β, thus indicating that glycosylation is probably not essential for full biological activity.

Because FDA's approval of Betaseron was based on data from Berlex' clinical trial involving patients with relapsing-remitting MS, the results pertain only to the relapsing-remitting patient group. The trial showed that injection of Betaseron every other day under the skin (subcutaneously) decreased the frequency of flare-ups and kept more patients free of flare-ups over a two-year treatment period. Adverse reactions to Betaseron included inflammation and pain at the injection site and flu-like symptoms.

About 1991 Biogen began to manufacture Avonex. On May 17, 1996, FDA approved Avonex for the treatment

of active relapsing forms of MS to slow the deterioration of physical ability and decrease the frequency of attacks. The definition of "active relapsing" MS included patients with both relapsing-remitting and relapsing-progressive forms of the disease, a more diverse population than the one studied by Berlex and approved for Betaseron.

When approving the licensing of Avonex, FDA relied on a randomized, double-blind, multicenter trial of active relapsing MS patients. In that trial, patients receiving a weekly injection of Avonex into their muscles (not subcutaneously) had a 37 percent reduction in the risk of clinically significant disability progression within the period of the study, compared with patients who received placebo. Furthermore 32 percent of placebo-treated patients had three or more exacerbations over the course of two years, compared with only 14 percent of Avonex-treated patients. Patients in the Avonex group also had a statistically significant reduction in active brain lesions seen on magnetic resonance imaging scans. The overall Avonex treatment was well-tolerated. Only 9 percent of the patients receiving the drug stopped treatment, half of them due to side effects (flu-like symptoms, muscle aches, fever, chills, and asthenia). Injection-site reactions associated with Avonex treatment occurred in only 4 percent of patients, not significantly different from patients in the placebo group. There were no reports of tissue death at the injection site, possibly because of the much lower dose, the much less frequent injections, and the intramuscular, rather than subcutaneous, injection route.

FDA's decision was based on its ODA regulations, providing that a new drug will not be considered the same as the previously approved drug if the new drug is "clinically superior" because it offers "greater safety in a substantial portion of the target populations" (15). Based upon these clinical results of the trial results, the FDA concluded that Avonex was clinically superior to Betaseron and therefore sufficiently different for ODA approval. The FDA decision rested on the substantially less frequent occurrece of injection site necrosis associated with Avonex (0 percent) compared with Betaseron (5 percent) and lower incidence of even lesser injection site reactions (4 percent of Avonex compared to 85 percent of Betaseron patients). FDA's decision was apparently based solely on the clinical data in support of the two manufacturer's applications. FDA did not attempt to determine whether the difference in the adverse effects of the two drugs was due to the dosing and administration differences or to the actual biochemical properties of the two drugs. Nor did FDA determine whether Avonex was more effective than Betaseron in treating the underlying disease.

SAME AND DIFFERENT FOR DRUGS IN OTHER FDA DETERMINATIONS

FDA Guidance Document Concerning Demonstration of Comparability of Human Biological Products

Just weeks before FDA approved Avonex, it published a Guidance Document regarding changes in the manufacturing processes of "well-characterized" biological drugs (16). FDA issued this Guidance Document to provide pharmaceutical manufacturers with more flexibility in bringing biological products to market more efficiently and expeditiously. Until that time, companies developing biotechnology drugs such as recombinant proteins or monoclonal antibodies faced a considerably more complicated regulatory pathway through FDA's Center for Biologics (CBER), than did companies developing small molecules through FDA's Center for Drugs (CDER).

Prior to the Guidance Document, CBER's approval of a biologic required two separate applications and approvals: (*1*) the approval of the product as safe and effective, through the submission of a product license application (PLA) and (*2*) the submission of an establishment license application (ELA) for approval of the manufacturing facility that produced the actual material that was used to generate the data in the PLA. The reason for the two-part, interrelated approval process for biologics was that the manufacture of biological molecules was such a highly variable process that a change from one facility to another could produce a change in the product itself. In addition the manufacturer was required to verify the product's identity for each lot produced. The FDA refused to approve applications unless clinical trials were conducted with the specific product to be licensed because any change in the manufacturing process could mean that the clinical data was in fact generated by different macromolecules and could not support the marketing approval for just one variant, or worse, an even newer and untested variant.

The Guidance Document was part of a greater policy to harmonize the requirements across the FDA for pharmaceutical manufacturers to produce "well-characterized" biotechnology drugs. The policy of more flexibility is possible because recent advances have provided the scientific ability to control the manufacture of biotechnology drugs and to determine the consistency of the identity of macromolecules produced by different processes. The Guidance Document thus recognizes this increasing technical ability of manufacturers to show that the protein produced in one facility and that was used in clinical trials is the same as the protein produced in another facility.

In allowing the clinical data for a "precursor" product be used for a later product, the FDA is may rely on the results of analytical testing, biological assays (in vitro or in vivo), assessment of pharmacokinetics and/or pharmacodynamics and toxicity in animals, in place of clinical testing, to determine whether or not the later product is the same as a prior composition. In other words, the FDA now feels that the science of genetic engineering of proteins has advanced to the point where a variety of tests can determine whether a change in manufacturing a protein yields the same protein or a different one for product identity purposes. Ironically, although the FDA relied on such principles to determine that Avonex could use clinical data generated by a prior compound, it ignored them completely in making its ODA determination that Avonex and Betaseron were different drugs based on its finding of clinical superiority.

When it proposed the FDA orphan drug regulations, FDA was concerned about determinations that involved

too much judgment and discretion from FDA officials. At that time FDA rejected any approach where the kinds of structural differences likely to be related to differences in pharmacological activity were not specified. With the Guidance Document now in place, FDA officials would appear to have more of the tools required to make judgments about when structural differences are likely to be related to differences in pharmacological activity for purposes of deciding the similarity of well-characterized biological drugs. FDA officials also have the discretion to decide whether further clinical trials are required. The Guidance Document provides a major science policy basis for a further clarification of ODA's "same or different" problem.

Generic Drugs

FDA has had a great deal of experience in determining when two drugs are identical, because that function is required during the approval of generic versions of drugs for which patent protection is expiring. Manufacturers of the generic versions need only submit abbreviated new drug approvals containing bioavailability and bioequivalence data. The process of generic drug approval is much simpler and less costly than the approval of new drugs. A manufacturer of an orphan drug that is also protected by patent exclusivity is unlikely to open itself to generic competition by seeking approval for a non-orphan indication. Instead, the manufacturer can build off-label usage for its orphan drug by distributing journal articles that support the additional uses. If the non-orphan indication were instead added to the label, a generic competitor could quickly gain approval for the non-orphan indication by submitting bioavailability and bioequivalence studies and, after approval, sell the cheaper generic off-label into the orphan market.

A generic version of a drug is chemically identical to the original pioneer drug and therefore can rely on the safety and efficacy studies done for the pioneer drug, submitting only manufacturing data and bioavailability and bioequivalence data to establish chemical identity (5). The tests for bioavailability and bioequivalence measure important aspects of the pharmacokinetics of a drug, the overall way in which a drug is processed by the body. If two drugs that are very similar in structure have nearly identical pharmacokinetics, one could presume that whatever differences in structure exist are not pharmacologically important or clinically relevant. That is because such important pharmacologically or clinically relevant differences, for example, in the two drugs' affinity for their target molecule (ligand), or their relative antigenicity, would cause marked differences in pharmacokinetics for most, if not all drugs.

CONCLUSION

The scope of orphan drug exclusivity for biotechnology drugs remains uncertain. In patent law the inquiry as to whether a second item that is more or less closely related to an earlier invention, is entitled to its own protection or infringes the earlier patentees rights, takes place under the doctrine of equivalence, the doctrine of reverse equivalents, and the doctrine of unexpected results. Each of these doctrines addresses a common, core concern: Does the second item make a substantial and independent contribution, or is it merely an attempt to profit from the exploitation of the earlier inventor's contribution? ODA, enacted in 1983, created an analogous problem in prohibiting FDA from approving the same drug for the same indication for a period of seven years. The obviously analogous problem for the administration of ODA is when a second, similar drug is the same and therefore barred from FDA approval and when a second, similar drug is sufficiently different to be entitled to its own market exclusivity.

Soon after the enactment of ODA, in the absence of any administrative interpretation of the scope of the "such drug" protection, the courts applied a much narrower interpretation to ODA protection than would have been the case under patent law. The damage that such narrow readings would do to the incentives of the Act was obvious. FDA ultimately responded with regulations that defined the scope of ODA exclusivity in a way that is much closer to the way that patent law resolves the issue. When FDA concluded that the ultimate basis for determining whether a second, similar drug could be approved would rest on its demonstration of clinical superiority over the prior drug, the agency was apparently searching for a test that essentially measured whether the second applicant made a substantial and independent contribution to the treatment of the orphan disease, or whether it simply sought to profit from the same work as the first applicant.

FDA's reliance on clinical superiority serves two purposes. One purpose is clearly furthering the interests of the patients in receiving the benefits of any significant advance in therapy. At the same time it serves a purpose much like that of the doctrine of unexpected results or the reverse doctrine of equivalents in patent law. Clinical superiority is intended by FDA to be sufficient evidence to support a determination that the difference between the two drugs is in fact significant.

Unfortunately, the Berlex Labs case reveals the shortcomings in FDA's attempt to solve the "same or different" drug problem by looking to clinical superiority. If clinical superiority can be demonstrated by the frequency of adverse effects where direct, head-to-head comparisons are not possible, it is difficult to know whether the second drug's "superiority" is due to the substantial independent contribution of the second applicant or the advantage of learning from the first applicant's data. At the same time the FDA has clearly indicated that it would not favor a policy that required head-to-head comparisons for all structurally similar second drugs. What remains is the question of when head-to-head comparisons are most necessary and whether the cases in which head-to-head comparisons should be required can be sufficiently clarified by further guidelines.

ACKNOWLEDGMENT

This entry is based in part of the article by R.A. Bohrer and J. Prince, *Harv. J. Law Technol.* **12**, 365 (1999).

BIBLIOGRAPHY

1. Orphan Drug Act, Pub. L. No. 97-414, 96 Stat. 2049 (1982), codified as amended at 21 U.S.C. §§360aa–360ee (1994), 26 U.S.C. §45C (Supp. II 1994), 42 U.S.C. §236 (1994).
2. 26 U.S.C. §28.
3. 35 U.S.C. §103.
4. 21 CFR §§312.80–88.
5. Drug Price Competition and Patent Term Restoration Act, codified at 35 U.S.C. §§156 (c)(2), 156 (d)(5)(F)(i), 21 CFR §314.55.
6. *Genentech, Inc. v. Bowen*, 676 F Supp. §§301, 306 (D.D.C. 1987).
7. *Amgen, Inc. v. Chugai Pharmaceutical Co.*, 927 F.2d 1200, 18 USPQ2d 1016 (1991) ("Amgen II") cert. den., *Genetics Inst. v. Amgen, Inc.*, 502 U.S. 856 (1991).
8. 56 Fed. Reg. 3338.
9. DHHS, Food and Drug Administration Agency, 21 CFR §316.3, Orphan Drug Reg., No. 85N-0483, RIN 0905-AB55, 57 FR 62076 (December 29, 1992).
10. 21 CFR §316.3(b)3(i–iii).
11. 57 Fed. Reg. 62076 at comment 18.
12. 56 Fed. Reg. 3338–3342, 3343.
13. *N.Y. Times*, October 12, F1 (1999).
14. *Berlex Laboratories, Inc. v. FDA*, 942 F. Supp. 19 (D.D.C. 1996).
15. 21 CFR §316.3(b)(ii).
16. FDA, *Guidance Concerning Demonstration of Comparability of Human Biological Products, Including Therapeutic Biotechnology-Derived Products*, U.S. Government Printing Office, Washington, DC, April 1996, Available at: *http://www.fda.gov/cber/guidelines.htm* and *http://www.fda.gov/cber/gdlns/comptest.txt*

(a) Bioavailability means the rate and extent to which the active ingredient or active moiety is absorbed from a drug product and becomes available at the site of action.

(e) Bioequivalence means the absence of a significant difference in the rate and extent to which the active ingredient or active moiety in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of drug action when administered at the same molar dose under similar conditions in an appropriately designed study.

ADDITIONAL READINGS

R.A. Bohrer, *Symposium issue, intellectual property and the FDA*: Forward: *Biotechnology Business Strategy: A Lawyer's Perspective*, 33 CAL. W.L. REV. 1, 6 (1996).

Brief of Amicus Curiae Chiron Corporation in Support of Respondent, Warner-Jenkinson Co., Inc. v. Hilton Davis Chemical Co., 117 S. Ct. 1040 (1997).

G.M. Cavalier, *Law Policy Int. Business* **27**(2), 447, 450 (1996).

P. Ducor, *Rutgers Comput. Tech. Law. J.* **22**, 369, 473 (1996).

V. Henry, *J. Legal Med.* **14**, 617 (1993).

Li-Hsien Rin-Laures and D. Janofsky, Note, *Harv. J. Law Technol.* **4**, 269, 271–272 (1991).

E. Knight, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **73**(2), 520, 522 (1976).

J.A. Levitt and J.V. Kelsey, *Food Drug Law J.* **48**, 525, 527 (1993).

J.T. O'Reilly, Food and Drug Administration, 2nd ed., §13.22 (1995).

E. Weck, in *From Test Tube to Patient: New Drug Development in the United States: an FDA Consumer Special Report 1988*, PLI Patents, Copyrights, Trademarks & Literary Property Course Handbook Ser. No. G4-3819, 1988.

See other entries FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE; GENE THERAPY, ETHICS, SOMATIC CELL GENE THERAPY; MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT; PATENTS AND LICENSING, ETHICS, INTERNATIONAL CONTROVERSIES; SCIENTIFIC RESEARCH, POLICY, TAX TREATMENT OF RESEARCH AND DEVELOPMENT.

G

GENE THERAPY, ETHICS, AND INTERNATIONAL PERSPECTIVES

ANDREA L. BONNICKSEN
Northern Illinois University
DeKalb, Illinois

OUTLINE

INTRODUCTION

An active debate has been ongoing about the ethics of human gene therapy since the early 1970s, when scientists announced innovations in recombinant DNA research. This debate intensified in the 1980s, when fertilization outside the human body brought closer the prospect of genetically manipulating human eggs, spermatozoa, and embryos. In the 1980s various nations set up ethics commissions to weigh ethical issues associated with embryo research, reproductive technologies, and genetic manipulations. From the recommendations of these commissions, officials in the United Kingdom, Germany, and other nations enacted laws that included provisions for gene therapy. Public attention broadened in the 1980s and 1990s, when regional and global international organizations began also to deliberate about the ethical implications of human genetic knowledge and applications. Faced with the need to achieve a consensus that crossed cultures, these organizations crafted more principle-oriented statements than did national governments, which tended to enact laws related to named reproductive techniques.

Gene therapy involves both somatic cell gene therapy (SCGT) and germ-line gene therapy (GLGT). In SCGT genes are modified and used to treat a patient without introducing changes that will be passed to subsequent generations through the germ cells. In germ-line gene therapy, genetic alterations affect the germ cells (oocytes, spermatozoa) or embryos in a way that presumably would be passed to the next generations. Research on somatic cell gene therapy, which is practiced in only a few countries and with mixed success, is generally regarded as ethical if standard safety protections and provisions for informed consent are followed. Because genetic changes are limited to the patient and are not passed to the next generation, SCGT approximates other forms of medical therapy and does not raise the same degree of ethical concern as the inheritable GLGT. One exception would be the use of SCGT on fetuses, which could inadvertently affect the fetal germ cells and become a germ-line intervention.

The still-hypothetical prospect of GLGT, in contrast, is not associated with the same degree of consensus as SCGT. According to one perspective, GLGT is a logical extension of medical research that differs by degree rather than by a clear line. If GLGT is shown to be safe through animal research and laboratory studies, according to this point of view, germ-line alterations will more efficiently eliminate genetic disease than will somatic cell technologies. Faced with this promising method of combatting human suffering, it is argued, medical personnel are obligated to develop new treatments for genetic diseases. In the process, scientific inquiry into germ-line manipulations will yield knowledge useful in its own right (1,2). According to an opposing perspective, GLGT is a labor-intensive and expensive alternative to other treatments that would be available primarily to wealthy people and would divert scarce resources from therapies that would benefit greater numbers of people. Those wary of GLGT have particular concern about the technique's inheritability, which means mistakes from this unusually risky treatment would pass from one generation to the next. Germ-line interventions are also thought to generate the potential for genetic enhancement, in which genes would be manipulated to promote socially desirable attributes that are not medically necessary. Germ-line enhancement, in turn, according to this argument, would lead to a differential evaluation of individuals, with those who have been enhanced for traits deemed socially important valued more highly than those who have not received a genetic enhancement. Germ-line interventions would also, according to the detractors of GLGT, undermine respect for human diversity, divert societal resources from ill to healthy people and impose unacceptable risks on individuals (3). For these and other reasons it has been argued there are simpler and safer ways of avoiding the birth of children with serious genetic disease, including preimplantation genetic diagnosis to identify and not transfer embryos with disease-linked genes (3).

National and international governmental and nongovernmental organizations have reached a range of conclusions about the ethical acceptability of GLGT, and the deliberations remain active as genetic knowledge advances. More consistency is apparent in the matter of germ-line enhancement, with consensus that nonmedical genetic interventions on the germ-line are not ethically acceptable (4). The proactive nature of the debate over GLGT, started well before human germ-line interventions

were remotely possible, underscores the sensitivity of the issue and the seriousness with which citizens regard the ethics of GLGT. Because a broader and more problematic array of ethical issues is associated with GLGT than with SCGT, this entry will focus on national and international responses to GLGT.

GENE THERAPY IN INDIVIDUAL NATIONS

Nations with laws in effect or in process relating to GLGT are primarily European and Anglo-American, and these nations have developed stable yet diverse approaches to assisted reproductive technologies ARTs in general and to GLGT in particular. National policies on GLGT tend to be primarily restrictive or permissive. Restrictive policies bar GLGT either explicitly by targeted provisions in national laws or implicitly by prohibitions on embryo research. Permissive policies leave a door open for GLGT through flexibly worded laws or through the absence of laws. Nations with laws governing assisted reproduction have been called "framework nations"; nations that rely on principles and rules developed by commissions and professional associations are known as "guideline nations" (5,6).

Restrictive Policies

Germany's Embryo Protection Act of 1990 (7), called the "world's most restrictive law" in reproductive medicine (8), was enacted to protect human embryos from nontherapeutic research and to draw lines for ethically contentious applications in assisted reproduction. The Act makes the misuse of human embryos a criminal act, where "misuse" is defined as buying, passing on, or acquiring a human embryo "for purposes other than preserving its development." Under this law, to cause the further development of an embryo "for purposes other than causing a pregnancy" is to commit a criminal act. An embryo is defined as a fertilized human oocyte that includes "any totipotent parts" that could develop into an individual being. According to this definition, the embryo is a human being and so are the totipotent (undifferentiated) cells in the early embryo.

The German law bars GLGT in two ways. First, it forbids preimplantation genetic diagnosis (PGD), in which a cell from the embryo of a couple at risk for passing a genetic disease to their offspring is removed and tested for the presence of the disease-linked gene. Not only is it illegal under the law to discard an affected embryo, but it is also a criminal act to discard the biopsied cell even if the embryo were found to be without the disease gene and transferred to the patient's uterus. By barring PGD, which would be a precondition of gene correction, the law precludes GLGT. Second, Germany's law explicitly makes it a criminal offense to "manipulate the genetic information of a human germ cell" or "use genetically manipulated germ cells for the purpose of fertilization."

Germany's embryo protection law was designed to exclude practices that might lead to eugenic applications. Particular features of German history, particularly the eugenic experiments and goals of Nazi Germany before and during World War II, make gene therapy a highly emotional issue in Germany. According to Mauron and Thevoz, the German approach to reproductive medicine, which is shared by other Germanic nations, reflects a cultural view that assumes negative consequences will follow from genetic applications. This view anticipates a technological imperative in which possessing a new capability is the same as using it (9).

Denmark also has a prohibitory assisted reproduction law that extends to gene therapy. Its Law No. 460 (1997) forbids the genetic modification of germ cells and fertilized eggs. The law provides that assisted conception will take place only for the purpose of "uniting a genetically unchanged (unmodified) oocyte with a genetically unchanged (unmodified) spermatozoan" (10). It also states that fertilized eggs used in therapeutic or nontherapeutic research will not be transferred "unless the fertilized oocytes are genetically unchanged." The Danish law allows PGD if there is "a known and considerable risk that the child will be affected by a serious hereditary disease." This underscores a value in Danish law that limits genetic procedures to serious medical situations only and that eschews selective embryo transfer for eugenic reasons. The Danish law was passed with significant involvement of the Danish public. It places fewer restrictions on ARTs and reveals less distrust of the procedures than the German law.

France's restrictive embryo research law forbids experimentation on the human embryo unless the embryo is to be transferred (11). It also forbids interventions that would adversely affect the embryo's "developmental capacities." Although French law forbids embryo research, it pairs PGD with prenatal diagnosis rather than with embryo research (12). Consequently, to remove cells for PGD is not regarded as a form of embryo experimentation. Physicians may perform PGD under limited conditions and if the couple is at demonstrated risk for bearing a child affected with a serious and incurable genetic disease (12). Ethical debates on PGD in France have revealed particular concern about eugenic uses, in which embryos would be selectively transferred for their socially preferred traits rather than selectively nontransferred for disease-related conditions (13). The French Senate's version of the 1994 embryo research law would have forbidden PGD altogether because of this concern, but the law's final version allowed PGD in limited circumstances (14). Despite the permissibility of PGD under French law, the technique is not practiced in France because the government only recently published the requisite decree rendering the embryo research law applicable. French law is more forthright about GLGT than about PGD. No study may be carried out if "its object is to change, or if it is likely to change, the genetic heritage of the embryo."

Other nations with restrictive GLGT policies include Austria, Switzerland, and Spain. Austria's embryo research law states that "interventions involving the germline shall not be permitted" (15). In Switzerland, a 1992 amendment to the Federal Constitution forbids "genetic manipulation" and imposes other restrictions on assisted reproduction (16). The referendum passed with a 71 percent approval vote in a national referendum and set up a committee of experts appointed by Parliament

to interpret the amendment (17). Spain's law on assisted conception and embryo research deems GLGT a "very serious offense" if the genetic manipulation is for "nontherapeutic purposes or for therapeutic purposes that are not authorized" (18). This wording leaves the door open for eventual authorized therapeutic GLGT.

Other nations limit GLGT indirectly in their restrictions on embryo research. Norway's law, for example, prohibits "research on fertilized eggs," which would include GLGT in its experimental stage (19). Such laws would not necessarily forbid therapeutic GLGT if it were effective and safe. In fact nonexperimental therapeutic GLGT might be more acceptable than PGD in some nations because it would aim to treat rather than discard embryos that have the disease gene in question.

At the subnational level, the Infertility Treatment Act of 1995 in Victoria, Australia, forbids specified techniques such as cloning and GLGT and regulates other ART procedures. It set up a Standing Review and Advisory Committee to advise the health minister on, among other things, "the use of treatment procedures or related procedures to avoid genetic abnormalities or disease" (20).

Permissive Policies

Nations are permissive regarding GLGT either by flexible laws or by default due to the absence of proscriptive laws. The United Kingdom exemplifies the former in that it has a national law that accepts embryo research and could be congenial to GLGT. British policy derives from the 1984 report of the Warnock Commission, one of the first national ethics commissions set up world wide to evaluate the ethics of ARTs. Its report recommended against specified interventions on the embryo but left GLGT open. Commission members pointed out that public anxiety over genetic therapy focused not on therapy but on "the deliberate creation of human beings with specific characteristics" premised on eugenics (21). Commission members presumed such uses would be covered by the controls recommended in the report, including a licensing body to review and approve embryo research.

Although pointedly critical of other techniques that were still speculative in 1984, the Warnock Commission was relatively permissive in the matter of gene therapy. Its report identified selective breeding but not therapy as onerous, and it invited the licensing body to determine what might not be ethical. Coming at the end of the report, the section on gene therapy was less restrictive than other sections recommending that certain activities, such as "the placing of a human embryo in the uterus of another species for gestation," should be criminal offenses. The Warnock Commission recommended that a licensing body be responsible for identifying research that "would be unlikely to be considered ethically acceptable in any circumstances and therefore would not be licensed."

From the Warnock Report emerged the Human Fertilization and Embryology Act of 1990 (HFEA) (22). This Act created a statutory licensing authority (Human Fertilization and Embryology Authority) to license laboratories working with human embryos and to monitor and set the conditions for embryo research. Under the law, researchers need a license to create, keep, or use embryos. The United Kingdom's policy on embryo research has been characterized as "highly permissive" (23). The Act lists categories of research for which licenses can and cannot be granted, and it adds a category of uncertain status into which gene therapy fits. Under the Code of Practice of the Human Fertilization and Embryology Authority, however, research licenses at present may not be granted for "altering the genetic structure of any cell" while the cell becomes part of an embryo. This also precludes gene therapy for nonmedical reasons. According to the 1992 report of the Committee on the Ethics of Gene Therapy, "in the present state of knowledge any attempt by gene modification to change human traits not associated with disease would not be acceptable" (24).

The British national policy on embryo research and assisted reproduction reflects a broader cultural optimism about medicine and the ability of humans to draw lines at unacceptable applications than appears in policies in Germanic-speaking nations. According to Mauron and Thevoz, the British policy reflects a utilitarian mode of thinking in which observers "seem as a rule more confident [than countries holding a pessimistic view] in their ethico-legal capability to promote good and prevent evil" (9). As policy the British HFEA is perhaps the most thorough on assisted reproduction in the world. It is flexible in the way it relies on a licensing system that has room for clinical judgments and gives government officials discretion in deciding categories of research for which licenses can be granted (25). It also encourages continued review of emerging techniques.

Other European nations have flexible policy climates for gene therapy by default in the absence of explicit embryo research or ART laws. These guideline nations rely on existing laws, norms, professional guidelines, and other protections to oversee innovations in ARTs. Some, such as Belgium and Canada, are in the process of formulating laws relating to bioethics. The process of developing legislation has been ongoing for over a decade in Canada. In 1989 the Canadian government appointed the Royal Commission on the New Reproductive Technologies to examine and report on the developments in and implications of the new reproductive technologies. Members of the Commission met over a period of 14 months, solicited written and oral testimony from interested citizens, and issued a two-volume report, *Proceed with Care* (3).

In its report, the Royal Commission concluded that somatic cell gene therapy was ethically acceptable provided principles governing informed consent and other protections were followed. Noting the risks of GLGT and the presence of PGD and selective nontransfer of embryos as a less risky alternative, however, it concluded that GLGT was "inconsistent with the Commission's guiding principles" and that it should be forbidden with no federal funds used to support human GLGT research. In 1995 Canada's Minister of Health formed a Discussion Group on Embryo Research. Several months later, she called for a voluntary moratorium on nine embryo research applications, including GLGT (26). Recommendations from the Royal Commission were used to draft Bill C-47 in 1996, which among other things would have forbidden

GLGT. The Bill died when the 1997 federal election was called, leaving the voluntary moratorium to limit GLGT. While it imposes no penalty, the moratorium is respected as a pronouncement of the federal government. If the Royal Commission report is an indication of public opinion, it reveals the desire of most Canadians to proscribe GLGT and enhancement interventions.

The U.S. policy on GLGT is de facto permissive except for several states with embryo research laws that might preclude GLGT. Concern about newly discovered abilities genetically to modify organisms through recombinant DNA research led the government to establish a Recombinant DNA Advisory Committee (RAC) in 1976 to review proposals from researchers seeking federal funding to conduct experiments involving genetic recombinations of nonhuman organisms. In 1984 the RAC formed a Working Group on Human Gene Therapy to consider the possibilities of human gene therapy. In the same year the Working Group disseminated for public review a Points to Consider document that posed questions about what would be necessary to establish ethical and safe therapy for SCGT (24). The final document was approved and implemented in 1985. One section excludes funding for GLGT research protocols:

> The RAC and its Subcommittee will not at present entertain proposals for germ line alterations . . . in germ line alterations, a specific attempt is made to introduce genetic changes into the germ (reproductive) cells of an individual, with the aim of changing the set of genes passed on to the individual's offspring (27).

This sole mention of GLGT in the U.S. federal regulations precludes federal funding for human germ-line research but does not forbid the research. At present, no central forum has been mobilized to weigh ethical issues associated with GLGT, although the National Bioethics Advisory Commission, situated in the Executive Branch, could be asked to issue a statement on gene therapy, as could a Gene Therapy Policy Conference under the auspices of the RAC. In 1998 the RAC announced it would review its SCGT submission procedures, which would present an opportunity to revisit the provision in those procedures that the RAC would not "at present" entertain proposals for germ-line alterations. A geneticist's proposal to conduct in utero SCGT, which could inadvertently affect fetal germ cells, was submitted to the government in 1998 to provoke a revisiting of the RAC funding guidelines (28). The absence of a proscriptive GLGT policy in the United States makes this guideline nation congenial for GLGT research relative to other countries.

In the mid-1990s Jones surveyed practitioners in twenty five nations in which in vitro fertilization is practiced, and he determined that nine had voluntary guidelines but no national regulations or laws on assisted reproduction and another four did not have either laws or clear guidelines (5). Schenker and Shushan surveyed practitioners in 16 Asian and Middle-Eastern nations and found that only Taiwan had legislation directed to ARTs (29). These surveys suggest that a guideline approach is common among nations in the area of ARTs and that a sizable number of nations lack even clear guidelines. Because policies on GLGT are often appended to ART laws, this suggests that when one takes numbers of nations into account, silence is a prevailing international perspective on GLGT policy.

Summary

A number of European and Anglo-American nations have debated the ethics of GLGT in national forums and have developed formal or informal policies related to gene therapy. Nations with GLGT policies cover a range of approaches, with some permissive and others restrictive. These nations, in which scientists engage in genetics research, make up a minority of the world's nations. Nations in which genetics research is nearly absent or peripheral generally have no laws directly related to GLGT and often have conducted no extensive debate on a national GLGT policy, although they may have set up review committees to review gene protocols (24). Although firm generalizations cannot be drawn, certain features of national responses to GLGT identify directions for future ethical inquiry and policy formation.

First, among nations in which gene therapy debates have ensued, it is generally the case that SCGT is regarded as ethically acceptable, provided safety and informed consent procedures are followed. Germ-line interventions for medical reasons, on the other hand, are generally proscribed but not necessarily permanently. Genetic interventions for enhancement, nonmedical, reasons are nearly universally held to be ethically unacceptable, although this may be by convention rather than through explicit mention in the law.

Second, national positions on GLGT reflect stable ideological perspectives related to the political culture and history of individual nations. Germany's restrictive embryo research law reflects fears of eugenic applications traced to reactions against the eugenic premises of the Nazi ideology during the World War II period and a notion of justice designed to protect vulnerable groups that include embryos and generations of humans to come. There is also in Germany a distrust of genetic alterations in general that extends to the genetic modification of plants and animals. The United Kingdom's policy, in contrast, reflects a greater trust in medical genetics. Not having directly experienced a national ideology that embraced the political strategy of selective breeding to bring in those with socially desirable traits and to select out those with traits deemed unworthy, British researchers conduct genetics research in a different milieu from that in Germany. In fact the United Kingdom, as the home of two significant events in the last quarter of the twentieth century — the first successful in vitro fertilization and the first cloning of a mammal, a lamb, from the body cell of an adult sheep — reveals an ethic of discovery and innovation in assisted conception. Distrust of plant and animal genetic engineering is, however, pronounced.

Third, legislation in European nations tends to emphasize the bioethical principle of justice. According to Knoppers and Chadwick, the ethical debate on germ-line interventions gains momentum from a desire for justice toward future generations, where future generations are deemed to be vulnerable populations. The debate also

stresses the importance of the principles of autonomy, privacy, quality, and equity (16). A European theme that regards human genes as a common heritage (see the discussion below) also orients thinking toward principles of distributive justice more than in the United States and the United Kingdom, where protection of the autonomous choices of patients is a primary value. Within Europe, however, GLGT policies differ according to national decision-making styles. According to Byk, bioethical decision-making processes are in place in Norway, Denmark, Sweden, and the United Kingdom as a result of shared moral values and a strong constitutional orientation (30). These nations also share a pragmatic style of thinking that tolerates the lack of a clear resolution for certain ethical issues. Other nations, such as Italy and Belgium, lack the same decision-making process for bioethics, argues Byk, and have fundamental divisions that can make consensus elusive.

Fourth, ethical debates and policies about GLGT in many European nations are anticipatory. On the one hand, this protects individual rights by addressing potential injuries before the science is imminent and before violations of human dignity occur, and it amounts to a preventive rather than curative ethics (31). On the other hand, restrictive anticipatory policies may cut off debate prematurely and before a context develops for reasoned, experienced-based inquiry (32). Moreover they may quickly become dated or difficult to interpret as technologies change. For example, GLGT policies revolve around definitions that consider GLGT to be the splicing of genes into nuclear DNA. Recently, however, researchers have suggested manipulations that might modify the definition of a germ-line intervention. Some genetic diseases are linked to mutations in mitochondrial DNA, a form of DNA found in the cell's cytoplasm. These diseases might be circumvented by transferring the nucleus from the egg of a woman with a mitochondrial disease to a donor egg from which the nucleus has been removed and discarded. This would result in an inheritable change in which the new mitochondrial DNA would appear in each cell of the child to be, including egg cells, which would make it a germ-line innovation. National laws cover interventions on nuclear DNA, but it is not clear in any nation except Germany whether they would apply to a procedure in which cytoplasms and hence mitochondrial DNA, is substituted (33). This suggests that anticipatory laws can produce formulas that are not sufficiently supple to oversee unexpected innovations in ARTs. A similar problem will conceivably arise when policy makers try to figure out what is meant by nonmedical (or enhancement) germ-line interventions. Nations with ART licensing procedures can arguably respond to definitional ambiguities more easily than those without such mechanisms.

Fifth, ethical debate on GLGT across nations is intertwined with views about the genetic modifications of plants and animals, which indicates the powerful symbolic content of genetics and genetic engineering. It is not coincidental that citizens in Germany and Switzerland, nations that restrict human genetic manipulations, also are distrustful of genetic modifications of plants and animals. The United States, in contrast, holds a relatively more trustful view of genetic interventions in humans as well as in agriculture and commerce, and many of the world's most active biotechnology companies are in the United States.

GENE THERAPY AND INTERNATIONAL ORGANIZATIONS

Variations in regulations across countries, especially within the close borders of European nations, generate concern that a "procreative tourism", will arise in which patients will travel from one nation to the next in order to circumvent restrictive ART laws in their home countries (34). A persistent sense that interventions on the human germ line are so important that protections must span national borders has caused policy makers to identify principles that ought to prevail in all nations. Efforts to identify cross-national guidelines for genetics research have taken place regionally in Europe and globally in governmental and nongovernmental decision-making bodies.

European Regional Organizations

Europeans, primarily through the Council of Europe, have taken a leadership role in identifying common principles to span variations in laws within European nations and in publishing the results of their deliberations (35). Europeans are well poised to do this in that they have been actively involved in protecting human rights in medical and scientific settings. From this tradition, as seen in the Nuremberg Code for medical research and the European Convention on Human Rights, emerged recommendations protecting the rights of dying persons and the mentally ill (36). In addition, in Europe, an historical collectivist ethic has undergirded the idea of the human genome as a collective heritage.

The Council of Europe has actively sought principles to govern new reproductive and genetic technologies. Established in 1949 to promote cooperation among nations, the Council currently has 40 member states. The Council's "leading conscience" is its Parliamentary Assembly, which is made up of appointed representatives of national parliaments of member states and geared to formulating resolutions and recommendations (36). In 1982, in response to widespread concerns about genetic engineering and inspired by a parliamentary hearing on genetic engineering and human rights held in Copenhagen in 1981, the Parliamentary Assembly delivered Recommendation 934 to the Committee of Ministers (37). This early public policy statement on genetics was positive in tone, and its scope was broad in defining genetic engineering as "artificially recombining genetic material from living organisms," which covered plants as well as humans.

Recommendation 934 affirmed the promise and responsibilities of scientific inquiry in general and gene therapy in particular by noting that gene therapy "holds great promise for the treatment and eradication of certain diseases which are genetically transmitted." Members of the Parliamentary Assembly developed Recommendation 934

with an eye to Articles 2 and 3 of the European Convention on Human Rights, which dealt with rights to life and human dignity. The members interpreted these articles to "imply the right to inherit a genetic pattern which has not been artificially changed." This right need not be absolute; Recommendation 934 stated that gene therapy on embryos, fetuses, and minors should not be conducted without the "free and informed consent of the parent(s) or legal guardian(s)," which leaves the door open for therapy on embryos. In its Recommendation 934, the Assembly encouraged a European agreement to be drawn up regarding genetic engineering so that national legislation could be aligned accordingly, and efforts could be made to "work towards similar agreements at world level." It called for an "explicit recognition in the European Convention on Human Rights of a right to a genetic inheritance which has not been artificially interfered with, except in accordance with certain principles which are recognised as being fully compatible with respect for human rights (as, for example, in the field of therapeutic applications)." Such a statement would leave the door open to GLGT if the intervention conformed with therapeutic and research principles and was not conducted for nontherapeutic ends. In 1989 the Parliamentary Assembly issued Recommendation 1100, which would have limited embryo research and would have forbidden "any form of therapy on the human germinal line" (38). The Recommendation did not pass the Committee of Ministers, however, so its broad proscription of any form of GLGT was not officially approved (39).

In 1991 the Council of Europe acted on the Parliamentary Assembly's recommendation to recognize a right to a genetic inheritance that has not been artificially changed and to seek a transnational harmonization of genetic principles. In that year the Council convened a bioethics convention. The convention was designed to create a legally binding instrument based on a "limited set of general principles founded on human rights values" (36). It was a prospective agreement designed to guide national legislation in a document flexible enough to reflect differing cultures and legal systems. The guiding assumption was that the convention would reflect values that would apply across different cultures and legal systems. It also was designed to accommodate rapid technological change and to address public fears about genetics (39). In 1996 the convention, known as the Convention for the Protection of Human Rights and Dignity with Regard to Biology and Medicine ("Bioethics Convention"), was endorsed by the Parliamentary Assembly. It was opened for signature of the Council's member states in 1997. Nations signing the convention were legally obligated to bring their own laws into conformity with the principles, unless they signed with a reservation that they would not conform with individual articles covered by their existing laws (40).

The Bioethics Convention uses human rights themes and is based on protections of human dignity inherent in the Universal Declaration of Human Rights (41). The Preamble to the Bioethics Convention reiterates the idea that individuals have the right to a gene line that has not been deliberately altered; "an intervention seeking to modify the human genome may be undertaken for preventive, diagnostic, or therapeutic purposes and only if its aim is not to introduce any modification in the genome of any descendants" (42). This provision bars GLGT if the genomes of descendants would be modified, which effectively precludes the most commonly envisioned forms of GLGT. It is more restrictive than Recommendation 934 because it adds a stipulation against introducing inheritable changes. It leaves the door open to noninheritable GLGT, which scientists suggest may be possible if, for example, genetic changes to the germ cell could be introduced via a dispensable extra chromosome that would be engineered not to pass to the next generation (43).

Twenty-eight of the 40 member states of the Council of Europe signed the Bioethics Convention within three years of its completion, which obligated them to harmonize their national laws with the principles of the convention. Fourteen others did not indicate whether they would sign. Germany was reluctant because the Convention was less restrictive than its own practices (e.g., it allowed embryo research) (44), and the United Kingdom was reluctant because certain provisions were more restrictive than its own legislation. Nonmember observer states that participated in the convention's development (Australia, Canada, Japan, the Holy See, and, the United States) were also invited to sign but none did (40).

European discussions reveal a wariness of GLGT and an intense distrust of enhancement interventions. As early as the 1980s the European Community set up a crossdisciplinary working party to examine ethical issues associated with the new reproductive technologies. With Jonathan Glover as chair, the group issued the Glover Report in 1989, which concluded that GLGT should be rejected for humans because of its "serious risks" and also because of the ethical concerns associated with changing the genomes of the patient's offspring. The authors argued that given the current state of knowledge, germ-line interventions should be rejected and that enhancement genetic engineering is unethical "at least until policies have been worked out to cope with the huge problems it raises" (45). In 1988 the European Medical Research Councils, representing the Medical Research Councils of Austria, Denmark, Finland, France, the Netherlands, Norway, Spain, Sweden, Switzerland, the United Kingdom, and West Germany, concluded that genetic enhancement "should not be contemplated" in light of the ethical problems it would raise (31). In 1989 the Council of Europe's Ad hoc Committee of Experts on Bioethics published a report on assisted conception and embryo research. Its nonbinding recommendations intimated that procedures on embryos were acceptable only to benefit or observe the embryo, but that states could allow other "investigative and experimental procedures" for a "preventive, diagnostic or therapeutic purpose for grave diseases of embryos." While this left the door open for GLGT, it proscribed enhancement interventions: assisted reproduction should not be used "for obtaining particular characteristics in the future child" (46).

A recurring concept in European organizations is the notion of a shared genetic heritage or genetic patrimony (*patrimoine genetique*). This notion, which is not generally embraced in Anglo-American nations, extends beyond the

genetic endowment of any one individual to regard the human germ-line as a shared, collective resource. Under this concept the genome represents "the collective assets of a community (or of mankind)" that is "both irreplaceable and of enduring worth, and therefore subject to specific forms of social protections" (9). As defined by Agius, "the collective human gene pool knows no national or temporal boundary, but is the biological heritage of the entire human species." Originating with discussions of the law of the sea in 1967, the concept of a common heritage embraces access to common resources. A common heritage cannot be owned; it is instead passed from one generation to the next as a set of openly accessible goods. As a consequence no generation has the right to use GLGT to "alter the genetic constitution of the human species" (47). Genes are part of the common heritage because they are inherited; moreover a sharing of genes across the generations unites humans as a species. The knowledge that humans share a common genetic structure helps humans see themselves as a "collectivity of rights and responsibilities." Thus humans as a species have the right "to inherit a healthy and diversified genetic heritage" that has not been appropriated by patenting or other actions that benefit the individual (48). Genetic patrimony implies an international sharing of genes as a resource and an international policy that involves broad participation on behalf of all members of the human species (48).

This concept brings GLGT into a framework of human rights and responsibilities. Germ-line interventions are more than mere science or medicine; they touch upon something with a nearly mystical aura — a collective genetic heritage. While the intensity of the concept varies among countries, at root it embraces notions that genetic interventions are societal rather than individual and that they pose harms as well as benefits. It also includes the idea, influenced by the German philosopher Hans Jonas, that to possess technological knowledge is eventually to use it in a kind of technological imperative (9).

Organizations and Associations Beyond Europe

International commissions and organizations beyond Europe have also weighed the ethics of GLGT. The Council of International Organizations of Medical Sciences (CIOMS), which holds international conferences to discuss medical and scientific topics, convened a meeting in 1990 on Genetics, Ethics and Human Values that was cosponsored by the World Health Organization (WHO) and the United Nations Educational, Scientific, and Cultural Organization (UNESCO). The 102 participants from 24 nations concluded in the CIOMS Declaration of Inuyama that GLGT is unique because of the possibility for "permanent genetic change," but that it "might be the only means of treating certain conditions," and "must not be prematurely foreclosed," (24). The statement limited GLGT to medical conditions that "cause significant disability," and it did not condone interventions to "enhance or suppress cosmetic, behavioural or cognitive characteristics unrelated to any recognized human disease."

The World Medical Association (WMA), founded in 1947 to promote high international standards in medicine and health care, advised physicians in 1987 to respect international ethical codes when engaging in research on genetic diagnosis and treatment. In response to what it perceived to be "uncompromising opposition" to the human genome project, the WMA in 1992 issued a Declaration on the Human Genome Project that urged a rational assessment of ethics using the same guiding principles to evaluate risks and benefits as used for any diagnostic or therapeutic innovations — respect for the patient as a human being and for his or her autonomy and privacy (49).

The Human Genome Organization (HUGO) was established in 1989 to promote cooperation among scientists engaged in human genome research and to encourage public debate on the multiple implications of genome projects (50). This independent body now has over 1000 members from over 50 nations. The Ethical, Legal and Social Issues Committee of HUGO, chaired by Bartha M. Knoppers, developed and released in 1995 an aspirational Statement on the Principled Conduct of Genetics Research. This Statement issued ten recommendations in response to concerns about genome research, none of which called for a ban on GLGT. It also identified four principles as essential in genetics research: regarding the human genome as "part of the common heritage of humanity"; respecting "international norms of human rights"; recognizing the "values, traditions, culture, and integrity of participants"; and upholding "human dignity and freedom" (51). The reiteration of the concept of the common heritage reflects in the HUGO document a European voice.

In early 1999 the Executive Board of the WHO also approved draft bioethics guidelines that repeat the idea that the genome is the "common heritage of humankind." The guidelines do not preclude GLGT, bar enhancement, or cover embryo research, but they do warn against "hurried and premature legislation in the rapidly evolving field of genetics" (52). The nonbinding draft guidelines were presented to the WHO General Assembly in mid-1999.

An aspirational document of particular importance is the Universal Declaration on the Human Genome and Human Rights, signed in 1997 by the 186 member states of UNESCO (53). In light of expanding genetic knowledge occasioned by the human genome project and increasingly voiced concerns about the impact of genetic knowledge on human rights, UNESCO charged its International Bioethics Committee (IBC), chaired by Nicole Lenoir, a member of the French Constitutional Council, to consider an international framework for protecting the human genome (50). IBC proposed a nonbinding, aspirational declaration rather than a binding treaty. The UNESCO General Conference accepted this proposal and requested a draft Declaration for its next meeting in 1997. It was hoped the Declaration could be ready for signatures of its member states to coincide with the 50th anniversary of the United Nations Declaration of Human Rights in 1998. Over the next several years, IBC broadly consulted parties and documents from around the world. IBC was an independent body made up of 55 individuals from 40 nations who were selected for their "competence and personal attributes," and not to represent individual nations. Overall, committee members drafted and examined eight versions of the Declaration before handing its final version to the General Conference in mid-1997 (50).

IBC faced numerous challenges in its deliberations. Dealing with a scientific area marked by rapid change, it aimed to create an enduring document that would not be dated within a short time. It also worked for a document that would be given to the nearly 200 nations of the world, each of which had different values, cultures, and national perspectives. In only a few of these nations were scientists engaged in genetics research, yet it was an aim of the document that research holding great promise for humans should not unduly be restricted (51).

Lenoir regarded UNESCO as an "ideal forum" for producing an international document linking genetics and human rights. For one thing, UNESCO is the only agency of the United Nations with jurisdiction over science. For another, its constitution, which underscores the ideals of "dignity, equality, and respect for human rights," links it with ethical ideals. The UNESCO Declaration was meant to balance individual rights with the promise generated for all humans through genetic advances. It was, according to Lenoir, meant to be a reminder of the solidarity of richer and poorer nations in enjoying the benefits of science. As a document embracing universal human rights, the Declaration was based on "the unity of mankind and the equal dignity of individuals, as upheld by the principle of the universality of human rights" (50).

The IBC draft Declaration was given to the General Conference in the summer of 1997, where it took its ninth and final draft. Differing national perspectives were expressed in these deliberations. Germany wanted more direction in the document, with lists of technologies to be banned with germ-line interventions as a priority. The United States, which acted as an observer with three representatives participating in the IBC meetings, favored a more open framework for genome research (54). The United Kingdom and Singapore also acted as observers even though they, along with the United States, were not UNESCO members.

By late 1997 the Declaration was presented for signature, whereupon it was signed by UNESCO's 186 member states. The document does not explicitly forbid GLGT or any other specific technique except cloning. It does, however, echo the European sentiment that the genome is a shared resource. The final wording of Article 1 states the following:

> The human genome underlies the fundamental unity of all members of the human family, as well as the recognition of their inherent dignity and diversity. In a symbolic sense, it is the heritage of humanity.

Other articles establish principles for the ethical conduct of genetic inquiry. For example, the document regards genetics research as a way to "offer relief from suffering and improve the health of individuals and humankind as a whole." This embraces the principle of justice in that it places a value on the equitable sharing of the benefits of genetic medicine among all individuals and not just the wealthy few.

The Declaration also places a high value on individual dignity and respect for diversity. Article 2, for example, states that individual dignity "makes it imperative not to reduce individuals to their genetic characteristics and to respect their uniqueness and diversity." The status placed on diversity echoes the preamble of UNESCO's constitution (41) and parallels other United Nations documents such as the Convention of Biological Diversity, which respects cross-species diversity. The concern about genetics as a common heritage, although modified in the final version, echoes UNESCO's tradition of protecting other entities of universal interest such as the moon and cultural achievements. It also reveals the influence of the European tradition (55).

Summary

International documents related to gene therapy generally rely more on principles than technique-driven rules. This strategy is designed to help ensure consensus among participants who represent a variety of political and cultural systems and to allow the documents to withstand the test of time better than were they geared to particular techniques. The documents present a general framework around which nations can harmonize their legislation; they do not impose rules on the nations, which are left to devise their own regulations (56). International statements also differ from national laws in that they avoid reaching conclusions on the moral status of the embryo or making detailed rules about embryo research. This contrasts with national laws that are intimately tied with presumptions about the embryo's moral status and the ethics of embryo research. The documents also provide protection to citizens in signatory nations that do not have germaine policies and are unlikely to craft their own legislation.

Differences exist between the regional European-based and internationally based positions on gene therapy, as seen by contrasting provisions in the European Bioethics Convention and the worldwide UNESCO Declaration. In the main, the UNESCO Declaration, which is oriented solely to genetics issues, has a broader reach that recognizes the global import of genetic interventions, addresses genetic reductionism, and warns of the dangers of eugenics by acknowledging the range of factors that contribute to individual personalities (41). The Bioethics Convention, in contrast, covers bioethical issues other than genetics and has a more procedure-oriented focus on genetics. Byk suggests a similar regional approach would be suitable for Asia in that a regional approach, which harmonizes rules when fundamental human rights are at stake, more easily accommodates regional cultural factors than a global approach (30).

One similarity among international documents is the concept of the genome as the common heritage of humanity, which arises in the Bioethics Convention, the UNESCO Declaration, HUGO's ethics statement and policies in Switzerland, Canada, Germany, France, Australia, and other countries (47). This reminds nations of the shared benefits as well as costs that can emerge from medical genetics.

THE UNITED STATES AND INTERNATIONAL PERSPECTIVES

International perspectives set a broad context for examining GLGT, but it is not clear how influential those

perspectives will be for U.S. policy. First, the United States has no specific GLGT policy and does not appear to be in the process of developing one. Second, while the United States was involved in drafting major international documents, it has not ratified them. It played an active role in developing the UNESCO Declaration, but it did not ratify the document even though it could have signed it as a nonmember. It could also have ratified the Bioethics Convention by virtue of its status as an observer state, but it did not (40). As Annas points out, the United States ratified the United Nations International Covenant on Civil and Political Rights but not the International Covenant on Economic, Social, and Cultural Rights, possibly because the latter was perceived to be inconsistent with the emphasis on private property in the U.S. economic system (57). The United States also failed to support the negotiated version of the 1998 Biosafety Protocol, which would have been the first global treaty to regulate genetically altered products, in part to protect its economic interests. The Bioethics Convention, which limits GLGT, may be inconsistent with U.S. economic interests even though GLGT is hypothetical and of uncertain ethical acceptability, and it may not turn out to be economically feasible.

Third, the U.S. and European perspectives differ in significant ways. The European perspective tends to be collectivist and infused with a worldview ethic. Justice, including respect for future generations, is a primary concern. The U.S. perspective, on the other hand, is more individualistic, with the autonomous choices of potential parents and freedom of scientific inquiries as primary values. Thus the idea of the genome as a common heritage of humankind is less compatible with American beliefs than with beliefs in Europe. Scholars on both sides of the Atlantic have voiced criticism of the concept. Juengst, for example, questions its validity as a scientific concept when he argues that the genome is not a thread connecting all humans from one generation to the next. On the contrary, "each organism's germ-line terminates in its gametes" and the embryo begins a new germ line that is the product of the mother's and father's germ lines, which now end (47). According to Juengst, the idea of a common gene pool is uncomfortably close to old and inaccurate ideas about blood lines. Moreover genetic modification is continually ongoing in the form of reproductive decisions, which opens to question why GLGT is singled out as particularly threatening. Cook-Deegan similarly asks why the genome should be considered a public resource in societies in which reproductive privacy is highly valued (58). In Europe, Sass recommends viewing the genetic heritage as an individual as well as familial concept characterized by an "ever-modifying and mutating pool of human DNA," (55) and Mauron and Thevoz question whether the human genome is a "resource" (9).

Although the genome as a human resource has not informed U.S. policy making, the interest in the ethical implications of germ-line interventions provides international forums for discussing GLGT issues. Reports such as those from the Royal Commission on the New Reproductive Technologies in Canada and the Warnock Commission in the United Kingdom and policies such as the UNESCO Declaration enlighten citizens and policy makers about emerging international norms. They also serve as an eloquent reminder that the repercussions of GLGT are global. This reminds policy makers of the need to craft policies aimed at avoiding potential injuries to future generations. Perhaps more important, if GLGT proceeds in the United States and other industrial nations and is shown to be safe and effective, this reminds us of the need to ensure that the benefits of germ-line alterations are shared equitably among citizens in nations of differing economic levels. If the genome is a resource, fellow humans deserve access to beneficial applications as well as protections from preventable harm.

CONCLUSION

In 1991 Walters examined 20 policy statements on human gene therapy issued by governmental and nongovernmental bodies worldwide between 1980 and 1990 (4). He found that most, but not all, opposed GLGT and none supported germ-line enhancement. Differing conclusions about the ethical acceptability of GLGT continue to permeate international perspectives today, and they have produced an array of policies at the national and international levels. Individual nations can be thought of having either frameworks or guidelines, where the former means having laws directly monitoring assisted reproduction and gene therapy and the latter means reliance on voluntary guidelines and tangential policies to protect human research subjects.

International policies echo this framework/guideline distinction as well. A framework international policy that is legally binding on its signatories is the Bioethics Convention of the Council of Europe, which states that interventions on the genome may be undertaken only if the aim is not to produce a modification in descendants (42). The UNESCO Universal Declaration on the Human Genome and Human Right, in contrast, in an aspirational, nonlegally binding guideline that sets forth principles for the just and equitable use of genetic knowledge (53).

Apart from a few notable national laws and the Bioethics Convention, which has not been signed by key nations in which scientists are engaged in innovative genetics research, the overall perspective on GLGT can be said to be cautionary but not absolutely prohibitory. It is not uncommon for warnings about GLGT to be paired with caveats about the state of knowledge at this time. Such utilitarian warnings are based on the assumption that the risks of correcting genes at the germinal level now far outweigh any benefits, given the primitive knowledge about human GLGT. Were animal and human laboratory experiments to reveal the requisite heightened level of safety needed to proceed, however, the cost–benefit calculation would be open to revision. Risk-based reasons for not proceeding with gene therapy at this time, in other words, leave the door open to eventual application.

A differing perspective regards the intentional manipulation of the germ line as fundamentally illicit because it would interfere with a resource that is shared across humanity. Still the idea of the genome as a collective human heritage is not autonomatically paired with a prohibition on GLGT. Instead, the notion of a collective

heritage, which has been moderated in meaning from the earlier, more rigid concept of genetic patrimony, can serve to guide eventual research. The Ethics Committee of HUGO, for example, used the collective heritage concept to guide scientists but not prohibit them from human gene therapy (51).

Many facets of human gene therapy remain to be explored in the international arena. For example, it is taken for granted that enhancement genetic interventions ought not be pursued, yet policies are bereft of guidance about what a nonmedical germ-line intervention might be and where and how a line would be drawn between medical and nonmedical applications. In addition definitions of GLGT are being challenged by research advances that suggest inheritable manipulations can occur in ways other than through nuclear DNA splicing or that germ-line interventions need not be inherited.

Ethical inquiry into GLGT remains fertile and provocative. As organizations, agencies, associations, and governments continue the discussion, the UNESCO Universal Declaration provides a backdrop, signed by most of the world's nations, that sets forth durable principles to advise the judicious use of genetic knowledge. International attitudes have modulated in the last quarter of the twentieth century to embrace concerns about whether benefits can be equitably shared as well as whether individual and societal harms can be prevented.

BIBLIOGRAPHY

1. E.T. Juengst, *J. Med. Philos.* **16**(6), 587–592 (1991).
2. N.A. Wivel and L. Walters, *Science* **262**, 533–538 (1993).
3. Royal Commission on New Reproductive Technologies, *Proceed With Care: Final Report of the Royal Commission on New Reproductive Technologies*, Minister of Government Services of Canada, Ottawa, 1993.
4. L. Walters, *Hum. Gene Ther.* **2**(2), 115–122 (1991).
5. H.W. Jones, Jr., *11th World Congr. In vitro Fert. Alternative Assist. Reprod.*, Vienna, 1995, pp. 121–125.
6. R. Scott, *Hum. Reprod.* **12**(11), 2342–2343 (1997).
7. A German Embryo Protection Act (October 24, 1990), *Gesetz zum Schutz von Embryonen*; reprinted in *Hum. Reprod.* **6**(4), 605–606 (1991).
8. H.M. Beier and J.O. Beckman, *Hum. Reprod.* **6**(4), 607–608 (1991).
9. A. Mauron and J.-M. Thevoz, *J. Med. Philos.* **16**(6), 649–666 (1991).
10. Denmark Law No. 460 of 10 June 1997 on artificial fertilization in connection with medical treatment, diagnosis, and research, etc. (Lortindende, 1997, Part A, 11 June 1997, No. 84, pp. 2195–2198), Den. 97.25; reprinted in *Int. Dig. Health Legis.* **48**(3/4), 231–233 (1997).
11. Decree No. 97-613 of 27 May 1997, on studies conducted on human embryos in vitro, and amending the Public Health Code (Second Part: Decrees made after consulting the Conseil d'Etat). (Journal officiel de la Republique Francaise, Lois et Decrets, 1 June 1997, No. 126, pp. 8623–8624), Fr.97.60; reprinted in *Int. Dig. Health Legis.* **48**(3/4), 355–357 (1997).
12. S. Viville and I. Nisand, *Hum. Reprod.* **12**(11), 2341–2345 (1997).
13. J. Testart and B. Sele, *Hum. Reprod.* **10**(12), 3086–3390 (1995).
14. *Science* **264**, 1655 (1994).
15. Austria, Federal Law of 1992 (Serial No. 275) regulating medically assisted procreation (the Reproductive Medicine Law), and amending the General Civil Code, the Marriage Law, and the Rules of Jurisdiction. Date of entry into force: 1 July 1992. (Bundesgesetzlatt fur die Republik Osterreich, 4 June 1992, No. 105, pp. 1299–1304) Austr. 93.1; reprinted in *Int. Dig. Health Legis.* **44**(2), 247–248 (1993).
16. B.M. Knoppers and R. Chadwick, *Science* **265**, 2035–2036 (1994).
17. A. McGregor, *Lancet* **339**, 1345 (1992).
18. Law No. 35/1988 of 22 November 1988 on Assisted Reproduction Procedures. Boletin Oficial del Estado: 24 November 1988, No. 282, pp. 33373–33378; reprinted in *Int. Dig. Health Legis.* **38**(4), 782–785 (1989).
19. Law No. 68 of 12 June 1987 on artificial fertilization. Norsk Lovtidend, Section I, 26 June 1987, No. 13, pp. 502–503, Norway; reprinted in *Int. Dig. Health Legis.* **38**(4), 782–785 (1987).
20. Australia (Victoria). The Infertility Treatment Act 1995. No. 63 of 1995. Date of Assent: 27 June 1995 (151 pp.), Ausl. (Vic) 97.1; reprinted in *Int. Dig. Health Legis.* **48**(1), 24–33 (1997).
21. M. Warnock, *A Question of Life: The Warnock Report on Human Fertilization and Embryology*, Basil Blackwell, Oxford, 1985.
22. Human Fertilization and Embryology Act 1990, reprinted in *Int. Dig. Health Legis.* **42**(1), 69–85 (1991).
23. D. Morgan and L. Nielsen, *J. Law, Med. Ethics* **21**, 30–42 (1993).
24. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997.
25. S. McCarthy, Presentation at Conference, Approaches to ART. Oversight: What's Best in the U.S.? Sponsored by The National Advisory Board on Ethics in Reproduction (NABER), RESOLVE and the U.S. Centers for Disease Control and Prevention, Washington, DC, 1998.
26. P. Baird, *Hum. Reprod.* **12**(11), 2343–2345 (1997).
27. NIH Guidelines for Research Involving Recombinant Molecules (NIH Guidelines), *Fed. Regist.* **55**, 7438–7448 (1990).
28. J. Couzin, *Science* **282**, 27 (1998).
29. J.G. Schenker and A. Shushan, *Hum. Reprod.* **11**(4), 908–911 (1996).
30. J.C. Byk, *Eubois J. Asian Int. Bioethics* **5**, 59–61 (1995).
31. B.A. Brody, *The Ethics of Biomedical Research: An International Perspective*, Oxford University Press, New York, 1998.
32. J.C. Fletcher, *Politics Life Sci.* **13**(2), 225–227 (1994).
33. A.L. Bonnicksen, *Politics Life Sci.* **17**(1), 3–10 (1998).
34. B.M. Knoppers and S. LeBris, *Am. J. Law Med.* **27**(4), 329–361 (1991).
35. European Commission, *Ethical, Legal and Social Aspects of the Life Sciences and Technology Programmes of the Fourth Framework Programme*, EUR 17767 EN, European Commission, 1998.
36. J.C. Byk, in D.R.J. Macer, ed., *Bioethics for the People by the People*, Eubois Ethics Institute, 1994, pp. 68–73.
37. Document from the Parliamentary Assembly of the Council of Europe, 1982. Thirty-Third Ordinary Session Recommendation 934 (1982) on genetic engineering; reprinted in *Hum. Gene Ther.* **2**(4), 327–328 (1991).

38. Parliamentary Assembly of Council of Europe Adopts Recommendation on Use of Human Embryos and Foetuses for Research Purposes; reprinted in *Int. Dig. Health Legis.* **49**(2), 485–491 (1989).
39. M.A.M. de Wachter, *Hastings Cent. Rep.* **27**(1), 13–23 (1997).
40. F.W. Dommel, Jr. and D. Alexander, *Kennedy Inst. Ethics J.* **7**(3), 259–276 (1997).
41. M.F. Munayyer, *Am. Univ. J. Int. Law Policy* **12**, 687–731 (1997).
42. Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine, Oviedo, 04.IV.1997. Available at: *http://www.coe.fr/eng/legaltext/160e.htm*
43. *Engineering the Human Germline Symposium*, Summary Report, University of California at Los Angeles (UCLA), 1998.
44. A. Abbott, *Nature (London)* **389**, 660 (1997).
45. J. Glover, *Ethics of New Reproductive Technologies: The Glover Report to the European Commission*, Northern Illinois University Press, DeKalb, 1989.
46. Council of Europe Publishes Principles in the Field of Human Artificial Procreation, *Int. Dig. Health Legis.* **40**(4), 907–912 (1989).
47. Quoted in E. Juengst, in E. Agius and S. Busuttil, eds., *Germ-Line Intervention and Our Responsibilities to Future Generations*, Kluwer Academic, Dordrecht, The Netherlands, 1998, pp. 85–102.
48. E. Agius, in E. Agius and S. Busuttil, eds., *Germ-Line Intervention and Our Responsibilities to Future Generations*, Kluwer Academic, Dordrecht, The Netherlands, 1998, pp. 67–83.
49. *Declaration on the Human Genome Project*, 44th World Med. Assembly, Marbella Spain, 1992.
50. N. Lenoir, *Kennedy Inst. Ethics J.* **7**(1), 31–42 (1997).
51. Statement on Research, HUGO Ethical, Legal and Social Issues Committee Report to HUGO Council. Available at: *http://www.gene.ucl.ac.uk/hugo/conduct.ht*
52. D. Butler, *Nature (London)* **398**, 179 (1999).
53. Universal Declaration on the Human Genome and Human Rights. Available at: *http://www.unesco.org/ibc/uk/genome/projet/index.htm*
54. Memo from E.M. Meslin to L. Staheli et al., *Report from UNESCO Meeting of Government Experts to Finalize the Draft Declaration on the Human Genome and Human Rights*, Paris, 1997.
55. H.-M. Sass, *J. Med. Philos.* **23**(3), 227–233 (1998).
56. D. Shapiro, *Politics Life Sci.* **13**(2), 233–234 (1994).
57. G.J. Annas, *N. Engl. J. Med.* **339**(24), 1778–1781 (1998).
58. R.M. Cook-Deegan, *Politics Life Sci.* **13**(2), 217–220 (1994).

See other GENE THERAPY and INTERNATIONAL ASPECTS entries.

GENE THERAPY, ETHICS, GENE THERAPY FOR FETUSES AND EMBRYOS

CAROL A. TAUER
College of St. Catherine
St. Paul, Minnesota

OUTLINE

INTRODUCTION

Trials of gene therapy for human fetuses and embryos have not yet been conducted. Discussions of in utero gene transfer have led to a consensus that it is too early to begin such trials. Gene transfer into early embryos has been debated as the most likely method for germ-line gene modification, and is similarly considered inappropriate for human trials at this time. However, animal research is making progress with both fetal and embryo gene transfer, indicating the importance of public discussion and education on possible human applications of these technologies.

TWO TYPES OF GENE THERAPY

Gene therapy in fetuses and gene therapy in embryos may appear to be quite similar, but the ethical and social issues raised by the two technologies are actually quite different. Gene therapy in fetuses, or in utero gene therapy, is technically and ethically similar to somatic cell gene therapy in born infants, children, and adults. Some additional ethical issues arise because interventions toward the fetus necessarily involve the pregnant woman. But these issues are similar to those that arise in any intervention or therapy directed toward the fetus, and are not specific to genetic interventions.

On the other hand, gene therapy in embryos is ordinarily understood to refer to genetic manipulation of in vitro fertilized embryos that have not yet been transferred to a woman. With the increasing availability of preimplantation diagnosis to identify genetic defects in early embryos, gene therapy offers a hope for correcting or ameliorating these defects rather than discarding the affected embryos. Since genetic intervention in embryos would occur at the developmental stage when all cells of the embryo are totipotent, or able to differentiate into all the cell and tissue types of the human organism, it is expected that genetic modifications of early embryos would affect germ cells as well as somatic cells. Thus genetic therapy in embryos is usually considered to be germ-line

gene therapy, hence raising all the ethical and social issues identified in the germ-line intervention debate. In addition the development of embryo gene therapy presupposes the ethical acceptability of research involving early embryos, an issue that is unresolved in American public policy. As of May 2000, federal funding of human embryo research is prohibited by act of Congress, although there is no prohibition or regulation of privately funded embryo research.

Because they involve significantly different ethical and social issues, gene therapy for fetuses and gene therapy for embryos will be discussed separately in this article.

GENE THERAPY FOR FETUSES

Prenatal diagnosis through amniocentesis became a possibility in the late 1960s, providing prospective parents the option of abortion if a genetically compromised fetus was identified. Even before induced abortion was legalized throughout the United States, such abortions were generally permitted under the heading of therapeutic abortion. While many opponents of abortion regarded prenatal diagnosis negatively, even describing it as a "search and destroy" mission, the hope was that eventually it would be possible to correct the genetic anomalies that were diagnosed (1).

Two streams of clinical research converge in fetal gene therapy. The first type is research on fetal treatment in general. Trials of pharmacologic and surgical interventions in fetuses have shown some promise in treating hereditary and other congenital conditions (1,2). The second type is research on somatic cell gene therapy, particularly gene transfer in infants and children, or transfer directed toward conditions where early intervention is preferable or even essential (3). Fetal gene therapy would combine these two types of innovative interventions by utilizing the procedures of somatic cell gene therapy to treat the fetus in utero.

One specific type of fetal treatment, the in utero transfer of hematopoietic stem cells, is particularly pertinent to the development of fetal gene therapy. Similar to a bone marrow transplant, the transfer of pluripotent hematapoietic stem cells is believed to hold great promise for the treatment of congenital blood disorders. Although trials reported thus far indicate only a handful of successes, all with fetuses having immunodeficiency disorders, proponents believe that interest and application are likely to increase. The advantages of stem cell transfer in utero rather than after birth are (*1*) the absence of an immune response to foreign cells in early gestation, (*2*) the possibility of developing "tolerance" to foreign cells that would allow for further treatment after birth, and (*3*) intervention that is early enough to correct a disorder before clinical and uncorrectable manifestations develop (4,5).

The RAC and In Utero Gene Transfer

When the Food and Drug Administration (FDA) began to receive applications for approval of investigational bone marrow and stem cell transplants into fetuses, it recognized that in adults, these procedures have been forerunners to trials of gene therapy. Thus in late 1994 it urged the Recombinant DNA Advisory Committee (RAC) of the National Institutes of Health (NIH) to begin to examine gene therapy in fetuses (6,7). In addition to ethical issues related to experimental gene transfer procedures and issues related to the involvement of the pregnant woman when treatment in utero is contemplated, scientists recognized a further issue. It is believed that in utero gene transfer, particularly at an early gestational stage, carries a risk of unintentionally affecting germ cells, a risk not incurred in trials of somatic cell gene therapy in children and adults. Thus in utero gene therapy could result in the first germ-line effects from gene transfer trials, brought about as unintentional side effects of a procedure classified as somatic cell gene therapy. Since the RAC has decided not to consider protocols that involve germ-line transfer at this time, such an unintended outcome would circumvent one of the ethical and social barriers currently believed to be prudent (6).

Also in 1994 NIH began a reassessment of the regulatory role of RAC regarding the approval of gene therapy propocols. Prior to 1994 all protocols had to be individually approved by RAC as well as FDA, resulting in what many perceived as duplication of review and unnecessary delays. A series of compromise proposals were considered by NIH administration, FDA, and RAC, and in 1996 Harold Varmus, director of the NIH, initially decided to eliminate RAC and to transfer its role in developing public policy to a new body within NIH. However, public comments weighed heavily against this plan, and Varmus decided to retain RAC, to continue requiring simultaneous submission of gene therapy proposals to RAC and FDA, but to require only FDA approval of individual protocols (8).

NIH's final policy decision was announced on October 31, 1997. It stated RAC's new functions as identifying specific human gene transfer proposals that raised novel issues deserving of public discussion, transmitting to the NIH director its recommendations on such proposals, and initiating consideration of forthcoming gene transfer procedures raising new ethical and social issues (8). Such consideration could be initiated in the absence of specific submitted protocols, with the intention of raising public awareness and obtaining public input prior to actual implementation of the novel procedures. Besides maintaining public access to its meetings, RAC together with NIH's Office of Recombinant DNA Activities (ORDA) would be expected to sponsor regular conferences on new developments in order "to serve as a unique public forum for the discussion of science, safety, and ethics of recombinant DNA research" (9).

These changes in the role of RAC are particularly pertinent to consideration of gene transfer in fetuses. Although no protocols for trials of in utero gene transfer had yet been submitted, in January 1999 RAC and ORDA sponsored a conference on "Prenatal Gene Transfer: Scientific, Medical, and Ethical Issues." This conference continued the discussion initiated when FDA referred the matter to RAC in 1994. The topic became even more timely as a result of the submission of two "preprotocols" for in utero gene transfer brought to RAC in July 1998 by

W. French Anderson and Esmail Zanjani. Their intention was not to seek approval for research trials but to stimulate discussion of the issues raised by this new application of gene transfer. While such preliminary discussion would be similar to discussions that took place before the first somatic cell gene therapy protocol was approved in 1990, it was particularly appropriate to the new role and function of RAC.

The two preprotocols on in utero gene transfer were the major topic of discussion for the RAC meeting of September 24, 25, 1998. Besides reviews submitted by individual RAC members, eight ad hoc consultants contributed reviews and six of them participated in the discussion at the meeting (10). Issues from this meeting were brought to the January 1999 conference on "Prenatal Gene Therapy," which was organized around questions assigned to three working groups: Preclinical Research Issues, Clinical Research Issues, and Ethical, Legal, and Societal Issues. By the end of the conference, areas of consensus and of disagreement had been identified. The main conclusion of the conference was stated as follows:

> At present, there is insufficient clinical data to support the initiation of clinical trials involving prenatal gene transfer. A substantial number of critical scientific, ethical, legal, and social issues must be addressed before clinical trials proceed in this arena (9).

This statement was followed by a listing of 26 specific areas in which more data were needed before clinical trials could be considered.

At the RAC meeting of March 11, 12, 1999, chairs of the three working groups presented their reports, responded to questions, and led discussion of the issues. While there were many areas of consensus, there were also some points of disagreement (11). The committee decided, however, that it was prepared to make a public statement in order to clarify its position:

> The RAC continues to explore the issues raised by the potential of *in utero* gene transfer research. However, at present, the members unanimously agree that it is premature to undertake any human *in utero* gene transfer experiment (11).

The RAC returned to the topic of in utero gene transfer at its June 14, 1999, meeting. At this meeting it reviewed a document that would eventually be published as a more detailed report of RAC's findings regarding prenatal gene transfer and that would be incorporated into the NIH *Guidelines for Research Involving Recombinant DNA Molecules* (12,13).

Summaries and minutes of RAC meetings, available through the NIH Web site and eventually through publication in the journal *Human Gene Therapy*, are an indispensable source of information regarding current issues in prenatal gene transfer. While committee discussion attempts to separate scientific and clinical questions from ethical and social issues, these areas overlap in many respects. For example, questions about the safety of gene transfer procedures may be regarded as preclinical or clinical questions, but the presence of significant risk to the fetus or pregnant woman also raises an ethical problem. In the following discussion, the focus on ethical and social issues will incorporate preclinical and clinical questions.

Areas of Ethical Consensus

Both in the ethics literature and in the discussions at RAC meetings, some areas of broad consensus have been identified, though consensus does not preclude the possibility of dissenting voices (1,10,11). Some consensus points relate to whether and when fetal gene transfer trials should begin, while others relate to requirements or conditions for protocols at a time when such trials are implemented.

Terminology. In the literature on research ethics there is continuing concern that research subjects may expect therapeutic benefit from their participation in research, even though there is no evidence for such benefit. This expectation, called the "therapeutic misconception," may lead subjects to accept risks that they would not accept if they realized that the procedures to which they are asked to consent are intended to gain knowledge for the treatment of future patients but are unlikely to benefit the research subjects (14). The therapeutic misconception may be reinforced by describing research procedures as "therapies" and by identifying research subjects as "patients."

Historically the research on gene transfer procedures aimed at therapeutic goals has been characterized as gene therapy research. However, nearly a decade of this research "has quite understandably failed to produce results swiftly leading to viable gene therapies" (15). In this situation, continuing to describe protocols as "gene therapy" contributes to a therapeutic misconception by potential research subjects as well as unwarranted expectations in the public. Because it is important that subjects have a realistic grasp of risks relative to expected benefits in order to give fully informed consent, misleading terminology should be avoided.

Several commentators have urged that the terms "gene therapy," "treatment," and "patient," be replaced by "gene transfer," "research," and "subject" in all gene transfer protocols (15). This proposal was supported by LeRoy Walters, former chair of RAC, at the meeting of September 24, 25, 1998, when he urged the committee to adopt the revised language (10). Examination of recent RAC documents indicates compliance with this proposal.

Genetic Counseling. In utero gene transfer would be performed on fetuses who are diagnosed as having a genetic condition. In turn, such prenatal diagnosis would be offered to couples who are at risk for transmitting a genetic disease. Thorough and unbiased genetic counseling must be available to all couples who have such risk factors. All options, including the option of abortion of an affected fetus, must be presented to them. After a genetic condition is diagnosed prenatally, the time frame for decision making is often brief. Yet time must be taken for the necessary information to be provided, for all options to be explained, and for reflection to occur.

Informed Consent. The consent process and form must clearly state that fetal gene transfer is an experimental procedure or intervention, that it may not benefit the fetus and will not medically benefit the woman, and that it carries specific (enumerated) risks to the fetus and to the pregnant woman. The pregnant woman provides consent for all interventions toward the fetus and herself. Though it is highly recommended that her partner be involved in discussions, he can neither consent to nor refuse interventions against her wishes. However, with regard to experimental interventions such as in utero gene transfer, it would be inadvisable to proceed if there were disagreement between the two partners.

Prenatal gene transfer cannot be made conditional on the woman's agreement not to seek a later abortion, nor on her agreement to have an abortion should the gene transfer not prove successful (1). (Note, however, that it may not be possible to test the effectiveness of the transfer until after birth.) Because an autopsy may be important in assessing the results of a clinical trial, the woman may be asked to agree to autopsy of the fetus or herself if either of them should die. However, if she changes her mind, an autopsy cannot be forced on her. The same thing is true of a requirement for long-term follow-up for herself and the prospective child. She may be asked to consent to follow-up assessments, but as with any research protocol, she cannot be forced to continue should she choose to withdraw.

Efficacy. As indicated in the RAC consensus statement, there is agreement that not enough is known about the potential effectiveness of in utero gene transfer to support clinical trials at this time (11). There is some promising research with animals, for example, the apparently successful reversal of cystic fibrosis in mice by gene therapy in utero (16). Trials of somatic cell gene therapy in born humans, however, have been disappointing. Since the first clinical trial in 1990, over 300 human gene transfer protocols have been registered with NIH, but none has been a clear therapeutic success. In the words of W. French Anderson, pioneer and proponent of gene therapy:

> The efficiency of gene transfer and expression in human patients is ... still disappointingly low. Except for anecdotal reports of individual patients being helped, there is still no conclusive evidence that a gene-therapy protocol has been successful in the treatment of a human disease (17).

Before subjecting pregnant women and their fetuses to experimental gene transfer procedures, researchers must have adequate evidence of potential efficacy in order to balance the risks involved. Many preclinical and clinical questions remain unanswered as of 1999. What is the most efficient way to transfer genes to the fetus? How can gene transfer be targeted to particular cell types without inadvertently exposing other cells, especially germ cells, to modification? When gene expression requires regulation to be successful, how would this be accomplished? The RAC working group on clinical issues agreed that clinical trials should not be undertaken until animal studies indicated that expression of transferred genes occurred at a level "conducive to correction of the phenotype rather than merely a slight change," or in other words, at a high level of effectiveness (11).

Safety. Gene transfer studies indicate that "the procedure appears to carry a very low risk of adverse reactions" (17). On September 17, 1999, the death of an 18-year-old man four days after a gene transfer into his liver marked the first fatality attributed to an experimental gene transfer procedure (18). One could argue that one fatality in over 300 protocols is a small number, but the event serves as a warning that gene transfer can involve serious risks.

In the case of treatment in utero, safety to the pregnant woman as well as to the developing fetus must be considered, and risks should be weighed separately for woman and fetus. The pregnant woman may consent to accept some additional risks for the sake of her fetus, but risks to her health should be minimized as they would be in any research involving a healthy subject. In most cases risks to the pregnant woman should be limited to risks arising from the method of gene delivery, for example, the risks of a surgical procedure.

In relation to the fetus, some of the safety issues are parallel to issues recognized in adult gene transfer procedures: the potential risk of mutagenesis caused by insertion of the transfer vectors, the possibility of vectors being replication-competent, and the risk of harmful unregulated gene expression. Additional questions raised by fetal gene transfer include: How might the process of fetal development be affected by introducing a vector? Is transplacental migration of virus vectors a possibility (thus potentially affecting the pregnant woman also)? Though the fetus may not have the same immune response to foreign DNA that is seen in adults, what other immune response problems might arise (11)?

Areas of Ethical Disagreement

Points of disagreement largely focus on which diseases, or types of diseases, would be the best candidates for the first trials of human in utero gene transfer. Should the first trial be for a disease where there is an animal model, and where effectiveness and safety have been clearly demonstrated in that model? Some diseases do not have animal models but nonetheless might be good candidates for human trials.

Should the first trial be for a disease where postnatal gene therapy of infants, children, or adults has been successful? For example, trials provide some evidence for the effectiveness of gene therapy in children with the immunodeficiency adenosine deaminase-severe combined immunodeficiency disorder (ADA-SCID) (1). However, in these trials an insufficient number of cells have been transduced to produce adequate ADA, so children in the gene transfer research have continued to receive supplementary administration of polyethyleneglycol (PEG)-ADA by shots that are both painful and expensive. It is possible that in utero gene transfer would improve transduction efficiency and hence eliminate the need for supplementary PEG-ADA. For this reason one of the preprotocols submitted to RAC in 1998 by Anderson and Zanjani related to ADA-deficient SCID (10). This preprotocol suggested a trial of in utero gene transfer for a condition where

there is some evidence for the effectiveness of postnatal transfer efficacy, and where there is a belief that prenatal treatment would be more effective.

While there is agreement that the first trials of fetal gene transfer should be for serious diseases, and certainly not for mere improvement of desirable traits, disagreement remains on these questions: Should in utero gene transfer be limited to diseases for which there is no effective nongenetic treatment? An affirmative answer would eliminate ADA-SCID, since PEG-ADA is an effective but burdensome treatment. Should fetal gene transfer be limited to diseases in which irreversible damage will occur in utero if the disease is not corrected at an early gestational stage?

A further disagreement among RAC members relates to the treatment of diseases that are ordinarily fatal in utero. This issue arose in connection with the second of the Anderson-Zanjani preprotocols, gene transfer for α-thalessemia. Besides being lethal to the fetus, this disease also produces toxic symptoms in the pregnant woman, and is thus an indication for therapeutic abortion. Although various in utero therapies for α-thalessemia have been attempted, none has been therapeutically effective, although there have been cases of partial correction resulting in the live birth of a severely affected infant (10).

Comments on the α-thalessemia gene transfer preprotocol expressed two contrasting points of view. Some RAC members and consultants argued that the disease was a good choice for a gene transfer experiment because it would otherwise be fatal in utero, and because there were no existing treatments for the disease. Others maintained that a disease that was fatal in utero was not a good choice for gene transfer, but that it would be better to let nature take its course. The risk of contributing to the birth of a severely ill newborn for whom there was no effective treatment, plus the serious risk to the woman from carrying a potentially toxic pregnancy to term, both argue against the selection of α-thalessemia for prenatal intervention.

The controversial issues involved in determining which types of diseases would be most appropriate for the first trials of fetal gene transfer are typical of the kinds of questions on which RAC and NIH seek public input. Given the RAC position that clinical trials of any in utero gene transfer would be premature at this time, there is a window of opportunity for public education and discussion to occur.

GENE THERAPY FOR EMBRYOS

Embryo Gene Transfer as Germ-Line Modification

Discussions of germ-line gene therapy agree that genetic modification of the one-celled zygote or early embryo is probably the most feasible way to make germ-line alterations. Because early embryonic cells are totipotent, or able to differentiate into any of the cell or tissue types of the organism, a genetic modification introduced at this stage has the potential to affect the developing germ cells are the sperm and the eggs. The only way to make the zygote or early embryo available for gene transfer procedures is through in vitro fertilization (IVF), thus requiring use of this reproductive technology as preparation for gene transfer. It may also be possible to achieve germ-line changes through modification of sperm stem cells. However, since females possess all their egg cells at birth, direct modification of egg cells is highly unlikely (19).

Animal research involving the genetic modification of mouse embryos shows that gene transfer before cell differentiation can produce changes that are transmitted to offspring. For example, Leroy Hood's experiments with "shiverer mice" demonstrate that genetic alterations of affected embryos can correct expressions of this phenotype, not only in the mice that develop from the treated embryos but also in their descendants. Other research groups have prevented the transmission of serious diseases in generations of mice by genetically altering mouse embryos (20,21). According to LeRoy Walters and Julie Palmer, there has been more success with germ-line genetic intervention than with somatic cell gene therapy in laboratory animals (19).

The ethical and social issues involved in gene therapy for embryos encompass all the controversies raised by germ-line transfer in general. In fact RAC has not explicitly considered the topic of gene therapy in embryos, while it has taken a position regarding germ-line gene therapy. Its position is stated in the NIH *Guidelines*: "[The] RAC will not at present entertain proposals for germ-line alterations" (13). The RAC sees its role as providing a forum for discussion of germ-line transfer, both at its meetings and through sponsored conferences, as long as it is clear that such discussion does not imply RAC endorsement of germ-line transfer (22).

Advantages of Gene Therapy for Embryos

The most obvious benefit from genetic modification of early embryos would be the correction of lethal and other serious diseases so that they will not be transmitted to offspring and later descendants. While somatic cell gene therapy may cure or help an individual who suffers from a genetic disease, it will not prevent transmission of a heritable disease to the next generation, since somatic cell gene transfer will not alter germ cells. When a disease is undesirable for an individual, it is equally undesirable for that individual's descendants, and its elimination may appear to have only good consequences.

Not only does embryo gene therapy offer benefits to the individual and family involved, but it promises long-term benefits related to improving the human gene pool. Somatic cell gene therapy may cure or ameliorate a disease in an existing individual, but in extending that person's life and enabling him or her to reproduce, it could have the effect of increasing the total number of persons (offspring) suffering from the disease. In preventing this outcome, germ-line gene therapy may be preferable to somatic cell therapy.

For some genetic conditions, somatic cell gene transfer would not work in principle. Diseases expressed in nondividing cells, like the cells of the neural system, are intractable to many somatic cell gene transfer techniques. Thus a disease like Lesch-Nyhan syndrome might respond to gene transfer into affected embryos, while somatic

cell gene transfer would be ineffective. Similarly somatic cell techniques that require removal of cells before their modification would not work with diseases expressed in nonremovable cells (19).

Some genetic diseases, for example, cystic fibrosis, affect many different organs and cell types in the body. Gene transfer at the early embryo stage would allow delivery of a corrective gene to all affected cell types, while later somatic cell transfer could require repetition of a variety of gene transfer procedures. This argument also applies to cancers that result from inborn genetic factors and subsequent mutations. Such cancers have the propensity to affect more than one organ or system, and hence might be prevented most effectively through correcting the genetic defect at the early embryonic stage (19).

Finally, some genetic conditions result in irreversible damage, often to the brain, during the first trimester of pregnancy. Corrective gene transfer into the fetus or after birth could not ameliorate this damage. Since the embryo in utero is not accessible to treatment until some weeks after implantation, early intervention requires that the gene transfer be done in the laboratory after IVF. Gene transfer into the zygote or early embryo may be the only way to prevent the intrauterine harms that are foreseen.

Arguments Against Gene Therapy for Embryos

Standard arguments against gene therapy for embryos are summarized by Walters and Palmer (19). These arguments generally view embryo gene transfer as a form of germ-line gene therapy, and hence focus on the issue of modifying the genetic heritage of individuals and of the human race.

Safety. A germ-line gene transfer procedure may have negative effects, including some that do not show themselves until later generations. In the case of somatic cell gene transfer, negative effects harm only the individual who is treated (or in the case of a fetus in utero, possibly also the pregnant woman). But when a genetic modification affects the germ cells, as is the case with modification of early embryos, then all descendants of the transfer recipient could be negatively affected. There is speculation that techniques will be developed to remove or render inoperable an inserted gene that is causing trouble for later generations (19). However, there are no guarantees that such reversal will be available.

Safety issues must be resolved through animal studies before attempts are made to modify human embryos. Even when animal studies demonstrate that gene transfer techniques are safe and effective, there is some risk in moving to human applications because of possible species differences. For successful gene transfer into embryos, the added genes must integrate without disrupting normal development of the resulting fetuses. The genes must integrate into all cells of the early embryo rather than merely some of them, a situation that could produce a genetic chimera. And the integrated genes must be properly expressed later in the born human being (19).

Use of Resources. Gene transfer into embryos would be a complex, multi-step procedure. First, the couple at risk for transmission of a genetic defect would have to conceive through IVF. While this procedure is ordinarily chosen by couples who are infertile, here it would be used in order to make embryos available for diagnosis and manipulation. Second, it would be necessary to identify the specific embryos that have the genetic defect by using preimplantation genetic diagnosis (PGD). This procedure is itself experimental, and its application has involved some erroneous diagnoses (23). Third, some type of gene transfer would be utilized in order to correct the genetic defect in the embryos that were identified.

Each of the three steps in this treatment protocol is expensive and technology intensive. Most likely the protocol could be made available only to affluent families, unless alternate funding were available during the research phase. In a society where millions of people lack access to basic health care, it is questionable whether such a procedure is a prudent use of medical, scientific, and technological resources.

Inviolability of Genetic Heritage. Some opponents of germ-line gene transfer argue that making genetic modifications that will be passed to future generations involves an improper tampering with the future of evolution and the human gene pool. These critics may rely on religious arguments, claiming that scientists would be "playing God," or they may hold that there are natural limitations built into the universe that humans ought not to exceed (19).

A related argument claims that human beings have a right to receive a genetic heritage that has not been tampered with. Stated as a type of human right, this prerogative has been enunciated by the Parliamentary Assembly of the Council of Europe. In its strongest form, the presumed right has led to adoption of a constitutional ban on germ-line intervention in Switzerland (24).

Positions like these are deontological, or based on a theory of rightness and wrongness that is independent of the consequences that are actually produced. Hence they cannot be refuted by arguments that point to the potential elimination of lethal or other serious genetic diseases. They might be countered, however, by making analogies to other medical procedures that seem to interfere with the ordinary course of nature, or to other ways in which we alter genetic heritage, for example, by selective breeding of animals. Opponents might still argue that human genetic heritage is different and should be regarded as inviolable.

Alternatives to Gene Therapy in Embryos

One particular form of opposition to embryo gene therapy focuses on alternatives that are regarded as preferable. Two options are offered. Since embryo gene therapy must be preceded by preimplantation genetic diagnosis, the genetic disease under consideration could be avoided by simply discarding affected embryos. Alternatively, in order to avoid the complex procedures of IVF and PGD, prenatal diagnosis during an established pregnancy followed by abortion of an affected fetus would achieve the same goal (25).

Either of these options may be viewed as safer, both in terms of achieving the desired result and in order to avoid harmful side effects or possible negative outcomes in a later generation. However, there may be situations in which the options are not available; for example, when both partners have two copies of the same malfunctioning gene, that is, both are afflicted with the same recessive genetic disorder. In this situation all genetic offspring of the couple would necessarily have the same disease as their parents (19).

Additionally some couples may have moral objections to aborting an affected fetus, or to discarding embryos that are identified as affected. Walters and Palmer suggest that attempting to treat a genetic disease, rather than discarding embryos or eliminating fetuses that have the disease, fits more closely the mission of the health sciences and shows greater respect for born persons who have a genetic disease or disability (19). Munson and Davis argue that medicine by its very nature has a therapeutic obligation to pursue the development of genetic therapy as a way of curing and eliminating disease (26).

The close relationship between preimplantation genetic diagnosis and gene therapy in embryos suggests that the ethical and social issues posed by these procedures should be considered together. Yet, as Pergament and Bonnicksen note, the form that public policy discussions have taken has led to their separation, with PGD seen as an issue related to research on reproductive technologies, and embryo gene therapy viewed as germ-line gene transfer and thus coming under the purview of RAC (27).

Research Involving Early Human Embryos

Public policy in the United States does not allow federal funding of research that involves early human embryos. Thus research on preimplantation genetic diagnosis, a necessary forerunner to gene therapy in embryos, may not be funded or sponsored by the National Institutes of Health.

In 1994, as a result of a change in the congressional language on appropriations for NIH, it appeared that some research involving IVF and early embryonic development would be fundable. At that time the director of NIH appointed a panel, the Human Embryo Research Panel (HERP), to develop guidelines for such funding. In its report HERP recommended federal funding for a variety of types of research related to infertility. It also gave its approval to funding for research on preimplantation genetic diagnosis (28).

The Panel was explicitly directed not to consider the issue of research on germ-line gene transfer. On the assumption that such research fell within the purview of RAC, the charge to the Human Embryo Research Panel stated that "Research involving human germ-line gene modification ... is not within the Panel's scope" (28, p. ix). This somewhat artificial separation of tasks eliminated the possibility of joint consideration of two closely related procedures, preimplantation genetic diagnosis and embryo gene transfer (27). Because of the limitations stated in its charge, HERP did not discuss embryo gene transfer in any way.

Similarly RAC has given essentially no attention to the fact that research on germ-line gene transfer would most likely involve research on early human embryos. If germ-line transfer is ever to be seriously considered as a public policy issue, this aspect of the issue would have to be debated.

Debate would be further complicated by the fact that shortly after the HERP report was approved by the Advisory Committee to the Director of NIH, Congress rescinded its approval of federal funding for IVF research. Beginning with fiscal year 1996 the appropriations bill for NIH has specifically prohibited funding of any research in which early human embryos are harmed or destroyed. It is unlikely that the congressional prohibition will be removed in the foreseeable future, with the possible exception of allowing research on embryonic stem cells that may have great therapeutic potential for treating diseases in already-born people.

Embryo Gene Transfer for Enhancement Purposes

While many advocates of gene transfer, whether germ-line or somatic cell, have stressed that it ought to be used only for therapeutic goals, others have recognized the likelihood that its use will eventually be extended to the enhancement of desirable human traits. Wivel and Walters believe that "Germ-line gene modification for serious disease will inevitably lead to the next step, genetic enhancement" (29). RAC has raised this concern about germ-line transfer, acknowledging that once the technique is successful as a medical procedure, it has the potential for "off-label" use for enhancement purposes. The issue of genetic enhancement is thus an important topic for debate by RAC, one where the committee recognizes its role in facilitating public discussion, education, and input (30).

A number of examples illustrate the tendency to extend use of FDA-approved therapies to enhancement purposes. When recombinant DNA procedures made human growth hormone (HGH) available in large quantities, its therapeutic use to treat dwarfism resulting from HGH deficiency was quickly expanded to the treatment of children who were short but had no medical condition (31). Breast implants that were developed for women who had radical surgery because of cancer became cosmetic prostheses for women who desired larger breasts. In the gene transfer area, recent reports of mice whose hair growth was stimulated by insertion of the Sonic hedgehog gene note that therapeutic use in humans would be directed to persons with hair loss due to chemotherapy. However, cosmetic uses of gene transfer to reverse baldness are anticipated, either with enthusiasm or with hesitation, depending on the perspective of the commentator (32).

While early discussions contrasting gene transfer for therapy with gene transfer for enhancement suggested that the distinction is clear, this is by no means the case (19,33). The preceding examples suggest that the line between a supposed therapeutic application and an enhancement or cosmetic application may be somewhat murky. Walters and Palmer distinguish between enhancement goals that are health related and

those that are not. An enhancement that is health-related alters one's physical condition so that one is more resistant to disease. For example, immunization is widely accepted as a physical enhancement that renders the body immune to certain infectious diseases. Similarly a gene transfer procedure might provide immunity to acquired immunodeficiency syndrome (AIDS), or enable the body to fight cancer more effectively (19).

Arguments by analogy with current practices thus seem to give some support to genetic modification for health-related enhancements like disease immunity. But once again, it is not always easy to distinguish health-related enhancements from non-health-related enhancements. Cognitive improvements such as improved memory might initially seem non-health-related. Yet, if one were able to use gene transfer to improve the cognitive functioning of a mentally retarded person, this enhancement might be seen as health related, or even therapeutic.

Some authors have viewed the prospect for genetic enhancement as the most perplexing problem in the germ-line gene transfer debate (25). For those who view enhancement uses of genetics as undesirable, the potential for these applications provides a strong argument against pursuing germ-line gene transfer. The fears of opponents are reinforced by assertions that "the only plausible reason to insert genetic material into embryos would be for genetic enhancement" (34). Given the availability of less complex means to avoid the birth of genetically compromised children (prenatal or preimplantation diagnosis), enhancement may eventually become the main reason for germ-line gene transfer, or for gene transfer in early embryos.

BIBLIOGRAPHY

1. J.C. Fletcher and G. Richter, *Hum. Gene Ther.* **7**, 1605–1614 (1996).
2. M.I. Evans et al., in A. Milunsky, ed., *Genetic Disorders and the Fetus: Diagnosis, Prevention, and Treatment*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1992, pp. 771–797.
3. C. Coutelle et al., *Nature Med.* **1**(9), 864–866 (1995).
4. A.W. Flake and E.D. Zanjani, *J. Am. Med. Assoc.* **278**(11), 932–937 (1997).
5. Gene Therapy Advisory Committee, Health Departments of the United Kingdom, *Hum. Gene Ther.* **10**, 689–692 (1999).
6. H. Gavaghan, *Nature Med.* **1**(3), 186–187 (1995).
7. P.D. Noguchi, *Food Drug Law J.* **51**, 367–373 (1996).
8. R.A. Merrill and G.H. Javitt, in T.J. Murray and M.J. Mehlman, eds., *Encyclopedia of Ethical, Legal & Policy Issues in Biotechnology*, Wiley, New York, 2000.
9. National Institutes of Health, *Gene Therapy Policy Conference: Prenatal Gene Transfer*, Available at: *http://www.nih.gov/od/orda/gfpcconc.htm* (1999).
10. National Institutes of Health, Recombinant DNA Advisory Committee, *Minutes of meeting, September 24, 25, 1998*, Available at: *http://www.nih.gov/od/orda/9-98rac.htm#IV* (1999).
11. National Institutes of Health, Recombinant DNA Advisory Committee, *Minutes of meeting, March 11, 12, 1999*, Available at: *http://www.nih.gov/od/orda/3-99rac.htm* (1999).
12. National Institutes of Health, Recombinant DNA Advisory Committee, *Minutes of meeting, June 14, 1999*, Available at: *http://www.nih.gov/od/orda/6-99rac.htm* (1999).
13. National Institutes of Health, *Guidelines for Research Involving Recombinant DNA Molecules*, Appendix M., Available at: *http://www.nih.gov/od/orda/apndxm.htm* (1999).
14. P.S. Appelbaum, L.H. Roth, and C.W. Lidz, *Int. J. Law Psych.* **5**, 319–329 (1982).
15. N.M.P. King, *Hum. Gene Ther.* **10**, 133–139 (1999).
16. J.E. Larson et al., *Lancet* **349**, 619–620 (1997).
17. W.F. Anderson, *Nature* **392**(Supp. April 30), S25–S30 (1998).
18. R. Weiss and D. Nelson, *Washington Post*, September 29, p. 1 (1999).
19. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997.
20. J.W. Gordon, *Int. Rev. Cyt.* **115**, 171–229 (1989).
21. R.D. Palmiter and R.L. Brinster, *Ann. Rev. Gen.* **20**, 465–499 (1986).
22. National Institutes of Health, Recombinant DNA Advisory Committee, *Hum. Gene Ther.* **10**, 489–533 (1999).
23. A.L. Trounson, *Hum. Reprod.* **7**(5), 583–584 (1992).
24. A. Mauron, *Politics Life Sci.* **13**, 230–232 (1994).
25. N. Holtug, *Cambridge Q. Healthcare Ethics* **6**, 157–174 (1997).
26. R. Munson and L.H. Davis, *Kennedy Inst. Ethics J.* **2**(2), 137–158 (1992).
27. E. Pergament and A. Bonnicksen, *Am. J. Med. Genet.* **52**, 151–157 (1994).
28. National Institutes of Health, Human Embryo Research Panel, *Report of the Human Embryo Research Panel*, National Institutes of Health, Bethesda, MD, 1994.
29. N.A. Wivel and L. Walters, *Science* **262**, 533–538 (1993).
30. National Institutes of Health, Recombinant DNA Advisory Committee, *Hum. Gene Ther.* **9**, 911–932 (1998).
31. C.A. Tauer, *IRB* **16**(3), 1–9 (1994).
32. N. Wade, *N.Y. Times*, October 5, pp. C1, C9 (1999).
33. E. Juengst, in T.J. Murray and M.J. Mehlman, eds., *Encyclopedia of Ethical, Legal & Policy Issues in Biotechnology*, Wiley, New York, 2000.
34. J.R. Botkin, *J. Law Med. Ethics* **26**, 17–28 (1998).

See other GENE THERAPY entries; HUMAN SUBJECTS RESEARCH, ETHICS, RESEARCH ON HUMAN EMBRYOS; REPRODUCTION, ETHICS, MORAL STATUS OF THE FETUS; REPRODUCTION, LAW, WRONGFUL BIRTH, AND WRONGFUL LIFE ACTIONS.

GENE THERAPY, ETHICS, GERM CELL GENE TRANSFER

ROBERT NELSON
University of Pennsylvania
Philadelphia, Pennsylvania

OUTLINE

INTRODUCTION

Germ cell gene transfer (GCGT) is distinguished from somatic cell gene transfer by the fact that the intervention alters the DNA of the germ cells (with or without altering the DNA of somatic cells) and the alteration is transmitted to the individual's progeny. Thus GCGT can be directed to the germ cells of a differentiated organism (sperm or egg) through gametocyte modification, or to the undifferentiated organism at an early embryonic stage (prior to the cellular distinction between somatic and germ cell lines), hereafter referred to as "pre-embryo modification." Somatic cell gene transfer, whether for therapeutic or enhancement purposes, is limited in effect to the individual who is the recipient of the technique. Gametocyte modification is directed solely at the transmission to an individual's progeny of either undesired or desired genes, independent of whether the parent has undergone the same modification. Pre-embryo modification would likely result in the genetic alteration of both somatic and germ cell lines so that the resulting individual and his or her progeny would benefit from the gene transfer. In addition to altering intergenerational transmission of genes, pre-embryo modification may emerge as an effective technique for somatic cell gene transfer. As such, we need to consider whether GCGT as a result of pre-embryo modification can be considered an acceptable albeit indirect effect (1). Nevertheless, any technique of GCGT would be precluded if there were compelling moral reasons why we should never directly or indirectly intend to alter the human genome so as to impact on our future progeny.

DISCUSSION OVERVIEW

This article begins with a brief review of some technical issues in GCGT to set the stage for considering the ethics of GCGT research. It is inappropriate to discuss the issue of GCGT as therapy before considering GCGT as research in light of current federal research guidelines. Four conditions will be identified which must apply for initial GCGT research to proceed: (*1*) no alternative treatments exist (including somatic cell gene transfer), (*2*) the phenotypic injury occurs early in fetal development, (*3*) the outcome is uniformly fatal, and (*4*) the condition is caused by a single gene defect. With this as a foundation, we then explore the major objections to GCGT based on an illegitimate "tampering" with our genetic endowment balanced by an obligation to heal those afflicted with a genetic disease. This obligation to treat an individual, however, may be interpreted as an obligation to prevent genetic disease in a population, raising concerns about the eugenic use of GCGT. Nevertheless, from an individual perspective, a parent may appropriately choose to enhance a child's opportunities through the use of genetic technology. If so, the challenge will be to distinguish between legitimate and illegitimate eugenic uses of GCGT — a challenge that raises the issue of the social and political control of GCGT technology. Before addressing the control of genetic technology, we need to address the similarities and differences between pre-embryo discard as part of GCGT procedures and the controversial issue of abortion. In addition GCGT needs to be considered in the context of the alternatives — most of which currently involve either embryo discard or abortion.

TECHNICAL ISSUES WITH GCGT

Gene transfer techniques that may be applied to germ cells fit into three broad categories: gene augmentation, gene modification, and gene excision and splicing. Gene augmentation in which a functional gene is inserted into a cell to direct the synthesis of an otherwise missing or defective gene product, or gene modification in which a functional gene is inserted into the nuclear DNA, have been used for somatic cell gene transfer. Gametocyte modification using gene augmentation probably will not be an effective technique given the need for the gene to be distributed predictably to subsequent cells (such as gametes) (2). Gene modification through the insertion of a replacement gene into nuclear DNA would allow for the transmission of the inserted gene to the subject's gametes; however, the insertion may disrupt otherwise functional genes, uncover or create proto-oncogenes, or lead to gene expression in inappropriate tissues (2). Given the need for precise timing and expression of gene products during embryonic and fetal development, it is likely that only gene excision and splicing targeted to particular missing or nonfunctional genes will meet the necessary standards of safety, accuracy and cellular integrity (3). However, accurate gene excision and splicing techniques do not exist at the present time.

Gene transfer techniques targeted at either somatic or germ cells may be effective only with single gene disorders given the necessary causal assumption that the simple absence of a functional gene results in a diseased phenotype. The correction of a dominant disorder may require the removal of the offending gene, whereas the simpler technique of gene augmentation or modification may correct a recessive disorder. Pre-embryo modification also depends on the ability to diagnose the targeted genetic disease at an early stage of in vitro development in order to determine the need for or success of treatment. Finally, there are a number of diseases involving mitochondrial DNA that may be amenable to either the insertion of a functional gene or the complete exchange of functional for dysfunctional mitochondria (3,4). GCGT techniques that target mitochondrial DNA can be considered a special

case of either gene augmentation or modification and will be discussed separately.

ETHICS OF GCGT RESEARCH

The ethical application of GCGT to either the treatment of human disease or the enhancement of human characteristics would first require that the techniques undergo a thorough research evaluation. Before the development and implementation of a human protocol, the gene transfer techniques to be used would need to be thoroughly tested in animals, including trials in nonhuman primates (5). The lack of an animal model for many human diseases that may be appropriate candidates for GCGT inevitably gives rise to uncertainty in moving from the nonhuman to the human subject. However, given that the human applications of GCGT require that the procedures first undergo research evaluation, it is useful to reflect on GCGT in the context of current federal research regulations (6).

The first question to consider is who is the subject of the research. The adult is clearly a participant in any research involving gametocyte modifications; however, one could argue that a child who is the anticipated product of the research should also be considered a research subject. To eliminate this possibility, it is conceivable that an adult who is otherwise at risk for passing a genetic defect onto his or her progeny, yet does not want children, may volunteer for early testing of gametocyte modification. The safety of the technique for the adult subject, as well as the efficacy in altering the genome of the resulting gametes, then could be studied without concern for a future child. With voluntary and informed consent, we currently allow competent adults to participate in nontherapeutic research which places them at some personal risk. However, the ethical acceptability of this early testing of gametocyte modification assumes that the adult subjects will not procreate — a stipulation that cannot be guaranteed if the subjects remain fertile after the research. A requirement that adult subjects be sterilized as the last step of a research protocol would prevent harm to any resulting progeny; however, such a requirement would be impossible to enforce as current research guidelines permit an individual to exit voluntarily from a research study at any point.

It is problematic to impute the right to be free from harm to an individual who does not exist (7); however, if a fetus is injured during research involving gene transfer, the ensuing child both exists and suffers as a result of the research. Since an adult participant may have a child after undergoing gametocyte modification, the risks and benefits of GCGT research protocols involving either gametocyte or pre-embryo modification should be evaluated from the perspective of the fetus/child who is both the product and the subject of the research (8). Although the definition of a "human subject" would need to be expanded beyond "a living individual," GCGT then would be judged according to the federal guidelines governing research using fetuses and children (6). Finally, the woman who becomes pregnant after pre-embryo modification and implantation could not be considered the research subject, for this would preclude the research as not being directed toward the "health needs of the mother" and the risk to the fetus is clearly greater than minimal (6).

The general requirements for approval of a research protocol involving an adult stipulate that the risks of the procedure are minimized and are reasonable in relation to any anticipated benefits and/or the importance of the knowledge (6). Unless a technique for targeting gametocytes can be developed that does not alter the subject's somatic cells, the subject will be at risk for the complications of somatic cell gene transfer such as the inappropriate expression of gene products in different tissues and the disruption of otherwise functional gene products or the unmasking of oncogenes through insertional mutations (2). Precise gene excision and splicing in which the defective gene is removed and replaced with a functional gene theoretically would reduce or eliminate these risks; however, this technique is not possible at this time. In effect, in vivo GCGT techniques involving the adult gametocyte should not be approved until the development of selective targeting or until these same techniques are deemed safe for somatic cell gene transfer. These difficult technical requirements for an appropriate research protocol along with the availability of in vitro fertilization after either in vitro gametocyte or pre-embryo modification make it unlikely that an in vivo GCGT protocol will be developed or approved.

Any GCGT protocol involving in vitro gametocyte or pre-embryo modification necessarily involves an adult as either a gamete donor or a pre-embryo recipient. A GCGT protocol may present no greater than minimal risk to an adult participant provided that (*1*) gamete procurement uses standard nonresearch collection techniques, (*2*) any in vitro genetic modification takes place prior to implantation, and (*3*) the pre-embryo modification does not alter the biology of pregnancy, so that the woman who is the pre-embryo recipient is not placed at any additional risk beyond that associated with standard in vitro fertilization procedures. Appropriate nonhuman primate studies may be necessary to establish this third condition, for otherwise the pregnant woman would need to be considered a research subject and the research would be disallowed under current federal guidelines governing research with pregnant women (6).

Currently, neither the Recombinant DNA Advisory Committee of the National Institutes of Health (NIH) in the United States nor the Joint Medical Research Council in the United Kingdom will consider a GCGT protocol (5). Although there is no direct ban on GCGT research, there is a ban on the use of federal funds either for the creation of a human embryo for research purposes or for research in which a human embryo is "destroyed, discarded, or knowingly subjected to risk of injury or death" beyond that allowed by existing federal guidelines (9,10). Unless the research is privately funded, this ban would need to be modified to allow for the creation of a human embryo even if the intent was to implant and not discard all created embryos. Once the pre-embryo is implanted, federal guidelines require that the research is designed to meet the "health needs of the particular

fetus" and restrict any risk to the minimum necessary to meet such needs (6). In effect, the fetal guidelines are similar to those for "greater than minimal risk" research with children. The risk of the research must be justified by the "anticipated benefit" for the child. The balance of risk and benefit must also compare favorably with alternative approaches. Treating the future child as a research subject does not preclude the development of a pre-embryo modification protocol due to the inability to obtain consent, for the assent of the child may be waived if the benefit is not available outside of the research (6).

The development and approval of a GCGT protocol, as discussed above, would need to build on prior experience using somatic cell gene transfer for the same or at least similar conditions (5). Consequently somatic cell gene transfer may serve as a viable alternative for certain conditions. The benefit of germ cell modification would be to eliminate the need for both the individual person and his or her future progeny to undergo somatic cell modification. The risks of germ cell modification would need to be balanced against those of somatic cell modification — a balancing that likely would favor somatic cell modification given the complexity and uncertainty of embryonic, fetal, and child development. The claims of the adult suffering from the genetic condition under consideration to want to free his or her progeny from the putative guilt of defective reproduction or the burden of somatic cell treatment may not be compelling if a safe and effective somatic cell treatment is available. Nevertheless, for conditions that require a more extensive and uniform distribution of the transferred gene, or that will have an impact on the future child at an early stage of in utero development, GCGT may be the only effective method for somatic cell modification. Thus concern about scientific uncertainty is met by a specification of the conditions under which such uncertainty is worth the risk. The conditions then that may be suitable for initial testing of GCGT techniques are those for which there are no available alternative treatments (including somatic cell gene transfer), the impact of the genetic defect occurs early in fetal development, the outcome is uniformly dismal or fatal, and the phenotypic condition is caused by a single gene defect (11). Once GCGT is shown to be safe and effective for this limited range of genetic disorders, the techniques could be extended (again within a research protocol) to other less ominous diseases, or to diseases for which safe and effective somatic cell therapies may exist. GCGT aimed at the enhancement of human characteristics would not be approved, if ever, until GCGT has been shown safe and effective for a variety of diseases.

MAJOR OBJECTIONS TO GCGT

The primary motivation leading to the development of GCGT technology is therapeutic; that is, it begins with a parent's desire to bear a child who is free from the burden of an otherwise untreatable genetic disease. In the absence of strong counterobjections, it can be argued that medicine has a prima facie duty to pursue research on the therapeutic use of GCGT. Some argue the strong moral claim that the character of medicine as seeking knowledge for healing purposes mandates the exploration of GCGT in the absence of compelling objections (2,5,12,13). If objections to the development and implementation of GCGT rests primarily in the fear of unknown and potentially disastrous consequences, the cautious and gradual implementation of germ cell protocols guided and controlled by an already established research review process may provide for an early recognition and mitigation of untoward consequences (13). Nevertheless, we would prevent the transmission of our defective gene(s) to the children of our children. Is this "cleansing of our family line" an unacceptable outcome, even if it is not our primary intent? Are GCGT techniques fundamentally wrong under all circumstances such that we can never choose it as a means to an otherwise acceptable end?

Munson and Davis interpret principled objections to GCGT as involving a basic claim of illegitimate "tampering" with either individual rights, social order, or nature itself. First, concerning individual rights, the Council of Europe asserts the right of the unconceived and unborn to a genetic inheritance which has not been artificially altered. An exception is made for therapeutic interventions based on a distinction between pathological and nonpathological conditions, that is, between treatment of disease and enhancement. The Council bases such a right on appeals to human dignity, integrity, the "normal" or "natural," genetic divergence or an appeal to the preservation of being human — concepts that neither entail an "alleged right to an untouched genome" nor escape the definitional ambiguities inherent in the concept of disease. Even if we reject the position that our future progeny have the right to be left alone, we may still hold that they have the right not to be knowingly harmed. The limits of this right would involve our ability to predict the results of our otherwise well-intentioned interventions, relieving us of the unreasonable burden of knowing in advance all of the potentially negative consequences of our gene transfer technology. This balancing of the risks and benefits of intervention is involved when GCGT is considered as a research protocol aimed at freeing an individual of genetic disease (12).

Second, GCGT may give rise to social disorder or conflict. Will a parent be required to submit to germ cell modification in order to prevent the transmission of a genetic disease? The apparent conflict between community interests and individual freedom is not new. The availability of genetic testing for such conditions as Huntington's disease, and the knowledge of the fetal impact of maternal drug or alcohol abuse, already have led some to advocate for restrictions on individual behavior. If GCGT is used for the enhancement of desirable biological traits, existing socioeconomic inequalities may be exacerbated. However, current inequalities in access to and distribution of health care resources already reinforce existing socioeconomic differences. These are important problems that we must face in the design, implementation and control of health care technology; however, they are not unique to GCGT (12).

The third version of "tampering" concerns "playing God" or altering the "very order of nature." Munson and Davis

identify three arguments against such tampering. First, the concern that GCGT will inevitably lead from treatment of disease ("negative eugenics") to enhancement ("positive eugenics") is met with an empirical claim that eugenic practices such as selective breeding have existed for centuries yet rarely been used. Acknowledging that the distinction between positive and negative eugenics may be difficult to sustain, Munson and Davis prefer to question the assumption that our desire to improve ourselves through genetic technology is wrong. Munson and Davis also take a cautious empirical approach to the second concern that unforeseen hazards may exist, for example, in the loss of biological diversity as genes are eliminated. They see no reason to fear a disaster from the development and application of GCGT more than from other applications of genetic technology (13,14). Finally, they address the concern that GCGT threatens our "humanity" by creating the possibility of human/nonhuman hybrids or an evolution into a superior yet nonhuman species. As an empirical concern, it is unlikely that the elimination or addition of genes through GCGT will corrupt or eliminate a genetic structure that is somehow essential to our humanity. By linking our sense of self-worth to the inviolability and integrity of our genetic structure, the opponents of GCGT appear to adopt the same reductionist assumption of genetic determinism that proponents of genetic technology are assumed to hold. Munson and Davis propose that behind all three objections is a belief in either the "wisdom of evolution" or the "design of a good and wise Being." However, not only is it impossible to establish the "sanctity" or "special moral standing" of human nature apart from specific theological or moral commitments, the very existence of genetic disease belies the wisdom of natural selection and the coherence of providential design (12,15). It is as likely that our ingenuity in developing and applying GCGT technology is either an evolutionary adaptation that enhances our chances for human survival or an affirmation of our God-given stewardship over creation (13).

Practical objections to the development of GCGT involve scientific uncertainty and the unpredictability of long-term risk, both addressed through an incremental process of research as outlined above. Some assert that GCGT will never be sufficiently cost-effective to merit the allocation of necessary social and economic resources. This objection, however, is true of all unproved technologies prior to our attaining sufficient knowledge to evaluate the costs and benefits. For some, the strongest pragmatic argument against GCGT is the existence of pre-embryo screening and selective implantation (16,17).

EUGENICS AND THE DESIRE TO PREVENT GENETIC DISEASE

GCGT has the potential to be more effective than somatic cell gene transfer in preventing the onset of a genetic disease in an otherwise affected individual through providing for broader cellular coverage. Putative alternatives such as in vitro fertilization and preimplantation embryo screening reduce the risk of disease through avoiding the birth of an affected individual, rather than in preventing the onset of disease. Juengst is critical of the ease with which the therapeutic use of GCGT to prevent the expression of a genetic disease in an individual (so-called phenotypic prevention) is equated with the use of GCGT to prevent the transmission of the genetic disease to future generations (so-called genotypic prevention). Juengst believes that it is this confusion over the purpose of germ cell gene transfer that fuels concerns about the eugenic use of this technology (18).

Juengst identifies four problems that geneticists face in adopting as a goal the prevention of certain genotypes (18). First, genotypic prevention, Juengst argues, understands the diseases it prevents as being caused by the associated gene abnormality rather than at the level of pathophysiological expression. He implies that this is a limited metaphor by which to understand genetic disease; however, those disorders that are amenable to GCGT techniques necessarily may fit within this model of causality. Nevertheless, the deterministic causal assumptions on which genetic therapy relies may be overly simplistic and subject to the risk of both false negative and false positive predictions of clinical disease.

Second, the decision to prevent the birth of an individual affected by a certain disease, Juengst points out, assumes that the burden of living with the disease outweighs any other value that the individual may experience or bring to the life of the family or community (18). However, the decision to risk GCGT in order to avoid the burden of genetic disease relies on the same calculus. Furthermore decisions to limit or withdraw life-sustaining treatment may be predicated on the value judgment that death (or nonexistence) is preferable to life under certain conditions. Our ambivalence in applying this same calculus to pre-implantation embryo selection may rest in the conceptual difficulties associated with so-called wrongful life and the active nature of the selective intervention. Beyond the stigmatization of those affected with a genetic disease is the prejudicial impact on carriers, either on the parents of those who undergo correction or on those whose parents either could not afford or who chose not to undergo carrier correction.

Third, Juengst is concerned that the traditional commitment to an individual's voluntary reproductive choice may give way to the economic and public health interests of society if medicine endorses "genotypic prevention." Finally, Juengst is concerned that the definition of "pathological genotypes" will inevitably be influenced by "larger cultural ideologies and social values" such as contemporary concerns to prevent either "reproductive anxiety and interpersonal aggression" (18). However, the need to define the domain of genetic conditions deemed suitable for the use of GCGT techniques remains a problem, even if geneticists eschew the professional goal of genotypic prevention. Unless individual requests are given unfettered access to genetic technology, some social and political definition of warranted requests will need to be established. The conceptual difficulties in defining a disease so as to exclude inappropriate requests for, say, genetic enhancement remain salient. Phenotypic prevention can be for the purpose of treating disease or enhancing human characteristics. Genotypic prevention can be aimed at the elimination of diseases or the selection of desirable characteristics in the population as a whole. Thus the distinction

between treatment and enhancement remains a problem independent of whether we accept or reject the goal of genotypic prevention.

TREATMENT AND ENHANCEMENT: A ROLE FOR POSITIVE EUGENICS

With the acceptance of somatic cell gene therapy and the continued discussion of germ cell gene therapy, Fletcher and Anderson observe that the moral differences between these two forms of gene transfer technology appear less significant than the distinction between the treatment of disease and the enhancement of human characteristics. Reflecting the general condemnation of the use of genetic technology for enhancement, they argue for drawing a moral line, not between somatic cell and GCGT techniques, but between treatment of disease and enhancement of human characteristics (5). However, this distinction is difficult to maintain as concepts of disease and illness involve complex and subtle evaluations regarding the scope of medical interventions. Any attempt to distinguish between human needs (treatment) and desires (enhancement) also falters on the historical and cultural diversity of conditions that medicine in fact has treated. Finally, the concepts of disease and health incorporate moral and nonmoral values and goals, making it difficult if not impossible to discover an objective yet morally significant line between disease and various competing, positive notions of health (15). Still, the concept of malady has been put forward in an attempt to draw such a line.

A malady is defined as an existing condition that causes a person to suffer or risk suffering an evil such as death or disability in the absence of a distinct sustaining cause. A genetic condition that fulfills the definitional criteria of malady can be treated through gene transfer as negative, not positive, eugenics. Berger and Gert claim that the definition involves the avoidance of universal (and thus objective) evils, such as death, pain, disability, loss of freedom or pleasure, which are not dependent on particular cultures. The notion of a "sustaining cause" seems to hinge on a distinction between an internal or physiologic cause and external or environmental cause — whose removal leaves the individual at increased risk of injury or evil. A full discussion is beyond the scope of this entry. Suffice it to say that the concept of malady will not escape the problems associated with distinguishing between the treatment of disease and the enhancement of health. Notions of disease and disability are notoriously culture dependent (19); the notion of causality itself deeply intentional (20). The recent controversy over the use of recombinant human growth hormone to treat children with short stature who are not growth hormone deficient illustrates the difficulty in defining a disease so as to draw a line between treatment and enhancement (21).

To render this distinction more problematic, Torres presents an example of a somatic cell gene transfer treatment protocol which, strictly speaking, qualifies as an enhancement technique. Gene transfer techniques are used to enhance the natural resistance of hematopoetic stem cells to the effects of anticancer drugs in patients receiving chemotherapy for solid tumors. Given that a clinical intervention may or may not be justified given its purpose, Torres proposes that gene transfer techniques may be used for enhancement provided that "such enhancement constitutes a necessary condition for the success of treatment designed to suppress the causes, symptoms or effects of severe pathology." The enhancement is thus a means and not an end. Torres cites two reasons that should dissuade us from the use of gene transfer techniques for enhancement: the risks of toxicity, and the discrimination involved in unequal access to genetic technology and the devaluing of those left with the unenhanced trait (22).

It may be inevitable that the introduction of GCGT for the treatment of disease would, after further research and development, be applied to the enhancement of human traits. If the purpose of the enhancement is an accepted social goal, such as providing for minimal functionality given the design of social space (i.e., short stature), or if it's purpose is an individual goal (i.e., enhancement of height in order to procure a lucrative sports contract) that is equally available to all — is enhancement necessarily precluded? It would clearly be unjust and a violation of parental autonomy for governments to perform genetic enhancement without parental consent. However, parents have the prima facie right both to attempt to prevent disease and to offer any advantage or benefit to their children through human GCGT. Although this right could be overridden by the risk of harm to future generations or by grave social, political, and economic injustices, Resnik proposes that the goods achieved by enhancement can be regulated according to accepted principles of justice so as to not exacerbate social inequities. The challenge then will be to distinguish between legitimate and illegitimate enhancement, both conceptually and through appropriate social policy and regulation (2,23).

THE ANALOGY BETWEEN PRE-EMBRYO DISCARD AND ABORTION

Mauron and Thevoz suggest that society can be protected from the use of GCGT for illegitimate enhancement through reinforcing the patient or client-centered ethics of genetics (16). Similarly Juengst wants to preserve the traditional allegiance of geneticists to respecting the freedom of parental reproductive choices — an allegiance that is threatened by the acceptance of genotypic prevention as a social and professional goal. The professional obligation to support the personal goals of the patient, Juengst claims, has little to do with the content of those goals and can peacefully co-exist with the parental motivation to eliminate specific genetic diseases and thus genotypes from an individual family heritage (18). Whereas Munson and Thevoz assert that this traditional focus enables medicine to "keep its sphere of action technically and morally manageable," Juengst recognizes that this traditional allegiance to reproductive freedom does not resolve a number of ethical problems (16,18). For example, as genetic testing increasingly becomes a prelude to prenatal interventions to improve the health of the fetus, the fetus emerges as

a "patient" with its own associated moral claims that may be in tension with parental choice. In addition, as mentioned above, a simple allegiance to freedom of reproductive choice may result in a laissez faire genetic economy in the absence of social or professional limits to the range of offered genetic services — limits which may reintroduce the problem of defining genetic conditions which are deemed pathological (18). Finally, if Juengst wants to insulate GCGT from antiabortion arguments against embryo discard, an appeal to reproductive freedom is less than reassuring given the importance of this argument to establishing access to abortion in the first place. The argument for parental reproductive freedom and autonomy may obviate the need to benefit any particular embryo and undermine appeals against pre-embryo discard to the extent that parental autonomy implies ownership of gametes and pre-embryos (11).

The analogy, though disputable, between the discard of defective pre-embryos and the controversial issue of abortion may impede the development of GCGT. During the development phase of pre-embryo modification, even after extensive animal testing, it is unlikely that the procedures would be perfected such that all modified pre-embryos are appropriate for implantation. The inevitable presence of a defective pre-embryo, even if created with the intent to implant, raises the question of discard given the woman's voluntary and informed consent (or dissent) to have such a pre-embryo implanted. This dilemma does not appear to be different from current in vitro fertilization practices, or from the decision to abort a fetus that is determined to be genetically defective based on prenatal testing. Thus GCGT does not present any new or different problems with respect to embryo discard (or abortion) than currently exists. In vitro gametocyte modification, if technically feasible, would allow for in vitro fertilization procedures to proceed as currently supported, given that only modified gametes could be used to create a pre-embryo suitable for implantation. Our experience with the abortion debate over the past two decades would suggest that if moral and political opposition to GCGT is based on the potential destruction of living pre-embryos, a simple principled or political solution may not be forthcoming (2).

Juengst's proposal would allow GCGT techniques to be developed for the purpose of more effectively preventing the "onset of a genetic disease in a patient." It undermines the practical argument that GCGT is unnecessary given the existence of pre-implantation embryo selection by highlighting an essential conceptual difference between the discard of defective embryos and the treatment of affected embryos (18). However, although individuals who are against the discarding of the products of human conception may support this distinction, the discarding of human embryos may be a necessary component of any research program seeking to establish the safety and efficacy of such techniques (11).

Juengst also suggests that rejecting the goal of genotypic prevention would allow for the correction of an adult's carrier status through GCGT techniques involving gametocytes, while precluding pre-embryo modifications directed toward the same goal of eliminating an offspring's carrier status (18). Although this option is rejected as a form of genotypic prevention through limiting the reproductive freedom of a child to transmit a deleterious gene, it is unclear that such a decision is outside of the range of discretionary choices that parents may make on behalf of their children. It may be difficult to insulate GCGT techniques from supporting the goal of preventing the transmission of deleterious genes. Juengst's arguments depend on two questionable assumptions: first, that the professional morality of geneticists can be divorced from an analysis and critique of the consequences of parental action; second, that personal moral choices can be supported while rejecting the social, cultural and political context and implications of such choices. A more productive approach may be to address directly the fundamental issue underlying the affirmation of individual reproductive freedom, that is, the social and political control of genetic technology.

ALTERNATIVES TO GCGT

The alternatives available for individuals who are at risk for transmitting a genetic disease to their progeny include (*1*) selection of a reproductive partner who reduces or eliminates the risk, (*2*) selection of a reproductive mechanism that reduces risk such as donor gametes, artificial insemination, or pre-embryo selection and in vitro fertilization, (*3*) avoid procreation altogether, (*4*) prenatal testing and selective abortion, (*5*) somatic cell gene therapy for affected offspring, and, finally, (*6*) GCGT techniques.

Restricting access to GCGT technology in light of the alternatives of avoiding procreation or selective termination of pregnancy generally is believed to be an unacceptable infringement on reproductive choice (2,13). The ability to conceive, nurture a pregnancy, and give birth to a healthy baby is accepted as an aspect of reproductive health appropriately addressed by medical technology. Individual screening and then selection of appropriate reproductive partners is technically feasible. However, rarely do affairs of the heart submit to such a rational and premeditated approach. For many individuals, the acceptability of alternative approaches using surrogates and donor gametes founders on the desire to produce a genetically related offspring. For such individuals who knowingly bear a genetic defect, the only options, other than the identification and termination of defective pre-embryos or fetuses, are germ cell and somatic cell gene transfer techniques.

Current in vitro fertilization techniques involve the discard of pre-embryos that are not suitable (for whatever reason) for implantation. Once we are able to identify pre-embryos that carry a certain genetic defect, we will be able to selectively eliminate these pre-embryos prior to implantation. Such a procedure may be technically and morally preferable to prenatal diagnosis and selective abortion, primarily due to lower maternal risk; however, it is conceptually no different from the point of view of a pre-embryo or fetus in the absence of any ontological distinction between the two. The development of gamete selection coupled with in vitro fertilization and implantation may eliminate the conceptual link with

abortion. However, there would remain rare instances such as couples who are homozygous for a genetic disorder where their desire for a genetically related yet healthy offspring could not be met in the absence of GCGT technology.

GCGT IN THE TREATMENT OF MITOCHONDRIAL DISEASE

Recently Rubenstein and colleagues proposed a protocol involving the treatment of genetic diseases associated with defects in mitochondrial DNA. The inheritance of mitochondrial DNA is strictly maternal, given that the cytoplasm of the ovum contributes most if not all of the mitochondria incorporated into the developing embryo. In effect, the procedure they propose transfers the nucleus of the carrier's ovum into the enucleated cytoplasm of a donor ovum, followed by standard in vitro fertilization and implantation. The procedure thus involves the transfer of genetic material into a germ cell for the purpose of correcting both the phenotypic expression and the vertical transmission of an otherwise debilitating and potentially fatal disease (4).

Although Rubenstein and colleagues present a complex and less than compelling classification of GCGT based on the level of cellular penetration, the essence of their argument hinges on an ethical distinction between manipulation of mitochondrial DNA and nuclear DNA (3). The protocol appears to satisfy the arguments in favor of GCGT, that is, medical utility and necessity, prophylactic efficiency, respect for parental autonomy and the pursuit of scientific knowledge within the bounds of ethical research. The protocol also renders the arguments against GCGT less compelling (24). Preserving the integrity of nuclear DNA may reduce the scientific uncertainty and risks to future generations inherent in other forms of GCGT, while avoiding concern about the putative right of an individual to an unmodified genetic endowment. The technology appears feasible and likely not to generate significant costs beyond those currently associated with some in vitro fertilization techniques. Finally, it is unlikely that the mitochondrial DNA will be amenable to enhancement. The development and application of GCGT techniques involving mitochondrial DNA may lead to a greater acceptance of germ cell manipulations for the treatment of inherited disease. However, given the special nature of mitochondrial DNA, it is unlikely that such acceptance will impact significantly on the debate concerning other forms of GCGT.

THE SOCIAL AND POLITICAL CONTROL OF GENETIC TECHNOLOGY

The ability to prevent the illegitimate use of GCGT technology will depend, not on the strength or weakness of such conceptual distinctions as the treatment of disease versus the enhancement of human traits, but on how we conceive of and establish the social and political control of genetic technology. The right to be free from harm, though not to be left alone, is the right of children with respect to their parents, not with respect to the state. The reproductive freedom of parents is asserted as a limit to state intervention. However, an appeal to individual freedom in the application of genetic technology neglects issues in the social control over development of this same technology. For example, the choices of parents who are at risk for transmitting either cystic fibrosis or sickle cell disease will be shaped by the resources committed to the investigation and development of these same alternatives — resources committed through a complex process of political advocacy, community activism, and private marketing.

The therapeutic application of somatic cell gene transfer, when combined with genetic diseases that may remain resistant to somatic cell approaches, likely will create moral and political pressure for the incremental development of GCGT techniques to address these diseases. As we gain experience with genetic technology, the development of GCGT protocols could proceed within the context of federal research guidelines assuming appropriate modifications to allow for the involvement of pre-embryos. The ethics of GCGT technology reduce to the ethical issues of the creation and use of pre-embryos. In the development phase, the discard of pre-embryos is likely despite attempts to conceptually link GCGT with the treatment of individual pre-embryos and thus divorce it from the issue of pre-embryo discard and abortion. However, in the application phase, it is likely that the availability of GCGT technology would greatly reduce the use of either pre-embryo selection or selective abortion for the prevention of genetic disease.

BIBLIOGRAPHY

1. M. Lappe, *J. Med. Philos.* **16**(6), 621–639 (1991).
2. B.K. Zimmerman, *J. Med. Philos.* **16**(6), 593–612 (1991).
3. M.D. Bacchetta and G. Richter, *Cambridge Q. Healthcare Ethics* **5**(3), 450–457 (1996).
4. D.S. Rubenstein, D.C. Thomasma, E.A. Schon, and M.J. Zinaman, *Cambridge Q. Healthcare Ethics* **4**(3), 316–339 (1995).
5. J.C. Fletcher and W.F. Anderson, *Law, Med. Health Care* **20**(1–2), 26–39 (1992).
6. 45 *Code of Federal Regulations* 46 (1983).
7. B. Steinbock, *Life before Birth: The Moral and Legal Status of Embryos and Fetuses*, Oxford University Press, New York, 1992.
8. T.H. Murray, *The Worth of a Child*, University of California Press, Berkeley, 1996, pp. 96–114.
9. Section 512 of the FY 1997 Appropriations Act, Public Law 104–208 (September 30, 1996).
10. J.C. Fletcher and G. Richter, *Hum. Gene Ther.* **7**(13), 1605–1614 (1996).
11. K. Nolan, *J. Med. Philos.* **16**(6), 613–619 (1991).
12. R. Munson and L.H. Davis, *Kennedy Inst. Ethics J.* **2**(2), 137–158 (1992).
13. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life: The Social and Ethical Issues of Genetic Engineering with Human Beings*, U.S. Government Printing Office, Washington, DC, 1982.
14. E.M. Berger and B.M. Gert, *J. Med. Philos.* **16**(6), 667–683 (1991).
15. H.T. Engelhardt, *Soc. Philos. Policy* **13**(2), 47–62 (1996).

16. A. Mauron and J.-M. Thevoz, *J. Med. Philos.* **16**(6), 649–666 (1991).
17. W. Glannon, *Bioethics* **12**(3), 187–211 (1998).
18. E.T. Juengst, *Hum. Gene Ther.* **6**(12), 1595–1596 (1995).
19. A. Kleinman, *Writing at the Margins: Discourse between Anthropology and Medicine*, University of California Press, Berkeley, 1995, pp. 19–92.
20. H. Putnam, *The Many Faces of Realism*, Open Court, LaSalle, IL, 1987, pp. 23–40.
21. R.M. Nelson, in J.P. Kahn, A.C. Mastroianni, and J. Sugarman, eds., *Beyond Consent: Seeking Justice in Research*, Oxford University Press, New York, 1998, pp. 56–57.
22. J.M. Torres, *J. Med. Philos.* **22**(1), 43–53 (1997).
23. D. Resnik, *J. Med. Philos.* **19**(1), 23–40 (1994).
24. E.T. Juengst, *J. Med. Philos.* **16**(6), 587–592 (1991).

See other GENE THERAPY entries; HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY: OVERVIEW.

GENE THERAPY, ETHICS, RELIGIOUS PERSPECTIVES

GERALD P. MCKENNY
Rice University
Houston, Texas

OUTLINE

INTRODUCTION

Theologians were among the first to address ethical issues in genetic science, and religious bodies and individual theologians have contributed to debates over the ethics of gene therapy from the beginning (1). Their contributions are important for several reasons. First, religious ideas inevitably play a role in shaping public attitudes to gene therapy. It is therefore important for those who have research, commercial, or policy interests in gene therapy to be informed of religious responses to the latter. Second, there is a common perception that religious traditions are hostile to, or suspicious of, genetic technology. This perception is fueled by a few instances in which religious leaders have taken high-profile stances against research on germ-line gene therapy or the patenting of genes, and by the use of religious or quasi-religious language by certain opponents of genetic technologies. However, as the following survey shows, the index of support for somatic cell gene therapy (SCT) among the vast majority of religious groups and writers is very high, while the range of responses to germ-line gene therapy (GLT) generally tracks that of the informed public as a whole. Third, many people who work in gene therapy or related areas adhere to a religious tradition and may consider it important to know what that tradition says about their work. Finally, many religious responses to gene therapy contain arguments or insights that are missing in secular debates.

CHRISTIANITY

Ecumenical

The World Council of Churches (WCC) and the National Council of the Churches of Christ in the USA (NCC) each produced multiple documents during the 1980s that addressed gene therapy along with other issues in genetics and biotechnology. With one exception (2), none of these documents represent the official position of the body in question, though all of them received official recognition at some level.

World Council of Churches. The WCC published two reports on genetic technology during the 1980s (3,4). The reports share a focus on science and technology as forms of power as well as knowledge, and emphasize social, political, economic, and ideological factors in biotechnology as a global enterprise. While the specific conclusions regarding gene therapy are conventional, attention to these factors affects the analysis of the broader context in which gene therapy is or will be carried out.

SCT is regarded as no different from other forms of experimental therapy; accordingly, it should be undertaken only in the absence of adequate alternative treatments and carried out under the usual protections governing research with human subjects. GLT is more problematic insofar as germ-line interventions involve alterations that will persist over many generations. Of course, the effect on future generations can be a strong argument in favor of GLT: "By overcoming a deleterious gene in future beings, the beneficial effect of such changes may actually be magnified" (3). However, weighing these potential benefits against risks would require extensive knowledge of long-term consequences that we do not yet possess; research on GLT should therefore be banned at present. But this ban need not be permanent; the report that advocates it goes on to call for ethical reflection leading to future guidelines (4).

The earlier of the two reports exhibits a tension found in other religious statements. On the one hand, the justifiability of gene therapy seems to be connected with its use for "recognized diseases of genetic etiology." On the other hand, there is considerable skepticism about whether reliable lines can be drawn between therapy and enhancement ("Correction of mental deficiency can move imperceptibly into enhancement of intelligence, and remedies of severe physical disabilities into enhancement of prowess") or between negative and positive eugenics ("There is no absolute distinction between eliminating

'defects' and 'improving' heredity"). There may be agreement that certain conditions "are so deleterious as to deserve first call" on gene therapy resources, but beyond these cases choice "becomes a matter of subjective judgment which will differ according to personal values."

In place of drawing lines between negative and positive eugenics or therapy and enhancement, both WCC reports raise other ethical issues. These issues all involve power and ideology as factors that arise in the social, political, economic, clinical, and research contexts in which gene therapy is or will be carried out. One ethical issue is the role of social and cultural prejudices in identifying certain genes as "defective" or, more generally, the potential for gene therapy to result in discrimination or in eugenic policies that institutionalize prejudices (4). When genetically transmitted characteristics become societal liabilities, gene therapy may be used to alter those characteristics rather than society altering its values and prejudices (3). A second worry at the societal level is that resources devoted to gene therapy will divert attention and resources from (*1*) nongenetic diseases, (*2*) protection of genes from avoidable damage (from mutagens, carcinogens, and man-made radiation), and (*3*) providing each person with opportunities to develop their existing capacities (3,4). At the political level, genetic interventions should remain options and not requirements (legal or otherwise) that parents are obligated to fulfil in their offspring (3,4). A similar worry pertains to economic pressures to secure certain characteristics for one's children (4). Also in the economic realm, one report calls for legal safeguards to protect individuals and their potential descendents with regard to quality control of materials and methods used in gene therapy, and misrepresentation of possible benefits by commercial advertisers or by scientists (3). In the clinical context, there is concern that desperately ill patients might try unproved techniques of doubtful efficacy. Patient-subjects should therefore be fully informed of possible negative effects. The report does not address the current question of whether desperately ill people who indicate awareness of possible negative effects should be allowed to enroll in early phases of experimental trials.

A different set of criticisms denounces the mechanistic worldview of contemporary science and technology that, the report claims, objectifies life for utilitarian and instrumental ends and whose primary goal is "the maximizing of material advantages for those few most able to appropriate and profit from the extraction of the earth's resources" (4). This conviction appears to undergird a concern that genetic engineering reduces persons to interchangeable parts. This reduction threatens the "inalienable dignity" of persons, which is the basis of mutual respect (3). However, it is not clear what exactly is under condemnation. At one point the target is the transformation "of offspring into interchangeable parts to be selected at will" — the concern, echoed in many religious statements, that genetic knowledge and interventions will turn children into products — but this is followed by a more general criticism that genetic engineering in and of itself, as a form of knowledge and practice, "converts the human subject into a composite object of interchangeable elements." This latter concern would seem to render all genetic interventions suspect, including those directed at conditions the WCC itself would consider serious diseases. The report goes on to ask, "In what ways do we, by manipulating our genes in other than simple ways, change ourselves to something less than human?" Unable to draw lines between acceptable and unacceptable *kinds* of intervention, the WCC latches on instead to a concern with the *degree* of intervention.

Finally, both WCC documents call for a prohibition of research on zygotes and embryos, with exceptions for therapeutic purposes under well-defined conditions (4). Both the prohibition and the exception derive from the status of zygotes and embryos as potential but not actual persons.

National Council of Churches. The NCC produced three reports during the 1980s. While not ignoring the warnings highlighted by the WCC, the NCC urges churches to take a positive stance toward genetic science and technology and to participate in public debates, and asks scientists to accept public scrutiny of genetic research. The reports applaud the role of genetics in fostering awareness of the interconnection of human life with other forms of life and the responsibility of humans for other life forms, though these convictions are in some tension with theological views that emphasize the special status of human life (5).

The reports are primarily designed to foster discussion of genetic technology; they therefore identify advantages and disadvantages of gene therapy rather than issuing prescriptions or proscriptions. SCT is no different in principle from other experimental therapies, but it poses potential dangers: The host cell into which the transferred DNA integrates may produce too little or too much of the desired product; the transferred DNA may disrupt the functioning of existing cells (5). GLT departs from standard medical therapy insofar as alterations are passed on to future generations. This holds out the promise of reducing the frequency of deleterious genes in the population but also raises ethical concerns that accompany eugenics (5). These concerns apparently include the possibility of involuntary participation in eugenic research, compulsory treatment in the name of eugenics, elitism in the determination of desirable and undesirable characteristics, and encouragement of an illusory quest for human perfectability (2,5). GLT also has the advantage of offering a possible treatment for genetic diseases that affect multiple tissues, but the technical difficulties with gene expression and the risk of disrupting cellular functioning remain (5). Future use of GLT is not ruled out, but the unknown and uncontrollable risks require extreme caution, and the interests of future descendents may have to be represented by a guardian ad litem (though the report is silent on the question of what those interests are) (2,5).

The reports emphasize the legitimacy, in principle, of intervening into genetic processes for the betterment of human life (2,5,6). They refute, on theological grounds, the objection sometimes attributed (usually wrongly) to Christianity that genetic interventions ipso facto exceed proper human limits or violate a normative natural

order. However, the reports are consistently skeptical of efforts to identify and prioritize genetic conditions for possible intervention: Some so-called bad genes may serve beneficial purposes. Social and environmental factors, which often can and should be altered, make some conditions liabilities. Efforts to prioritize genetic conditions by placing them on a spectrum from trivial to serious are relative to social, economic, medical, and value variables (6). Rather than attempting to prioritize interventions, the NCC invokes human dignity and distributive justice to impose certain limits and requirements on any such interventions. However, the conception of human dignity is vague and inconsistent; its grounding is unclear [is it an *alien* dignity "conferred by God's love" or is it "related to human powers and to human transcendence over the rest of nature" along with "human reverence and human relations with the rest of nature"? (6)]. Human dignity functions primarily as a placeholder for four ethical concerns: the sanctity of life that prohibits deliberate distortion or destruction of human beings in genetic research, rejection of the notion that genetic health or normality is a criterion of human worth, affirmation of the possibility that some kinds of suffering can serve a purpose, and recognition of the limits of some kinds of control over nature. The report concedes that no firm prescriptions follow directly from these concerns, but insist that the latter "establish a context of awareness" (6).

The claims on behalf of distributive justice are more specific, though they are short on argument and attention to practical implementation. The reports question the development of procedures and products that, because of demand and cost, will benefit only a few; pose the problem of how to balance the current treatment needs of afflicted individuals with research that might someday cure these diseases; and express a concern that basic health care not suffer neglect due to the pursuit of "exotic techniques of genetic control" (5,6). The reports also call for the benefits of genetic technology to be made available to all "regardless of geographic location, economic ability, or racial lines," especially when the products of genetic research result from public funding, and strongly oppose the disparity in standards for the protection of human subjects in the case of products used in the United States but initially tested elsewhere (2,5,6).

Orthodox

Few Orthodox individuals or groups have directly addressed issues of gene therapy. Two exceptions are John Breck (7) and Demetrios Demopulos (8). The lack of official statements by patriarchs or bishops and the scarcity of work by individual thinkers makes it difficult to determine how representative these commentators are, though both approach genetic technology with characteristic Orthodox themes and concerns.

Both Breck and Demopulos welcome SCT but oppose GLT. Breck, citing the "unacceptable risk" of transmitting irreversible consequences of errors to future generations, calls for a moratorium on GLT research. Demopulos points to the unknown consequences of eliminating genes from the gene pool and the likelihood that the development of gene therapy techniques would involve the discarding of embryos (a concern that other Orthodox are likely to share). These objections do not seem to rule out the justifiability of GLT in principle or in perpetuity; they address risks that may someday fall within acceptable levels and moral wrongs (in the case of discarded embryos) that may eventually be avoidable. However, Breck and Demopulos set the knowledge conditions and the moral strictures on the processes by which GLT would be developed very high. Even if GLT is eventually acceptable on their terms, its process of development will almost certainly have been unethical.

Breck and Demopulos both begin with the characteristic Orthodox view of humanity as the "icon of God." "Human nature in this sense is a process of moving toward the Archetype which is Christ incarnate" (8). As such, humanity is also a microcosm of creation and the link between God and the rest of creation. The purpose of humanity is "to proceed toward union with God and achieve ontological actualization, and to bring the rest of creation with it" (8). These convictions might at first seem to support ambitious efforts at genetic enhancement and eugenics, which might be understood in terms of ontological actualization or cooperating with God's "intent to transfigure the cosmos" (7). However, Breck and Demopulos sharply reject this interpretation. Breck supplies the theological reason: *theosis*, or union with God, is not achievable through genetic means but only "through a process of continual repentance and the free exercise of moral choice" that permits the practice of virtue. Breck therefore distinguishes "therapeutic" from "innovational" interventions. The latter (which appear to encompass enhancements and positive eugenics) are unable to produce the characteristics (repentance and moral choice) that truly matter while, if ever successfully developed, they would likely be used for traits associated with enhancing competitiveness, which Orthodox Christianity would consider suspect. In addition their development would almost certainly violate moral norms, including moral limits on the treatment of embryos. On the other hand, Breck is surprisingly supportive of negative eugenics (though he would not support the use of GLT for this purpose or, presumably, involuntary measures). Demopulos also restricts the role of genetics in ontological actualization to the reduction of sickness, which would enable persons to live longer, giving them more time to pursue union with God by nongenetic means.

These analyses raise serious questions about the role of genetics in various characteristics and the distinction between therapy and enhancement. Regarding the latter, Breck admits that the line is unclear, but argues that since enhancement of character traits is, at best, far in the future, more pressing issues deserve primary attention. These issues include whether access to beneficial technologies will be limited to those who can pay for them, and how standards for research using human subjects will be set. To these concerns, Demopulos adds the question of the priority of gene therapy research relative to other medical needs. However, beyond identifying these issues as priorities, Breck and Demopulos do little in the way of analyzing or resolving them.

Roman Catholic

Catholic approaches to bioethical issues traditionally rely on natural law theories that claim validity apart from appeals to revelation or distinctively Christian theological claims. However, debates over the role of human experience and culture in interpreting natural law, whether natural law issues in absolute prohibitions, and the relation between natural law and virtue divide Catholic thinkers. Some of these debates are reflected in the following treatments of gene therapy.

Pope John Paul II. In the early 1980s Pope John Paul II issued two declarations on genetic technology (9,10). The declarations establish moral norms that no genetic interventions may violate and that cut across distinctions between SCT and GLT, between therapy and enhancement, and the like. These norms are derived from a broad notion of human dignity ultimately grounded in the biblical notion of humanity as "created in God's image, redeemed by Christ and called to an immortal destiny" (10). However, in accordance with Catholic natural law theory, these norms are knowable apart from biblical revelation. For clarity's sake they may be grouped under three headings: respect for life, human dignity in a narrower sense, and liberty. Respect for life entails the right to life "from the moment of conception to death" and status as an end and not a mere means to the collective good (10). This rules out genetic interventions that destroy embryos or subject them to experimentation (9). And, should they ever become possible, it rules out "manipulations tending to modify the genetic store and to create groups of different people, at the risk of provoking fresh marginalizations in society" (presumably because of their inferior status) (9,10). Dignity in the narrow sense refers to the integral unity of humanity as one in body and soul (9,10). This rules out genetic interventions that might make use of forms of reproduction that separate the procreative act from the biological and spiritual union of husband and wife (e.g., artificial insemination and in vitro fertilization). It also rules out interventions that would distort or destroy this integral unity, though it is unclear what kinds of interventions the pope has in mind (10). Finally, liberty is violated when a genetic intervention "reduces life to an object, when it forgets that it has to do with a human subject, capable of intelligence and liberty . . ." (10). Again, it is not clear what kinds of interventions or procedures would violate liberty in this sense.

These considerations lead to more specific conclusions about gene therapy, some explicitly drawn by the pope, others that may be inferred. The primary distinction is between therapeutic and nontherapeutic interventions, a distinction the pope assumes without elaboration. Both declarations express strong support for therapeutic interventions so long as they tend to improve one's overall condition. In principle, this could apply to both SCT and GLT. But the restrictions on treatment of embryos and forms of human reproduction would make most GLT research currently envisioned morally unjustifiable. Nontherapeutic interventions, whether somatic or germ line, must avoid the violations of human dignity and liberty identified above. They also must avoid racist assumptions and a materialist view of human happiness (10). This certainly stops well short of ruling out enhancements or positive eugenics; one can imagine circumstances in which such interventions would not create inferior beings, distort or destroy body-soul unity, objectify persons, or carry out racist and materialist attitudes. But the pope's final remark, contrasting "adventurous attempts aimed at promoting I know not what superman" with "salutary efforts aimed at correcting maladies, such as certain hereditary maladies," seems to cast a general suspicion on enhancements and eugenic efforts (10).

Catholic Bishops' Joint Committee on Bioethical Issues (U.K.). In 1996 a working party of this committee, consisting mostly of physicians and fellows of the Linacre Centre, published what is probably the most comprehensive statement on gene therapy from a religious perspective (11). The report analyzes gene therapy in light of fundamental questions about human nature and fulfillment, the role of medicine in promoting human fulfillment, and responsibilities for the genetic health of oneself and one's children.

Human nature, as in the papal declarations, is a unity of soul and body. Since the soul is the body's "life principle," a living human body is never without a human soul; personhood therefore coincides with being a distinct living organism, which usually begins at conception. In accordance with a recent natural law theory, human fulfillment consists in the pursuit of certain basic goods such as life, knowledge, and sociability (12). (Health is one of these basic goods but is also a condition for pursuing the other goods.) These goods are never to be deliberately attacked (though they need not be promoted in all situations). Hence "[t]he life and health of some may not be promoted by means of an attack on innocent others; for example, by means of destructive experimentation on human subjects." Since personhood begins at conception, this principle extends to research using human embryos. The role of medicine is to promote the basic good of health, defined as "the complex of *functional, goal-directed, psychophysical systems* . . . in the contribution they make to the good of the whole." Health is best promoted "through the normal channels of human *activity*, whether conscious or non-conscious"; medicine should intervene, in the form of prevention, cure, or palliation, only when a functional defect renders the normal means to fulfillment unavailable or unsatisfactory. Finally, while persons have some responsibility for their genetic health and that of their children, there are natural and environmental limits to the elimination of genetic disease, and often the most appropriate response to a genetic disorder will be social or environmental rather than medical.

These general considerations enshrine theories of natural law and of medicine that are controversial in themselves and not fully integrated with each other. However, they lead to specific judgments about SCT, GLT, and genetic enhancements. SCT is no different in principle from other forms of medical treatment. As such, it should be evaluated according to principles that govern other experimental therapies, including consent, independent review, proportionality between risks and

burdens of treatment and degree and likelihood of benefit, consideration of the risks to others (e.g., the risk to the mother from SCT performed on a fetus), and restricting initial use of SCT to cases of serious disease for which there is no satisfactory alternative treatment. There is a risk that SCT will have inadvertent germ-line effects, but many other therapies pose the same risk.

GLT is problematic for several reasons aside from current technical difficulties. First, its development and use, at least in early stages, would involve in vitro fertilization (which violates the respect owed to life in its transmission) and the destruction of embryos. However, these objections would not apply to GLT performed on ova or spermatagonia followed by normal marital intercourse [a conclusion also drawn by the Catholic Health Association (13)], to treatment of the pre- or postimplantation embryo in situ, or to the removal and treatment of the pre-implantation embryo followed by implantation. None of these germ-line interventions need involve IVF or the discarding of embryos. A second problem with GLT is the risks it poses to various existing and future persons. At the present stage of technology, the risk to the embryo would be considerable, as recent animal experiments demonstrate. GLT procedures involving IVF and/or therapy on ova pose risks to the mother (as would treatment of the embryo in situ or its removal and replacement, though these are not mentioned in this context). Finally, there are long-term risks of adverse effects on the germ line. These risks are still too great to consider GLT even in cases involving serious diseases. But what if technological advance reduces these risks sufficiently to justify GLT in cases of extremely serious conditions? Three problems would remain. First, the risks of GLT would still be significant and would apply to descendants — assuming they would have existed at all — who are not affected by the condition, and who therefore would be subjected to risks without having the condition that justified taking those risks. Second, GLT would be costly and would compete against other pressing medical needs. Since GLT would mostly affect those individuals who will exist only because GLT will have been available, it is harder to justify the cost (or the risks) over against the needs of individuals and their descendants who would have existed whether or not GLT had been available. Both of these objections raise significant questions about future persons that the report does not address. Third, GLT would almost certainly become safe and effective due to immoral research on human embryos. While it would not necessarily be immoral to make use of it at that point — parents could always request that no embryos be harmed in their case — appearing to condone the means by which a technique was developed and the witness to the sanctity of life entailed in refusing it would be significant factors to take into account. However, the report explicitly rejects the common European claim of a right to inherit an unaltered genome (14). Changes to the genome do not affect the uniqueness of the person any more than do changes to other parts of the body. The genome "like other parts of the body ... may *in principle* be altered, to cure some defect of the body." Indeed, assuming that the objections noted above could all be met, the possibility of eliminating a devastating disease from a family would in many cases be not simply a right but an obligation.

The report addresses nontherapeutic genetic interventions (enhancements) by distinguishing between "environmental" and "mechanical" interventions. The former involve "a mere *response* to *selected existing potential* of the child" and are "open-ended" in that they do not specify the exact characteristics or the degree to which the intervention will prove favorable. The latter, which include genetic interventions, involve "an *amendment* of existing potential" and "are something that *happens* to the child rather than something the child *does* in a certain environment." There are two arguments in favor of environmental over mechanical interventions. First, because health should, whenever possible, be promoted through the normal channels of conscious or nonconscious human activity (see above), there is a presumption in favor of the former. Second, while both types of intervention run the risk that parents will consider the child as a product or something they control, this is more likely in the case of mechanical interventions. Mechanical interventions, then, "would at least sometimes be unjustified, and conducive to further acts of parental manipulation." This stops short of a prohibition, but the report does not address questions of what circumstances would justify overriding the presumption against mechanical interventions, whether the latter are justifiable for conditions (e.g., short stature) that do not generally admit of environmental solutions, and what conditions should be candidates for enhancement at all. However, the report notes that if the common belief that strangers are not entitled to perform nontherapeutic mechanical interventions on children is justified, then germ-line enhancements would be immoral, since we are strangers with regard to our future descendants in a way that we are not with regard to our children.

Other. In place of pre–Vatican II natural law theories and the new natural law theory adopted by the Catholic Bishops' Joint Committee, Richard McCormick proposes a criterion for genetic interventions that relies heavily on human experience, asking of each proposed intervention whether it will "promote or undermine human persons 'integrally and adequately considered'" (15). This criterion "is necessarily inductive, involving experience and reflection upon it." Experience and reflection alone, however, could be appealed to in support of almost any conclusion; fortunately, McCormick identifies certain values that are meant to supplement, or perhaps specify, the criterion. One is the sacredness of human life, which opposes undue risks and especially discriminatory distribution of risks, and requires informed consent. McCormick does not discuss embryo research here, but elsewhere he argues for a presumption against the latter with exceptions approved "by an appropriate authority" (16). A second value refers to the interconnection of life systems, on which grounds McCormick rejects genetic interventions that accomplish short-term benefits at the risk of long-term harms. Since GLT risks eliminating deleterious genes that may have long-term beneficial effects, it is suspect on these grounds. Third, human diversity and individual uniqueness are

important aspects of the human condition that are threatened by some eugenic interventions. A related concern is that genetic enhancements could lead us to evaluate persons "not for the *whole* that they are . . . but for the *part* that we select." Finally, social responsibility requires distributive justice in both research priorities and access to medical benefits. These values, however, are still quite general; it is not clear how one moves from them to judgments regarding specific interventions and policies.

James Keenan criticizes the dominance of concerns about rights in discussions of the ethics of gene therapy (17). "We must ask not whether we have a right to enter these areas or not. We must ask what type of people we could become by entering into any of these areas." Keenan's major concern is the potential of gene therapy for objectifying persons. This potential apparently resides in a combination of reductionism and the mode in which genetic technology intervenes into evolution. While humans have always intervened into evolution, genetic technology does so from within human nature rather than from without; just as the process of directing nature has objectified external nature, so genetic interventions will objectify the human subject. Keenan does not show how this kind of objectification is more significant than more familiar interventions that also work "from within," ranging from psychotherapy to ascetic practices, nor does he question the viability of the idea of a free subject that underlies his analysis. Instead, he describes a progression of objectification: SCT objectifies the disease, GLT collapses the distinction between person and disease, enhancement takes the genotype itself — not simply correction of a disorder — as its object, while eugenics aims precisely at the objectification of the genome so that the person is an object before being a subject. GLT also risks the objectification of parenting (because it is difficult to see a gamete or zygote as more than an object) and research (because consent cannot be obtained). Keenan does not argue that the threat of objectification renders gene therapy unjustifiable; rather, the moral challenge "will be the creation of conditions in which the person, though objectified, is not solely treated as an object." Despite the problems with the category of objectification and its applications, Keenan's central question — what kind of people could we become through genetic interventions — opens up a promising line of inquiry that too few commentators, whether religious or secular, have followed.

Protestant

Anglican. The Episcopal Church offically adopted a brief resolution on gene therapy in 1991 stating that there is no theological or ethical objection to gene therapy if proved effective without undue risk, if aimed at "prevention or alleviation of serious suffering," and if benefits are available to all who need them for these purposes (18). None of these conditions are elaborated in any detail, and "serious suffering" is not defined. In 1992 the Governing Body of the Church of Wales commended to the church a report on genetic screening and therapy written by the church's Division of Social Responsibility (19). The report finds no objection to the use of SCT to correct serious genetic defects for which there are no alternative cures, but supports a 1988 statement by a group of European medical research councils (20) that GLT should never be carried out (presumably because of its unknown consequences to future generations, though this is not clear). Finally, the report rejects any "attempt to manipulate the human genome for other than therapeutic reasons. Any proposal to engineer particular traits or characteristics in human beings should be rejected as frivolous and regarded as unethical." No effort is made to distinguish enhancement from therapy or to address hard cases that may not be readily assignable to either category. No reason is given for the opposition to enhancements, though the first of three concerns the report raises about genetics in general, namely the concern that parents will choose characteristics of their children based on individual whim, could be one such reason. The other two concerns — that the state will require that parents carry out genetic interventions on their children, and that discrimination against or diminishment of respect for those who will continue to be born with congenital anomalies despite genetic interventions will occur — would remain even if, as the report recommends, gene therapy were restricted to serious diseases.

Evangelical/Holiness/Adventist. Ethical issues of gene therapy have been addressed chiefly by individual evangelicals rather than by churches or organizations. If those individuals are representative, evangelicals may be less interested than some others in drawing lines between different kinds of genetic intervention. In place of distinctions between SCT and GLT or between therapy and eugenics, John Feinberg proposes a line between gene therapy "to fight something in human beings that is clearly a result of the consequences of sin and living in a fallen world" and gene therapy to alter what is simply part of the diversity of creation (21). Feinberg classifies cystic fibrosis, Huntington's chorea, Parkinson's disease, and other physical and psychological conditions under the former category, while hair color, skin color and left-handedness fall under the latter. In effect, Feinberg simply redefines the distinction between diseases and traits as a distinction between the effects of sin and genetic diversity. Genetic interventions designed to alleviate the former are in principle permissible, though the standard moral and scientific preconditions for performing any medical intervention may supply reasons not to perform them in a given case or in general (as is the case with germ-line interventions until they are proved to be safe and effective). What about the latter — is it permissible to change one's hair color, for example, even though this is a matter of human diversity rather than sin? It depends on one's motive. Belief that certain of these traits are inferior to others, that changing them will increase one's own value as a human being, that those who possess other such traits are less valuable, or that everyone should possess a certain trait — all of these motives are immoral according to Feinberg. But there are problems with this approach. First, as Feinberg concedes, it is difficult to determine whether some traits, such as aggressive behavior, are the results of sin or diversity.

Second, while certain characteristics — skin color or body shape, for example — are from Feinberg's standpoint due to genetic diversity and not sin, the discrimination some people face due to these traits is the effect of sin. Would Feinberg permit one to fight this consequence of sin by changing the trait? Third, one may question on theological grounds whether it is always justifiable in principle to fight the effects of sin. If, as Feinberg himself believes, death is a consequence of human sin, should Christians support the use of genetic technology to attain immortality? Like Feinberg, James Peterson questions the line between SCT and GLT (22). In practice, some somatic cell interventions (e.g., keeping those with deleterious genes alive long enough to reproduce) could have germline effects, while ethically, the generational factor in GLT brings not only greater risks but also greater potential benefits. Peterson also questions the line between therapy and enhancement, arguing that if one proposes a normal range of human functioning as normative, there is no reason for excluding some characteristics and functions from being brought into that range and, indeed, no reason for not trying to increase the normal range itself. Instead Peterson proposes five criteria any genetic intervention should meet. They should be (*1*) incremental in order to minimize the degree and extent of unanticipated harms, (*2*) choice-expanding in the sense of not limiting one to a particular kind of life, (*3*) parent-directed in order to decentralize choice and thus avoid large-scale eugenic programs, (*4*) kept within societal boundaries that set minimum conditions that parents must meet with regard to their children but allow flexibility within those conditions, and (*5*) carried out by acceptable means, namely with limited risks and as noninvasive as possible. Peterson recognizes the limits of this approach: Parents, in seeking advantages for their children, and corporations, in promoting the use of genetic technologies, will ignore these limits. Nevertheless, because of their potential benefits, humans are responsible to God to pursue genetic interventions — "[e]ven some instances of germline enhancement" — within the five conditions.

The Church of the Nazarene, in the holiness tradition of American evangelicalism, included in its Manual from 1993 to 1997 a statement approving of gene therapy for the prevention and cure of disease but opposing "any use of genetic engineering that promotes social injustice, disregards the dignity of persons, or that attempts to achieve racial, intellectual, or social superiority over others (Eugenics)" (23). The statement does not articulate what is meant by social injustice or human dignity.

The Seventh-Day Adventist Church adopted a document on genetic interventions in 1995 (24). The general justification for gene therapy is theological: The genetic endowment of Adam and Eve was perfect; hence gene therapy is welcomed as a form of cooperation with God in recovering more of the original condition of creation and alleviating the results of sin. However, gene therapy must be carried out in accordance with Christian principles. These include (*1*) the rejection, given current knowledge, of GLT on the ground that it could affect the image of God in future generations; (*2*) the exercise of "great caution," in light of human sinfulness, the possibility of abuse, and unknown biological risks, in attempts to modify physical or mental characteristics in persons free of genetic disorders; and (*3*) the availability of the benefits of genetic research to all who need them. The document does not tell how to distinguish therapy from enhancements, nor does it explain why genetic interventions affecting future generations, but not those affecting present individuals, threaten the image of God.

Lutheran. Lutheran theologians in the United States and theologians and church groups in Europe together constitute a spectrum of theological and ethical evaluations of GLT. Among Americans, Ted Peters has called for "keeping the door open" to GLT and eugenics (25). Peters argues that if God's creative activity is understood as giving the world a future and humanity is understood as a "created co-creator," then "ethics begins with envisioning a better future." We should therefore keep open the possibility of improving the genetic makeup of the species. Peters rejects the various arguments against GLT and eugenics: Unforeseen consequences to future generations is no reason to prohibit GLT but rather to proceed in accordance with growing knowledge of those effects; eugenics can be dissociated from historical abuses; prejudice and discrimination against those who fall short of standards of genetic perfection already occur without GLT and need not accompany GLT. Peters addresses another criticism made by some European Lutheran theologians and by nonreligious writers in the United States such as Jeremy Rifkin (26) and Robert Sinsheimer (27): neither nature in general nor DNA in particular is sacred or represents, in its present form, God's final plan for humanity; rather, because nature is created ex nihilo, it has no ultimacy or sanctity in itself, while because God continues to create (*creatio continua*) nature as it is has no normative status (28). By describing creation entirely in terms of future and novelty and human co-creating entirely in terms of technology, Peters gives the impression that his theology simply reinscribes the modern narrative of progress. Also Peters offers no guidance regarding what future possibilities, aside from the treatment of fatal diseases, are and are not worth pursuing. With one notable exception, he also says little about the means by which they may be pursued. The exception concerns embryos. The latter may not be persons in the full sense, but they possess a moral status that makes genetic manipulation of gametes prior to fertilization morally preferable to genetic manipulation of zygotes (29).

In contrast to Peters, Gilbert Meilaender opposes GLT (30). Meilaender concedes that GLT could be therapeutic in intent and effect, and that it could spare future generations serious problems. This last feature, however, is precisely the problem; GLT exercises control over future generations. "Such interventions would aim ... at shaping the nature of others still to come. Not only a human being but humankind is then the object of our intervention." Medicine, apparently, should focus on the person with the disease, neither eliminating the person (as with selective abortion) nor eliminating the disease from humanity as a whole. However, Meilaender does not indicate why medicine should be judged differently from

other interventions (e.g., those involving public health and the environment) that aim at humanity as a whole. In any case, if SCT but not GLT is justifiable, can we distinguish between treating diseases and enhancing traits? Meilaender's definition of diseases as "disorders that bring pain or hinder an individual in carrying out the biological functions necessary for personal or species survival" is admittedly narrow; more promising is his suggestion that in place of such a line we cultivate "a renewed sense of the mystery of the human person and the limits to our own efforts at shaping and transforming character" together with the virtue of love "that in its open-hearted acceptance of an other disciplines and restrains the urge to transform and remake." Meilaender worries that without these attitudes and virtues children will become products — made rather than begotten; with these attitudes and virtues, and the discernment that comes with them, there is little to fear from SCT. But why, if these attitudes and virtues are capable of guiding the use of SCT, are they not also capable of guiding GLT? If GLT is ruled out because it exercises control over others (and not simply because it aims at humanity as a whole), then why not also rule out SCT performed on one's children? And if unqualified love and a sense of mystery are sufficient to prevent abuses of enhancements in the case of SCT, why are they not sufficient to prevent the same abuses in the case of GLT?

European Lutherans tend to side with Meilaender against Peters (31,32). A few isolated individuals view SCT as the first step toward breeding human beings or argue that placing the human genome at human disposal nullifies human dignity, but the vast majority support SCT provided that standard ethical conditions governing experimental therapies are met. GLT, however, is almost universally rejected, either permanently or in light of present knowledge and current (European) moral conventions. The reasons vary. Principled objections refer to the illicitness of embryo research and to the claim that GLT violates the genetic integrity of humanity (as humanistically inclined theologians argue) or that humanity precisely as it is, is created in the image of God (as more biblically oriented theologians argue). These latter sorts of claim seem to presuppose a genetic essentialism that most theologians and church groups attempt to avoid; moreover they raise the question again of why the germ line, and not somatic cells, is the locus of human integrity or the image of God. A second set of objections rules out GLT because of certain problems inevitably connected with it, namely the impossibility of drawing a line between GLT and eugenics or the impossibility of determining, in the final sense GLT implies, what are healthy and diseased genes. A final set of objections could be overcome by future developments. These refer to the unreliable results of animal research at present, to the consequences errors in GLT may have for descendants, and to the relative risks and cost of GLT.

Methodist. In 1992 the United Methodist Church adopted a comprehensive report on genetic science that endorsed SCT for the alleviation of suffering caused by disease, opposed GLT until its safety and certainty of its effects can be demonstrated and its risks to human life shown to be minimal, and opposed the use of gene therapy for eugenic purposes and for enhancements designed only for cosmetic purposes or social advantage (33). These conclusions are listed without supporting arguments, but the report as a whole provides reasons that at least partially support most of them. Human beings are understood as stewards of God's creation. This role emphasizes the sustaining of creation, which allows for enhancing creation but also requires acknowledgement of limits to human creativity and power. Genetic diversity reflects the goodness of creation and therefore must be preserved, while the unity of humanity in creation and in Christ rules out discrimination based on biological factors and requires recognition of the worth of the most defenseless. These convictions could serve as arguments against ambitious eugenic programs. The report cites several specific concerns regarding gene therapy. Three of these — the danger to individuals from experimental procedures, unanticipated adverse effects of combining genes from different species, and the larger numbers of people who are helped by gene therapy but who may be carriers of genetic diseases that are difficult or expensive to treat — apply to SCT as much as, or more than, to GLT. The other three — the long-term effects on the species, the unanticipated long-term health and genetic consequences of genetic enhancement, and the vision or goal that governs efforts to control evolution — apply primarily to GLT and/or efforts at eugenics and enhancement. Nevertheless, the report does not indicate why these concerns outweigh the potential benefits of GLT, enhancements and eugenics, but permit SCT.

J. Robert Nelson, a United Methodist theologian, supports SCT but proposes a present ban on GLT (34). Nelson is not opposed to GLT in principle — he recognizes its therapeutic potential, first for the organism itself, and then for its progeny — but only because of our insufficient knowledge, at present, of long-term generational consequences. Like most Christians, Nelson rejects any notion that the genetically unmodified person is normative; gene therapy cannot be rejected on grounds that it is unnatural.

Reformed. Both the Presbyterian Church, U.S.A. and the United Church of Christ have approved statements on genetics that include brief treatments of gene therapy. In 1990 the Presbyterian Church's General Assembly resolved to "[s]upport the discovery of new genetic knowledge that can improve the treatment and eradication of disease ..." (35). In 1983 the General Assembly approved a report that supported the potential of genetic research for "relieving suffering and enhancing life" but warned against the "threat of idolatry in the search for the 'perfect human being' ..." and concluded that "[t]he pursuit of 'superior' human beings through genetic manipulation should be explored only with great caution, if at all" (36). No elaboration or arguments clarify or support these declarations. The United Church of Christ statement, approved in 1989, is marginally more substantive, approving of SCT and noting that GLT may have unforeseen consequences that preclude it now, though future developments may alleviate this problem (37).

Support for gene therapy in principle is grounded in human covenantal responsibility to participate in God's creative and redemptive work, as made known in the healing ministry of Jesus.

In 1995 the General Assembly of the Church of Scotland welcomed a report written by a study group of the church's Board of Social Responsibility (38). The report, which focuses on genetic diagnosis, screening and therapy, is perhaps the most restrained of any religious treatment of gene therapy. Its authors believe most genetic interventions, including SCT beyond very limited uses and any use of GLT, are highly unlikely; rather than analyzing ethical issues they consider wildly hypothetical, they give reasons why they believe the techniques that pose such issues are so improbable. These reasons refer to technical obstacles, moral and regulatory requirements regarding research using human subjects, costs, and the presence of less risky (and less costly) alternative interventions (e.g., in vitro fertilization followed by selective implantation or fetal testing followed by abortion as alternatives to GLT). Not all of these reasons are convincing (e.g., GLT would be a superior alternative for those opposed to selective implantation or abortion on moral grounds), but all of them have been raised by other individuals and groups as well. In general, the report seems to endorse the conclusions of the Clothier Report of 1992 (39), which supported SCT under the current regulations governing experimental therapies and opposed GLT for the present. However, it also includes sociologist Margaret Stacey's criticism of the Clothier Report for ignoring the cultural context in which "genetic manipulation will itself inevitably change perceptions and beliefs about what it is proper for individuals to ask others to do to them or their children. . . ." Stacey's point, that technology in its cultural context generates new obligations and ideals, offers a promising line of inquiry which religious analyses of gene therapy have so far ignored.

Ronald Cole-Turner takes issue with most theological responses to genetic technology for leaving unclear the moral status of genetic disease and its cure (40). First, they do not resolve the question of whether genetic illness is natural or a defect of nature; they therefore have no basis for determining whether God wills illness or its cure. Second, they emphasize creation rather than redemption. As (created) co-creators, human beings are authorized to explore novel genetic combinations, but the question is whether they are permitted to identify and correct genetic defects. The latter requires a view of redemption as the restoration and reordering of creation. Cole-Turner seeks to rectify these problems. Nature is good, but also disordered: "A gene is identified for research and possible therapy because it causes human suffering. But it is regarded as a genetic defect because it is taken as a manifestation of the moral disorder of nature in reference to the intentions of the Creator. . . ." This gives gene therapy a moral ground: "That which is defective is that which may be changed or altered. Indeed, altering it would be seen as an act of participation in the redemptive work of God." What genetic interventions, then, may be welcomed as redemptive? Cole-Turner has no satisfactory answers to this question. He does, however, refer to the healing ministry of Jesus as paradigmatic of redemption. The conditions into which Jesus intervened — Cole-Turner identifies skin diseases, neurological conditions, and mental disorders — indicate what kinds of conditions are contrary to the purposes of God. And Jesus's special concern with the weak, the sick, and the poor, reversing natural selection by favoring the retention of their genes, indicates what research priorities and marketing arrangements should, from a Christian perspective, govern the development of gene therapy. But these claims are open to criticism. First, why is the ministry of Jesus the sole paradigm of redemption? Second, how strictly does Cole-Turner want to interpret this ministry? If Jesus healed only these sorts of diseases and restricted his healing to individuals whom he encountered more or less directly, does this rule out enhancements and GLT, respectively?

Judaism

Nearly all practicing Jews accept the normative status of the law, or *halakhah*, as given in the Talmud and in the commentaries, codes, and responsa that constitute the rabbinic tradition. However, there are important areas of disagreement. One disagreement is over how strictly to interpret the law. Another concerns whether there is an ethic outside the *halakhah*, and if so, what is its content, and how is it related to the law. Thus far the most detailed treatments of gene therapy have come from Orthodox commentators such as Barry Freundel, Azriel Rosenfeld, and Fred Rosner. Because they largely agree with one another and their views overlap, they will be discussed together, along with other Jewish thinkers whose less direct treatments of gene therapy supplement the work of the three Orthodox commentators. However, this should not give the impression that there is a single Jewish position on gene therapy or that, as other Jewish thinkers begin to address this issue, the consensus among these commentators will necessarily hold up.

Judaism brings to gene therapy a tradition of strong encouragment of therapeutic interventions, based on the importance of saving life, that amounts to a justification of gene therapy in principle. Arguments that genetic engineering falls under a class of illicit alterations of nature (along with sowing diverse seeds, mating different kinds of animals, mixing certain fabrics) are not unknown in Judaism, but Rosner's counterargument — that genetic engineering is a permissible alteration because it falls within the physician's divine license to heal — reflects a nearly universal view (41). Indeed, "even the most conservative Orthodox thinking provides no support for the view that such genetic manipulation would be an unallowable 'tampering' with nature" (42). Like other interventions that save or improve human life, gene therapy falls under the divinely ordained task of *tikkun olam*, namely healing or repairing and perfecting the world (43). In principle, then, gene therapy is permitted at all stages, including preconception, in utero, or following birth. SCT is judged by the same risk–benefit criteria used to judge other medical procedures, though a slightly higher risk factor may be tolerated for the fetus (who is not fully a person in Jewish law) or the infant prior to thirty days (who

is considered not yet fully viable). GLT also falls under the permissibility given to therapeutic interventions. What distinguishes Judaism from other perspectives, religious and secular, is its attitude toward the risk of unanticipated ill consequences posed by gene therapy (especially GLT). Freundel argues that Jewish law deals with what exists in the present: "A person who is ill today is to be helped to the extent possible. What results in later generations will be dealt with then." Freundel grounds this in "a Talmudic principle that enables us to assume that when we do our best G_d will take care of what we could not foresee or anticipate. If things do not work out, the theological question is G_d's to answer, not ours" (43).

Given this strong endorsement of gene therapy in principle, discussions have focused on whether or not Jewish law forbids certain means of accomplishing it. In an early article Rosenfeld considers possible objections to gene surgery on human ova (their removal, genetic modification, and replacement) and transplantation of genes from a donor into germinal cells (44). As for gene surgery, Rosenfeld appeals to an "indisputable principle" that any surgery permitted on a person is permitted on germinal cells, which are at most potential persons, prior to conception. If there were a surgical cure for hemophilia, it would be permissible; hence gene surgery to cure hemophilia is also permitted. Similarly, if (as many authorities agree) cosmetic surgery is permitted to relieve psychological distress, then gene surgery to achieve cosmetic effects should also be permitted. These permissions assume, however, that the procedures are safe enough that the ova are (almost) never destroyed; otherwise, they violate the prohibition against "destruction of the seed." Gene transplants raise the questions of whether they involve a possible illicit sex act and whether the child whose birth followed the procedure would be considered related to the donor, with implications for inheritance and restrictions on marriage into that family. Rosenfeld argues that no sex act is involved, since the donor genes need not come from the reproductive cells of the donor and the transplantation is carried out outside the body of the recipient. Nor would the child be related to the donor: If ovaries or testicles were transplanted, a child conceived after the transplantation would not be regarded as related to the donor; why, then, would he or she be in the case of genes? It is impossible that transplanting submicroscopic parts of sex cells — which, as invisible to the human eye, are generally excluded from consideration in Jewish law — would have more effect on the status of a child than whole sex organs. Since later commentators tend to repeat Rosenfeld's analysis, it is difficult to know what *halakhic* difficulties, if any, alternative methods of gene therapy would present.

Rosenfeld's reference to cosmetic surgery raises the question of enhancements more generally. Freundel and Rosenfeld both refer to a Talmudic story according to which Rabbi Yohanan sat on the road from the ritual bath so that women would see him before resuming sexual relations with their husbands. Yohanan's purpose was that women would think of him and thereby produce offspring as handsome as he and as accomplished a scholar. For both Freundel and Rosenfeld, the story authorizes the use of genetics for intellectual-ethical and aesthetic purposes. According to Freundel, genetic interventions for such purposes are no different from psychological or behavioristic interventions. Are there any limits to what characteristics can be modified? Freundel argues that certain characteristics — he mentions speech, intellect, love, and creativity — are "manifestations of the presence of the soul, but not the same as the soul." To discover the genetic sites of these characteristics is permissible. "Great concern, however, would exist about tampering with such sites either in terms of damaging something fundamentally human or in terms of potentially diminishing free will and individuality." Freundel's concern, however, stops short of a permanent prohibition; a judgment on such interventions would have to be made in accordance with the nature and impact of the interventions. On the basis of the stories about the creation of *golem* in Jewish lore, Jewish commentators tend to rule out the use of genetics to create humanlike beings lacking fundamental human qualities, though it is not clear what qualities would have to be absent or distorted (the *golem* often lacked speech).

Jewish attitudes to eugenics are shaped by the horrendous suffering of Jews at the hands of Nazi eugenic policies. Laurie Zoloth-Dorfman locates this history within a longer history of discrimination and violence in which the difference of the Jewish body from the male gentile body served to mark Jews as dangerous or less than fully human (45). Zoloth-Dorfman also notes the readiness of some Jewish physicians to use medicine in order to make the bodies of Jews conform to gentile "normality." Her point is that Jews and gentiles must remember this history as a warning not use genetic medicine to underwrite a suspect "normality" to which others are made or expected to conform. Freundel argues that Jews would not accept any eugenic program that kept those who wanted to procreate from doing so or that defined certain people as undesirable.

Rosner mentions two other (extra-*halakhic*) ethical issues. First, gene therapy is subject to common ethical principles governing novel therapies and research on human subjects. Second, gene therapy runs the risks of furthering the mechanization of human life. As with the extra-*halakhic* concerns regarding artificial beings and eugenics, however, it is not clear whether or how these concerns qualify the acceptance of gene therapy. Also missing from these accounts, but prevalent in analyses of other bioethical issues by Jewish commentators, Orthodox and other, is attention to questions of allocation of and access to gene therapy.

Islam

By almost all accounts, Islam places no limitations on the pursuit of scientific knowledge, including genetic knowledge. However, Hassan Hathout and B. Andrew Lustig note that applications of scientific knowledge are subject to five Islamic governing rules (46). The third of these rules refers to "changing God's creation," a phrase uttered by Satan in connection with his plans for leading humankind astray. Hathout and Lustig claim that there is a consensus among Islamic scholars that this

phrase does not support a ban on genetic engineering; otherwise, it would rule out many forms of life-saving and life-promoting surgery (appendectomy, tonsillectomy, cholecystectomy, and others) that involve a change in God's creation. They conclude that genetic engineering is permissible, but that the fourth rule — "Wherever the welfare exists, there stands the statute of God" — requires juridical sanctions to ensure that applications of genetic research will be used for human benefit. However, Gamal Serour, who has written widely on genetic and reproductive issues in Islam, restricts the justifiability of gene therapy to its therapeutic uses (47). Use of genetic technology for enhancement or eugenic purposes "would involve change in the creation of God" that could lead to imbalance in the universe as a whole or in humanity.

Other

The survey thus far presents an incomplete picture of religious attitudes to gene therapy. It excludes the many religious traditions, large and small (including the major traditions of South or East Asian origin), which have not yet addressed gene therapy in a substantive way. Among those it surveys, it concentrates on official or quasi-official statements and leading theologians, which may only approximately reflect the attitudes of large numbers of adherents. Finally, it ignores the views of the growing number of people who do not identify with an official religious body but who have spiritual commitments, often drawn from a wide range of sources. Popular culture often expresses possibilities or fears related to genetics in religious language (48), while opponents of genetic technologies, including Jeremy Rifkin (26) and Robert Sinsheimer (27), describe DNA as sacred or quasi-sacred. Is this latter view, rejected by most public spokespersons of mainstream religious traditions, widely shared, and does it indicate a principled stance against many genetic interventions or only an expression of ambivalence toward the latter? The importance public representatives of mainstream religious traditions place on criticizing such views indicates their precarious hold on the language of the sacred and, by extension, the religious response to gene therapy.

BIBLIOGRAPHY

1. J.R. Nelson, *Hum. Gene Ther.* **1**, 43–48 (1990).
2. National Council of the Churches of Christ in the U.S.A., *Genetic Science for Human Benefit*, National Council of Churches of Christ in the U.S.A., New York, 1986.
3. World Council of Churches, *Manipulating Life*, World Council of Churches, Geneva, 1982.
4. World Council of Churches, *Biotechnology: Its Challenges to the Churches and the World*, World Council of Churches, Geneva, 1989.
5. F. Harron, ed., *Genetic Engineering: Social and Ethical Consequences*, Pilgrim Press, New York, 1984.
6. National Council of the Churches of Christ in the U.S.A., *Human Life and the New Genetics*, National Council of Churches of Christ in the U.S.A., New York, 1980.
7. J. Breck, *St. Vladimir's Theol. Q.* **32**(1), 5–26 (1988).
8. D. Demopulos, in J.T. Chirban, ed., *Ethical Dilemmas: Crises in Faith and Modern Medicine*, University Press of America, Lanham, MD, 1994, pp. 91–99.
9. Pope John Paul II, *Origins* **12**(21), 342–343 (1982).
10. Pope John Paul II, *Origins* **13**(23), 385–89 (1983).
11. The Catholic Bishops' Joint Committee on Bioethical Issues, *Genetic Intervention on Human Subjects: The Report of a Working Party*, Linacre Centre, London, 1995.
12. G. Grisez, J. Boyle, and J. Finnis, *Am. J. Juris.* **32**, 99–151 (1986).
13. Catholic Health Association of the United States, *Human Genetics: Ethical Issues in Genetic Testing, Counseling, and Therapy*, Catholic Health Association of the United States, St. Louis, MO, 1990.
14. Europaishes Parlament, *Betr. die Gen-Manipulation*, Empfehlung 934, Bundestagsdrucksache 9/1373, 1982, pp. 11–13.
15. R.A. McCormick, *The Critical Calling: Reflections on Moral Dilemmas since Vatican II*, Georgetown University Press, Washington, DC, 1989, pp. 261–272.
16. R.A. McCormick, *The Critical Calling: Reflections on Moral Dilemmas since Vatican II*, Georgetown University Press, Washington, DC, 1989, p. 345.
17. J.F. Keenan, *Hum. Gene Ther.* **1**, 289–298 (1990).
18. Episcopal Church, *Resolution 1991-A095*, Archives of the Episcopal Church, New York, 1991.
19. Division for Social Responsibility, Church in Wales Board of Mission, *Human Genetic Screening and Therapy: Some Moral and Pastoral Issues*, Church in Wales Publications, Penarth, 1992.
20. *Lancet* (1988).
21. J.S. Feinberg, in J.F. Kilner, R.D. Pentz, and F.E. Young, eds., *Genetic Ethics: Do the Ends Justify the Genes?* Eerdmans, Grand Rapids, MI, 1997, pp. 183–192.
22. J.C. Peterson, in J.F. Kilner, R.D. Pentz, and F.E. Young, eds., *Genetic Ethics: Do the Ends Justify the Genes?* Eerdmans, Grand Rapids, MI, 1997, pp. 193–202.
23. Church of the Nazarene, *Manual/1993–97*, Nazarene Publishing House, Kansas City, MO, 1993, 904.1.
24. Seventh-Day Adventist Church, *Statements, Guidelines and Other Documents*, Seventh-Day Adventist Church, Silver Spring, MD, 1995, pp. 91–98.
25. T. Peters, *J. Med. Philos.* **20**, 365–386 (1995).
26. J. Rifkin, *Algeny*, Penguin, New York, 1983.
27. R. Sinsheimer, *Technol. Rev.* **86**, 14–17, 70 (1983).
28. T. Peters, in T. Peters, ed., *Genetics: Issues of Social Justice*, Pilgrim Press, Cleveland, OH, 1998, pp. 1–45.
29. T. Peters, *For the Love of Children: Genetic Technology and the Future of the Family*, Westminster John Knox Press, Louisville, KY.
30. G. Meilaender, *Christian Century* **107**, 872–875 (1990).
31. H. von Schubert, *Evangelische Ethik und Biotechnologie*, Campus Verlag, Frankfurt, 1991.
32. K. von Kooten Niekerk, in J. Hubner and H. von Schubert, eds., *Biotechnologie und Evangelische Ethik*, Campus Verlag, Frankfurt, 1992, pp. 324–347.
33. United Methodist Church, *New Developments in Genetic Science*, United Methodist Publishing House, Nashville, TN, 1992.
34. J.R. Nelson, *Human Life: A Biblical Perspective on Bioethics*, Fortress Press, Philadelphia, PA, 1984.

35. General Assembly, Presbyterian Church, U.S.A., *Social Policy Compilation*, Presbyterian Church, U.S.A., Louisville, KY, 1990.
36. Presbyterian Church, U.S.A., *The Covenant of Life and the Caring Community*, Presbyterian Church, U.S.A., Louisville, KY, 1983.
37. United Church of Christ, *The Church and Genetic Engineering*, United Church of Christ, Cleveland, OH, 1989.
38. W. Storer and I. Torrance, eds., *Human Genetics: A Christian Perspective*, The Church of Scotland Board of Social Responsibility, Edinburgh, 1995.
39. C. Clothier, *Report of the Committee on the Ethics of Gene Therapy*, H.M. Stationery Office, London, 1992.
40. R. Cole-Turner, *The New Genesis: Theology and the Genetic Revolution*, Westminster John Knox Press, Louisville, KY, 1993.
41. F. Rosner, *Modern Medicine and Jewish Ethics*, 2nd ed., Yeshiva University Press, New York, 1991.
42. R. Green, *Judaism* **34**(3), 263–277 (1985).
43. B. Freundel, in J.R. Nelson, ed., *On the New Frontiers of Genetics and Religion*, Eerdmans, Grand Rapids, MI, 1994, pp. 120–136.
44. A. Rosenfeld, in F. Rosner and J.D. Bleich, ed., *Jewish Bioethics*, Hebrew Publishing Co., Brooklyn, NY, 1979, pp. 402–408.
45. L. Zoloth-Dorfman, in T. Peters, ed., *Genetics: Issues of Social Justice*, Pilgrim Press, Cleveland, OH, 1998, pp. 180–202.
46. H. Hathout and B.A. Lustig, in B.A. Lustig et al., eds., *Bioethics Yearbook*, vol. 3, Kluwer Academics, Dordrecht, The Netherlands, 1993, pp. 133–148.
47. G.I. Serour, in B.A. Lustig et al., eds., *Bioethics Yearbook*, vol. 5, Kluwer Academics, Dordrecht, The Netherlands, 1997, pp. 171–188.
48. D. Nelkin and M.S. Lindee, *The DNA Mystique: The Gene as Cultural Icon*, Freeman, New York, 1995.

See other GENE THERAPY and RELIGIOUS VIEWS ON BIOTECHNOLOGY entries.

GENE THERAPY, ETHICS, SOMATIC CELL GENE THERAPY

NELSON A. WIVEL
University of Pennsylvania School of Medicine
Philadelphia, Pennsylvania

OUTLINE

INTRODUCTION

With the initiation of any new area of scientific research, a whole host of unknowns must be addressed. In the preclinical stages of development, moral restraints do not play a major role, but once human experimentation is contemplated, a significant number of ethical parameters have to be considered. This has been particularly true for human gene therapy, an application of recombinant DNA technology that involves the insertion of functioning genes into the somatic cells of a patient, either to correct an inborn genetic error or to impart a new function to the cell.

In that interim period between medical experimentation and actual development of medical treatment, researchers continually must confront the necessity of creating a positive risk–benefit ratio. This requires careful analysis with regard to the choice of disease target, the prognosis for a given disease, the cohort of research subjects to be studied, and the types of therapy currently available. This matrix of issues must be addressed both for gene therapy and all other types of human subjects research. In this particular sphere, gene therapy is no different than other types of experimental intervention.

One of the major ethical debates surrounding gene therapy has focused on the question of whether or not this particular application of biotechnology is qualitatively different from preceding types of medical therapies. A superficial analysis might suggest this to be the case, but a more careful consideration will lead to the conclusion that gene therapy is simply an extension of a therapeutic continuum (1). Since somatic cell gene therapy targets only nonreproductive cells, the genetic changes are limited to the patient, and there is little or no chance of affecting future offspring of that patient. Further the products of gene transfer are proteins that function in a manner analogous to drugs. In addition the use of gene transfer has the potential to produce many of the same results as allogeneic organ transplantation. For example, diabetes mellitus can be treated pharmacologically with insulin, or one can attempt to treat this disease with islet cell transplantation, or one could postulate treatment with gene therapy in which the gene encoding for insulin is given to the patient and the control of gene expression is physiologically regulated. In the foregoing context, gene therapy is not so intrinsically different from other treatments, and it takes advantage of the knowledge derived from more standard therapies.

A principal purpose of this discussion is to examine some of the early findings that led to the development of gene therapy, to look at origins of public oversight of human gene therapy, to review the major ethical questions concerning this type of research, and to look at the challenges for the future that are posed by this particular form of molecular medicine.

DEVELOPMENT OF SCIENCE AND CONCEPTS LEADING TO GENE THERAPY

If one looks at the history of molecular biology and molecular genetics, it encompasses a period of approximately

50 years, and the developments in this area of biological science have provided the necessary infrastructure for the establishment of human gene therapy. A critical discovery was made by Avery, MacLeod, and McCarty at Rockefeller University when they demonstrated that a gene-inducing transformation in bacteria could be transferred in nucleic acids (2). This satisfied a longstanding desire to identify chemical means by which hereditary traits could be transferred from one generation to the next. A few years later Watson and Crick were able to complete the model for the helical structure of DNA (3); this was followed by discovery of mRNA, and the subsequent development of the central dogma of molecular biology that postulates the flow of genetic information from DNA to RNA to protein (4). Another major advance occurred with the discovery of the restriction endonucleases, enzymes that could cut DNA at specific recognition sites, thus paving the way for creating DNA molecules in which sequences that are not naturally contiguous can be placed next to each other (5). Cohen and Boyer and their colleagues were able to construct functionally active recombinant DNA molecules (6), and it soon became possible to move genes from one species to another without loss of function. Several years later the development of gene delivery vehicles became a reality with the report that retroviruses could be modified to insert genes into cells (7). Although the first human gene therapy protocol was not approved until 1990, the techniques that made it possible had their origins in the many fundamental discoveries emanating from the continued interest in genetics.

As is so often true in science, its practitioners are able to frame concepts well in advance of the actual experiments. For a particularly intriguing history of gene therapy one is referred to the article of Wolff and Lederberg (8). They have pointed out that Edward Tatum, in 1966, was the first to suggest that viruses could be used to insert genes into cells and that modification and regulation of gene activities ultimately might be used to treat cancer (9). Lederberg proposed a potential gene therapy for hemophilia in 1968 when he stated that fractionated DNA containing the normal alleles of the hemophilia gene could be introduced into the liver in experiments analogous to the attempts at transforming bacteria (10). Arthur Kornberg, who successfully used DNA polymerase to synthesize DNA in vitro, predicted that hereditary defects might be cured by attaching a therapeutic gene to a harmless virus that would serve to infect the cell and deliver the gene (11). It is now popular to refer to DNA as a designer drug, but in 1970 Aposhian advanced the idea that if the purpose of a drug was to restore the normal function to a physiological process, then the time might arrive when DNA would become the ultimate drug (12).

If the basic understanding of molecular genetics suggested ideas for gene therapy itself, there were also discussions of the ethical implications of gene therapy and one of the first was presented by Marshall Nirenberg who played a key role in deciphering the genetic code. In 1967 he predicted that cells would be programmed with synthetic messages within 25 years, but that the technology might surpass our ability to assess the long-term consequences of such alterations (13). He expressed the particular concern that this knowledge might not benefit humankind unless it was applied with sufficient wisdom. Robert Sinsheimer accelerated his misgivings by suggesting that designed genetic change was simply a new form of eugenics that could escape the boundaries of the selective processes that occur in nature (14). It was his contention that these new types of choices imposed an extreme need for responsibility.

Although the first human gene therapy trial was not approved by local and national oversight bodies until 1990, there was an earlier attempt at gene transfer that occurred in the 1960s. Stanfield Rogers had done fairly extensive work with the Shope papilloma virus and observed that it apparently had arginase activity. Animals infected with this virus exhibited reduced blood levels of arginine, and laboratory personnel who studied this virus also had reduced blood arginine but no apparent side effects. Three siblings who suffered from arginemia because of arginase deficiency were injected with the Shope virus, but there was no reduction in blood arginine; fortunately there was no evidence of toxicity (15,16). There was some concern expressed about the ethics of these experiments, but the principal investigator defended his actions on the grounds that this type of intervention offered the only reasonable chance to prevent progressive deterioration in these children (17). Another pioneer in the field of gene therapy, French Anderson, supported the experiment on the basis that there were several decades of experience documenting the safety of the Shope virus, and that there was an absolute certainty of suffering and death associated with arginemia (18). Still other investigators expressed the concern that the experiment lacked sufficient preclinical data and that it would serve as stimulus for other groups to proceed in an unprepared fashion (19). There were no further human experiments until 1980 (these will be described at a later point in this discussion).

In attempting to characterize the confluence of events that supported the actual development of human gene therapy, one is properly forced to consider the analogies to pharmacology and surgery. While the administration of drugs preceded our current understanding of molecular genetics, it is our present-day knowledge of genetic principles that allows us to consider the use of DNA fragments in a pharmacotherapeutic context. Similarly gene therapy might be considered as molecular surgery since the incorporation of a therapeutic gene into cell chromosomes has the potential to modify tissues or organs for the life span of the treated patient (8). Thus human gene transfer for the treatment of disease does not require entirely new ethical paradigms, but rather it requires careful attention to those issues that bear on any type of human experimentation, the choice of disease, the choice of patients, the risk–benefit ratio, the need for informed consent, and the right of patient privacy.

ORIGINS AND EVOLUTION OF PUBLIC OVERSIGHT OF HUMAN GENE THERAPY

In order to develop the proper perspective for the review processes that were created explicitly for human gene therapy clinical trials, one has to look at the history of

recombinant DNA research. Some of the initial concerns that were applied to recombinant DNA technology abated over the course of five to six years when it became readily apparent that many of the postulated hazards were not going to occur, but there was a recrudescence of concern as the possibilities for human gene therapy became concrete.

Marshall Nirenberg's caution about genetic engineering was translated to the public sector and by 1968, Senator Mondale introduced a resolution for the specific purpose of establishing a Commission on Health, Science, and Society. One of the proposed tasks for this commission was the study of the moral and ethical questions surrounding genetic intervention. Several prominent geneticists testified at hearings held during the spring of 1968 and emphasized that there were few societal hazards associated with the "new" genetics. In the period between 1968 and 1982, Congress did not concern itself with human gene therapy, but in the mid-1970s there was a spirited debate about the potential hazards of recombinant DNA research.

As early as 1971 Paul Berg was having success in developing the first recombinant viral vector, using the simian virus SV40. His experiments provoked a spirited discussion at a meeting at the Cold Spring Harbor Laboratory. Because of the legitimate differences of opinion, the first Asilomar Conference was convened in 1973. There was a fairly systematic discussion of some of the potential hazards associated with certain kinds of experiments such as the insertion of antibiotic resistance genes into bacteria that do not normally possess such properties, or moving part or all of the genes of one animal virus into a plasmid or another virus. Because of the cumulative uncertainties, a voluntary moratorium was suggested and was communicated in the form of a letter to the journal, *Science*, with many of the most active investigators as signatories (20). In February 1975 the second Asilomar Conference was held, and the results were suggestive of a new order for the scientific community. Despite internal disagreements the participants in this meeting ultimately reached a consensus that there should be a scheme for control of recombinant DNA experiments that would at least minimize, or preferably eliminate, potential biohazards. Two kinds of containment were proposed: One was physical containment that would require appropriate facilities, and the second was biological containment that would require the engineering of microorganisms so that they would have a selective disadvantage for survival in the laboratory environment.

Immediately following this conference, the first meeting of the National Institutes of Health (NIH) Recombinant DNA Advisory Committee (RAC) was convened. The primary task for this group was to create the "NIH Guidelines for Research Involving Recombinant DNA Molecules"; the task required about 16 months, and this document was published on June 23, 1976, in the *Federal Register* (21). As a part of its charge, RAC reviewed all recombinant DNA research that was conducted in institutions receiving NIH support.

In the period from 1976 to 1980, it became apparent that many of the predicted hazards of recombinant DNA research did not materialize, and the NIH Guidelines were revised into a less stringent format. However, 1980 became a sentinel year when attention toward human gene therapy reached a new level. Shortly after the Supreme Court ruled that recombinant microorganisms could be patented, a letter was sent to President Carter. It was signed by the General Secretary of the National Council of Churches (Protestant), the General Secretary of the Synagogue of America, and the General Secretary of the United States Catholic Conference. The thesis of this document was that questions about the proper use of genetic engineering were moral, ethical, and religious questions. Misuse of this technology was seen as a threat to the fundamental nature of human life and the dignity and worth of the individual human being. There was a specific request for the formation of a body of wide-ranging interests and expertise that could advise the government in its necessary oversight role. A Presidential Commission was formed and its initial meeting took place in July 1980; it accepted the task of studying the ramifications of genetic engineering.

In the meantime, Dr. Martin Cline of the University of California, Los Angeles (UCLA) School of Medicine discovered that DNA from a methotrexate-resistant Swiss 3T6 cell line could be successfully transfected into mouse bone marrow cells (22). By using functional markers, he was able to establish that the transfected bone marrow cells could be successfully transplanted into irradiated mice, and that resistance to methotrexate was maintained. Based on this experimental evidence, he attempted to transfect the β-globin gene into human bone marrow cells that were then transplanted into two patients with thalassemia (one in Israel and one in Italy). Because this protocol had not received approval by the local safety committees at UCLA, an investigation was conducted by NIH with the result that Dr. Cline was censured and lost research funding (23,24).

By 1982 the Presidential Commission had published its report entitled, *Splicing Life*, and it concluded that there were no fundamentally new social and ethical questions raised by somatic cell gene therapy (25). In response to this report, Congressman Albert Gore convened hearings on human genetic engineering. Scientists, clinicians, ethicists, and lawyers all testified, and the recurring question from Mr. Gore alluded to the need for some kind of government body to oversee the development of human gene therapy.

In the final chapter of the *Splicing Life* report, it was suggested that a reconstituted RAC might be an appropriate oversight body for human gene therapy studies. Given the fact that RAC had been functioning as a recombinant DNA review body since 1976, it clearly possessed the most in-depth experience and expertise in this particular technology. Although the original RAC was composed only of scientists, the composition of its membership was changed in 1978 by Joseph Califano (then Secretary of Health, Education, and Welfare) so that two-thirds were scientists and one-third were "public" members.

In April 1983 the chairman of RAC asked the members of this committee if they wished to respond to the *Splicing*

Life report. The response was positive and addressed two issues, the establishment of a Working Group to develop guidelines and review procedures, and acceptance of the responsibility for review of actual protocols at such time that it would become necessary. This Working Group had two meetings in 1983 and then recommended that a larger interdisciplinary group be convened; both RAC and the NIH director concurred with this suggestion and a 15-member Working Group on Human Gene Therapy was created. Members represented clinical medicine, laboratory science, ethics, and law. Throughout 1984 a document entitled "Points to Consider in the Design of Human Gene Therapy Protocols" was prepared, and a first edition was published in the *Federal Register* in 1985. At its September 1985 meeting the full membership of the RAC accepted the revised version of the Points to Consider. By February 1986, the executive secretary of RAC sent a letter to all potential investigators, asking for the submission of preclinical data pertaining to the development of human gene therapy protocols. During 1986 and 1987 the Working Group was known as the Human Gene Therapy Subcommittee (HGTS), and a number of essential discussions were held that addressed such topics as retroviral vectors, the use of trangenic animals as disease models, and the FDA process for regulation of investigational new drugs (INDs). By 1987, a group of investigators, headed by French Anderson, submitted to HGTS a compendium entitled, "Human Gene Therapy: Preclinical Data Document." This was reviewed in the context of being an actual protocol but was actually a prelude to the first proposal for human gene transfer.

In July 1988, the first request for a protocol was submitted to HGTS by Steven Rosenberg, French Anderson, and others; this was not a true "gene therapy" protocol, but rather a "gene-marking" trial to determine if retroviral vectors containing the transgene encoding for neomycin resistance could be given to human subjects without untoward effects. Following an initial discussion, approval of the protocol was deferred, based on the need for additional data. In September 1988, HGTS again requested more information, but in October, the parent body, RAC, approved this protocol with considerably less than a unanimous vote. On October 18, 1988, the director of NIH, Dr. James Wyngaarden, did not approve the protocol and sent it back to the HGTS with the request for additional data. By December 1988, HGTS had approved the protocol and members of the RAC gave their approval via a telephone conference call.

In January 1989, the NIH director publicly announced the approval of this "gene-marking" protocol. Almost immediately a lawsuit was filed by the Foundation on Economic Trends, litigation that was designed to prevent patients from enrolling in the trial. It was argued that a telephone conference call among RAC members was not equivalent to a public meeting and thus a violation of the NIH Guidelines. Following several months of legal interactions, the matter was brought to a successful conclusion. It was during December 1988 that the two-stage process for national review was established. By joint committee agreement it was decided that initial review of protocols would be conducted by HGTS, and once full approval was granted, said protocols would be forwarded to RAC. At this stage of the oversight process, most of the expertise for gene therapy review was concentrated in HGTS, although there were some members who served both on HGTS and RAC.

In March 1990, French Anderson and Michael Blaese submitted a protocol for the study of adenosine deaminase (ADA) deficiency, a form of severe combined immune deficiency that often caused the death of patients within the first two years of life. At its June meeting, HGTS agreed to provisional approval of this protocol, but requested additional data relating to proof of a selective survival advantage for lymphocytes transduced with the normal ADA gene. On July 30 and 31, 1990, several groundbreaking events occurred. On July 30, HGTS approved two protocols, the ADA protocol and a cancer protocol designed to use tumor-infiltrating lymphocytes as the delivery vehicle for tumor necrosis factor, a means of treating melanoma by adoptive immunotherapy. On July 31, the RAC convened its meeting and approved both protocols.

Although the results of the "gene-marking" protocol were a necessary prologue, the first actual gene therapy trial was conducted on September 14, 1990, when Ashanti De Silva received approximately one billion of her own peripheral blood T lymphocytes that had been transduced with the normal ADA gene. Although the chronology of these early events has been compressed for the purposes of this discussion, a detailed history can be abstracted from the published minutes of the RAC meetings (26–30).

At the outset, there were two independent and parallel processes for the review of human gene therapy protocols, the entirely public review process conducted by RAC and the FDA review process, a mandatory regulatory exercise dictated by federal statute. There was no option for open review by FDA because of legal requirements that all information pertaining to the development of drugs or cell therapies or gene therapy be treated as proprietary and therefore confidential. In retrospect, it was probably fortuitous that NIH RAC accepted the responsibility for public review, since the absence of such an activity could have negatively influenced public acceptance of this experimental form of molecular medicine.

Two public opinion surveys, taken in 1986 and 1992, yielded results that are interesting, but also reflective of the fact that the general public had very little knowledge of genetic engineering and gene therapy. In 1986, before the first protocol was approved, 52 percent of a random sample of polled individuals felt that it was not morally wrong to genetically alter human cells to treat disease, while 47 percent strongly approved of gene therapy to treat genetic diseases and 41 percent somewhat approved of this approach (31). In 1992, after a number of gene therapy protocols had been approved, and several initiated, 30 percent of respondents were very willing to undergo gene therapy to correct a serious or fatal genetic disease before symptoms appeared in late life, and 49 percent were somewhat willing. When asked about the willingness to have a child undergo gene therapy for a usually fatal genetic disease, 52 percent were very willing and 36 percent were somewhat willing (32). Despite a lack

of in-depth understanding, these new technologies were seen as acceptable in the treatment of genetic disease, and this is an important finding since the majority of research monies supporting gene therapy, and all the basic research that augments gene therapy, represent tax dollars, and not private investment.

As gene therapy clinical research has developed, there has been a parallel with recombinant DNA research itself in that many of the fears about safety have failed to materialize. Experience has been an important factor in the rather constant modification of the oversight process. By 1991, it became apparent that the two-stage national review of gene therapy was becoming redundant, unnecessarily time-consuming, and of questionable value. In October 1991, RAC decided to consider the disbanding of the HGTS and transferring its membership to the parent body. It was correctly assumed that the major responsibility of RAC, at this point in its history, was human gene therapy and not bacterial genetics. By February 1992, HGTS was formally disbanded, and the sole responsibility for public national review was vested in RAC, whose meetings were increased from three times per year to four. A one-year transition period was established for the purpose of transferring the members of HGTS, not already on the RAC, to the RAC (33).

While the review process was time-consuming, investigators took their responsibilities seriously, and while the open dialogues often served to highlight legitimate differences of opinion, the end result was an improvement over the initial submission. However, in 1992, Drs. Ivor Royston and Robert Sobol requested that the director of NIH and the commissioner of the FDA grant them a compassionate plea exemption so that they could use gene therapy to treat a patient with a brain tumor. This request was particularly challenging because it would bypass all the usual oversight procedures that had been put into place. Although FDA had set a precedent for compassionate plea exemptions for single patients, no analogous mechanism was in place at NIH.

At the December 1992 meeting of RAC, members raised serious concerns about this request. Since the entire field of research was still so new, there was no unequivocal evidence of efficacy in any of trials. These particular investigators had a paucity of preclinical data, and therefore it was difficult to assess the risks associated with the proposed treatment. If the request had represented a minor variation on previously approved protocols, it might have been possible to approve it, but that was not the case. After vigorous discussion, the RAC declined to approve the protocol.

Subsequently the NIH director and the FDA commissioner approved the request on a compassionate plea basis, an action that was entirely within their prerogatives. Committees such as the RAC are simply advisory to the NIH director and their actions constitute recommendations that can be accepted or rejected. It is of note that this particular circumstance was never repeated, but RAC responded to the episode by creating procedures for expedited review, with provisions for using ad hoc reviewers who could be available on short notice (34). A further revision occurred in that protocol categories were established and some of these categories were exempted from full RAC review.

A further change in oversight was initiated in 1994 as a result of the formation of the National AIDS Task Force on Drug Development. This task force had an unusual charter in that the government officials on the task force were required to be at the level of agency head. Thus the Assistant Secretary of Health, the NIH director, and the commissioner of FDA were members of this working group. Other members included physicians, pharmacologists, AIDS clinical researchers, senior pharmaceutical company executives, and members of several AIDS activist groups. This latter contingent had a single element agenda and that was a change in the approval procedures for human gene therapy protocols. It was the contention of the AIDS activists that gene therapy offered one of the best hopes for the cure of this disease, and that the dual agency review of gene therapy protocols was unnecessarily inhibiting progress. They proposed that RAC be abolished and that sole review of human gene therapy protocols be confined to FDA. This clearly was not acceptable to the entire task force and a compromise was reached.

The new plan was crafted on the idea of consolidated review in which both NIH and FDA would review all new protocols simultaneously. Designated staff members from the two agencies would consult; if the protocol represented a marked departure in concept from previous protocols, it would be fully reviewed by both NIH RAC and FDA. If the protocol lacked any notable differences from previous protocols, it would receive only one review by FDA. First, the AIDS Task Force on Drug Development accepted this new scheme for protocol review, and RAC, at its September 1994 meeting, voted to make the appropriate changes in the NIH Guidelines (35). At this point, consolidated review became the standard operative procedure.

Another significant event occurred in 1995 when NIH Director Dr. Harold Varmus appointed an ad hoc review committee to assess the function of RAC, to develop recommendations about its future role, and to identify ways in which this committee could best support gene therapy research. This committee met several times and issued a summary of its findings in September 1995 (36). It affirmed the basic tenets of the consolidated review process, it recommended that RAC should continue to provide advice on gene therapy policy matters, and it recommended that a data management system should be devised to enable RAC and the Office of Recombinant DNA Activities to monitor all gene therapy clinical trials, even in the absence of case-by-case review.

In 1996, the NIH director announced that he planned to abolish RAC and to replace it with a small number of scientists and ethicists who would meet on an ad hoc basis to render advice on public policy issues relevant to human gene therapy research. After publishing a Notice of Intent in the *Federal Register*, a large number of comments was received and the vast majority were in favor of retaining RAC as the principal public advisory committee for human gene therapy research. In September 1996, the NIH director announced that RAC would be retained but would operate under the following conditions. Membership would be reduced from 25 to 15, the approval process for

gene therapy protocols would be relinquished, and a major new assignment would be the responsibility for organizing gene therapy policy conferences to cover such topics as lentivirus vectors, in utero gene therapy, and the use of gene transfer for the purposes of enhancement.

In 1996, the RAC began operating under its new mandate. Several gene therapy policy conferences have been held and selected protocols have been reviewed even though approval is no longer required. Thus RAC continues to provide a venue for public discussion of new scientific and ethical issues in the area of gene therapy clinical trials and can compliment the privately conducted review activities of FDA.

PRINCIPLES OF REVIEW FOR HUMAN GENE THERAPY

From the standpoint of process, there were two principal elements in the NIH oversight of human gene therapy protocols. At the local level, two separate committees were involved, the Institutional Review Board (IRB) or the ethics board responsible for protecting patients from unnecessary research risks, and the Institutional Biosafety Committee (IBC), a product of the requirements detailed in the NIH Guidelines (37). In looking at the division of labor, IRB placed a primary emphasis on the informed consent document, and IBC made its principal focus the science of the pertinent recombinant DNA technology. At the national level, both HGTS and RAC were initially involved, and later RAC itself assumed sole review responsibility. With the exception of the IRB, all of the review bodies relied on the Points to Consider or what is now Appendix M of the NIH Guidelines.

In a previous publication, Walters and Palmer have reduced more than 100 specific questions in Appendix M to seven central ethical questions. They are as follows:

1. What is the disease to be treated, and why is it a good candidate for gene therapy?
2. What alternative treatments are available for this disease?
3. What is the potential harm associated with the genetic intervention?
4. What is the potential benefit associated with the intervention?
5. What steps will be taken to ensure that participants in the study are selected in a manner that is fair to everyone who wants to take part in the study?
6. What steps will be taken to ensure that the consent of study participants is both informed and voluntary?
7. What steps will be taken to protect the privacy of participants and the confidentiality of medical information about them (38)?

A thorough perusal of Appendix M of the NIH Guidelines emphasizes that the great multitude of specific questions are directed toward the seven aforementioned general questions. In developing a background and rationale for a trial, the principal investigator must provide information concerning the disease to be studied (General Question 1). The following subset of questions must be addressed. Why is the disease selected for treatment by means of gene therapy a good candidate for such treatment? What is the natural history and range of expression of the disease selected for study? What objective and/or quantitative measures of disease activity are available? Are the usual effects of the disease predictable enough to allow for meaningful assessment of the results of gene therapy? Is the protocol designed to prevent all manifestations of the disease, to halt the progression of the disease after symptoms have begun to appear, or to reverse manifestations of the disease in seriously ill patients?

Under General Question 2 (alternative therapies), the principal investigator is required to describe what alternative therapies exist and characterize the groups of patients for whom these therapies are effective. It is also necessary to indicate the relative advantages and disadvantages as compared with the proposed gene therapy. This particular matter assumed great importance in the very first gene therapy trial that was approved. Because of the very prolonged review process for the ADA trial, two alternative therapies came into existence, partially matched bone marrow transplants and the use of conjugated enzyme therapy (polyethylene-glycol-ADA, or PEG-ADA). Both showed modest degrees of efficacy. In the final analysis, the first two patients enrolled in the trial were maintained on PEG-ADA, and this action had sound ethical origins. It would have been impossible to justify the withdrawal of PEG-ADA and substitute an entirely experimental gene therapy of unknown therapeutic potential.

Under General Questions 3 and 4 (potential harms and potential benefits), it is necessary to provide a wealth of information about research design, including appropriate preclinical studies. In an important sense, high-quality science carefully applied to the study of patients is a necessary precursor to establishing a positive risk–benefit ratio. Thus it is essential to provide a complete description of the methods and reagents to be employed for gene therapy. It is necessary to describe the structure of the cloned DNA and to completely characterize the gene vector or gene delivery vehicle. This means that one has to perform a complete nucleotide sequence analysis or a detailed restriction enzyme map of the total vector–gene construct. All the regulatory elements in the construct must be identified, including the promoters, enhancers, polyadenylation sites, and origins of replication. In essence, one has to describe the molecular structure of the material that will be administered to the patient. It is also incumbent on the principal investigator to demonstrate the safety, efficacy, and feasibility of the proposed procedures using animal and/or cell culture model systems and to justify why the chosen models are the most appropriate. Very early in this history of human gene therapy research, it became apparent that there would be a number of diseases for which there was no adequate animal model system; it was necessary to introduce some flexibility into the review system and data from cell culture models were allowed to stand in lieu of animal models, when necessary.

In order to accurately assess the risks and benefits the vector or delivery system is subjected to the following

kinds of questions. What cells are the intended target cells of recombinant DNA? Is the delivery system efficient? What percentage of the target cells contain the added DNA? Is the added DNA extrachromosomal or integrated? How many DNA copies are present per cell? What is the minimal level of gene transfer and/or expression that is estimated to be necessary for the protocol to be successful in humans? Is the gene expressed in cells other than target cells?

As a further adjunct to assessing risks and benefits, the clinical elements of the study have to be described in detail, including procedures for patient monitoring. In answering the following questions, many of the key elements related to the success or failure of the study will be characterized. Will cells be removed from patients and treated ex vivo? Will patients be treated to eliminate or reduce the number of cells containing malfunctioning genes (radiation or chemotherapy)? How will the treated cells be administered? How will it be determined that new gene sequences have been inserted into the patient's cells and if these sequences are being expressed? What studies will be conducted to assess the presence and effects of contaminants? What are the clinical endpoints of the study? How will patients be monitored to assess specific effects of the treatment on the disease? What is the sensitivity of the analyses? What are the major beneficial and adverse effects of treatment anticipated? What measures will be taken in an attempt to control or reverse adverse effects if they occur?

When one looks at the levels of risk, there are two types of risk to take into consideration, the direct risk to the patient and the risks to health care workers or family, namely the public health considerations. To address these issues, the following kinds of questions must be answered. Is there a significant possibility that the added DNA will spread from the patient to other persons or to the environment? What precautions will be taken against such spread? In light of possible risks to offspring, including vertical transmission, will birth control measures be recommended to patients?

In summary, there are many component concerns that must be addressed before one can establish a risk–benefit ratio. In effect, this ratio will be altered by the type of disease being studied and the principal aims of the study itself. Yet for all the seeming complexities associated with this particular type of clinical research, the approval of individual protocols rests on a positive risk–benefit ratio, and with almost no exceptions, protocols that have failed approval have lacked critical information concerning this issue.

The fair and equitable selection of patients (General Question 5) poses a problem for most emerging new technologies as they reach the stage of clinical trials. Patient interest often exceeds the resources available for a Phase I trial, which is really a pilot study. Very often the diseases are extremely serious, and not infrequently, they are uniformly fatal. Given what are often dire circumstances, it is not surprising that patients and their families are seeking any possible solution.

A significant problem occurred rather early in the development of human gene therapy protocols. Investigators at NIH were given approval to conduct a Phase I trial for the study of glioblastoma multiforme, a type of brain tumor that has a particularly relentless course. Most of the patients die within less than a year after diagnosis. As initially approved, a maximum of 20 patients could be studied. Because of the intense interest and the publicity given to the study, approximately 2000 patients made inquiries about participation. Because of this overwhelming response, a screening committee was organized, and by choice, the principal investigators did not participate in this screening exercise in order to avoid any potential conflict of interest.

General Question 6 addresses one of the most complex issues pertaining to clinical research, that of informed consent. Informed consent from adults presents one particular set of challenges, but obtaining informed consent from minors means that parents or guardians will be the actual signers of the document. Interestingly enough, there are two differing ethical perspectives with regard to entering children into clinical trials. Almost 30 years ago, a dominant position was that clinical trials should be done in adults first, and then children could be studied secondarily (39). More recently the prevailing opinion is that persons who participate in clinical trials actually have the primary access to potential benefits that are not available to the public at large. This being the case, no class of individuals whether they be members of ethnic minorities, or women, or children, should be denied access to timely participation in clinical trials (40).

A principal challenge for investigators in gene therapy involves conveying the critical information about the disease, the major alternative treatments and the procedures to be followed in the clinical trial. If one adheres to the dictum that the information has to be understandable to someone with an eighth-grade education, this requires that such elements as recombinant DNA technology, the nature of vectors, the insertion of genes into cells, and the possible complications of the trial, have to be simplified in very imaginative ways.

Throughout its period of gene therapy protocol review, RAC spent considerable time revising the portion of Appendix M of the NIH Guidelines pertaining to informed consent. In essence, it became more detailed with time. Investigators were asked to provide information about the manner in which the study would be communicated to potential research subjects, the personnel involved in the process, the measures taken to avoid conflict of interest, particularly if the researcher has responsibility for the patient's medical care, the length of time for decision making, any special arrangements in place for pediatric or mentally handicapped subjects, the nature of reproductive risks or the need for reproductive restrictions, the need for long-term follow-up, and the indication that permission to perform an autopsy will be sought from the family, whatever the cause of the patient's death (41).

RAC's review of gene therapy protocols represents one of the few forums for public discussion of informed consent, and not infrequently, the discussions were both lively and marked by some disagreements between investigators and reviewers. If there was a recurring theme, it was focused on unbridled optimism on the part of some investigators

that resulted in optimistic statements about potential therapeutic benefit in Phase I trials, an inappropriate inclusion for a study that is designed to answer questions about safety and not efficacy. Many of the informed consent documents lacked clarity when it came to a discussion about the charges associated with certain research procedures. Obviously, the third-party payers or insurance companies have little interest in paying for anything but standard medical procedures that are relevant to the treatment of a given patient.

Another complicating element in the RAC review of informed consent was derived from the fact that the primary responsibility for government oversight of human subjects research was posited in the Office of Protection from Research Risks (OPRR) of the Department of Health and Human Services (DHHS). Its mandate was derived from a Code of Federal Regulations [45CFR46] (42). Thus RAC was reduced to functioning in a purely advisory role because the final control of informed consent documents resides with IRB, or the local ethics committee of the institution where the research is being conducted. If the investigators chose to ignore RAC's requests or if the local IRB chose to ignore them, there was no means for appeal. However, there has been a positive outcome to all these discussions in that there is documentable evidence that RAC had an influential role in shaping both the process and the text of the informed consent.

Major Ethical Question 7 addresses the issue of patient privacy and much attention was directed to procedures for maintaining such privacy when the original Points to Consider document was being drafted. Since gene therapy primarily was designed for the treatment of the so-called single gene deficiency diseases, and since many of these diseases affect children, the need for patient and family privacy was given appropriate recognition. It is generally accepted that it is the responsibility of the principal investigator and the members of the research and medical teams to preserve privacy if that is the request of the patient and his/her family. What has happened since the advent of the first approved protocol is that there have been notable exceptions to the general request for privacy. While the first two girls to be studied in the initial ADA protocol remained anonymous for several years, they were later featured, including names and pictures, in a popular magazine (43). During the second phase of the ADA trial, in which a number of technical changes were made, the parents of the newborns who were treated, specifically requested that the pertinent information be disclosed (44). Within the last year there has been a rather intense public interest in gene therapy trials in Boston that are designed to treat severe peripheral artery disease in adults. One of the first patients to be admitted to that trial consented to a public interview, and later on, the Public Broadcasting System television network ran a special program describing this trial and conducting patient interviews (45). While it is impossible to predict the wishes of a particular patient or family, the principal investigator and his/her team have a primary obligation to protect the family's intentions if privacy is requested; conversely, there is no choice but to step aside if the decision is to make the trial participation public.

HUMAN GENE THERAPY AND THE INDUSTRY MODEL FOR DRUG DEVELOPMENT

Strangely enough, the failure of human gene therapy to demonstrate unequivocal evidence of efficacy in any of the clinical trials, has delayed the onset of another significant problem. At such time that gene therapy is successful in the treatment of one of the rare monogenic deficiency diseases, the question of industry funding will be a point of active debate. It is particularly instructive to look at the pattern of clinical trials over the past eight years. While the rare genetic diseases provide the best conceptual models for the use of gene therapy, they have not been a major focus of investigation thus far. Of the 244 trials approved up to this point, 73 percent have been devoted to the study of cancer, 14 percent have involved monogenic deficiency diseases, and 9 percent are for the study of AIDS (46). Since many of these early trials have received significant funding from industry, it suggests that large or commercially promising markets are a key element in product development.

The rare disease issue is not new to the pharmaceutical industry, and indeed, Congress passed the Orphan Drug Act so that companies could be granted financial incentives to develop treatments for uncommon disorders. This legislation has been useful, but it is not known if it could be adapted to cover gene therapy. It is possible that new statues would be required in order to stimulate companies to develop treatments when there is a limited potential for payback.

There are a number of theoretical circumstances in which gene therapy could ameliorate a disease on the basis of a single treatment, and it is probable that the theoretical will shift to the actual within the next 15 to 20 years. In order for this to occur, it will be necessary to successfully insert a therapeutic gene into a cell such as a bone marrow stem cell and for that gene to be permanently expressed; further the level of expression will have to be sufficient to correct the genetic disorder. When this series of events occurs, there will now be in place a therapeutic model that is the exact opposite of the drug model in which a patient takes a pill one or more times a day. How would one price a gene therapy treatment in this context? Already there are several genetic diseases that can be treated with recombinant protein products; recombinant glucocerebrosidase is available for patients with Gaucher's disease, and in some instances it may cost as much as $300,000 a year to treat a single patient. Assuming that gene therapy could be used to treat this disease and do it effectively with one intervention, how much would one charge? Even if one charged $300,000 for gene therapy that might be only 5 or 10 percent of the income derived from using the recombinant protein. There is little in this scenario to tweak the interest of the marketing department of a biotechnology company.

In recognition of this problem, FDA and NIH have begun considering alternative ways of product development for gene therapy of rare genetic diseases. This might involve altering the standard paradigm for drug development in that clinical trials could be compressed into fewer phases and would require fewer patients (47). When the patient base is very small, the amount of gene product

needed will be relatively small and it may be feasible to have nonprofit distribution centers, located in academic medical institutions, serve the role of traditional drug and biotechnology companies. Naturally all the standards of quality control and quality assessment would have to be met in precisely the same way that is required of the for-profit sector. It will be interesting to watch what models actually develop as gene therapy for rare diseases becomes a reality.

CHALLENGES FOR THE FUTURE

This discussion has focused on ethical issues when gene transfer is used in the context of treating disease, but another significant challenge may develop in the not-too-distant future when this technology may be adapted for purposes other than the treatment of disease. This alternative use has been designated as enhancement. Between 1980 and 1993, there have been 28 international policy statements concerning gene therapy, but there are no specific references to genetic enhancement in these documents (38). There are two exceptions to the aforementioned statements; these emanate from advisory committees in Canada and the United Kingdom and find genetic enhancement to be ethically unacceptable (48,49). There have been two polls in the United States that have attempted to define public attitudes toward genetic enhancement, one taken in 1986 and one taken in 1992. In 1986, respondents were asked about their attitude toward genetic manipulation to improve physical characteristics in children, and 16 percent strongly approved, 27 percent somewhat approved, 21 percent somewhat disapproved, and 33 percent strongly disapproved; in 1992, the responses to the same question revealed that 16 percent strongly approved, 28 percent somewhat approved, 23 percent somewhat disapproved, and 31 percent strongly disapproved (31,32). When asked about their attitude toward genetic manipulation to improve intelligence in children, respondents in 1986 indicated that 17 percent strongly approved, 25 percent somewhat approved, 20 percent somewhat disapproved, and 35 percent strongly disapproved. In the 1992 survey, 18 percent strongly approved, 26 percent somewhat approved, 22 percent somewhat disapproved, and 31 percent strongly disapproved (31,32). Given the limits of error in this kind of poll, it appears that a slight majority of Americans are opposed to genetic engineering for enhancement, but a significant minority are in favor of it.

In 1997, NIH sponsored a Gene Therapy Policy Conference that was organized by the Office of Recombinant DNA Activities. Two issues were defined by the discussions that took place. First, there was the prediction that procedures for enhancement might be available in the near future, and second, it was agreed that the gradation from treatment of disease to enhancement might be subtler than once perceived. With regard to the second issue, it was pointed out that a biotechnology company had been successful in inserting the tyrosinase gene into the cells of the hair follicle (50). Current research is directed toward genes that promote hair growth. One of the primary objectives is to develop a product to treat alopecia or hair loss associated with cancer chemotherapy, just as erythropoietin is used to treat the anemia associated with cancer drugs. Beyond the primary indication for treating hair loss in cancer patients, there is a large audience of men and some women who would not object to using a product to counteract baldness. The latter use is perfectly legal and would constitute a so-called off-label use of the product, a practice commonly allowed by FDA. As an example, it is well known that certain anticonvulsant drugs are regularly used to effectively treat patients with a variety of affective disorders, although the original drug approval was solely for the treatment of seizures.

What about other scenarios? Recombinant proteins have already been used in settings that could be described as enhancement. Some young children of very small stature have been treated with the human growth hormone (HGH) protein, even though there were no data to support HGH deficiency. Admittedly, short stature in males could be a serious social problem, but it does not constitute a disease. One could envision the use of the gene encoding HGH instead of using the protein. Another area that is vulnerable to enhancement is the world of athletics where the participants are always seeking to gain a competitive advantage. An athlete wishing to increase his/her muscle mass might be more than willing to use the gene encoding for vascular endothelial growth factor (VEGF) to increase the blood supply to this tissue or perhaps to use a combination of VEGF and a gene encoding for one of the dystrophins to further augment bulk and muscle strength.

Undoubtedly, there are a number of risks in using gene transfer either for disease treatment or for enhancement, but it is much easier to justify said risks when one is attempting to treat a serious disease. The fact that people will take risks for a purely cosmetic result attests only to personal interests. Should this technology be made available for enhancement purposes with a warning label attached? It would be difficult to find a consensus on this point. However, the matter of control could turn out to be a very elusive one. One nation or series of nations could elect to ban gene transfer for enhancement, but another sovereign state could take advantage of a commercial opportunity, knowing that there would be a ready supply of customers for whom personal expense and travel are negligible barriers.

Various writers have commented on the matter of gene transfer for various types of intellectual enhancement including efficiency of memory, general cognitive ability or intelligence, or other behavioral traits such as antisocial behavior (38). While such treatises provide fascinating reading and address such issues as the fundamental importance of individual liberty, they fail to take cognizance of the scientific infrastructure. If one considers that our ability to treat single gene deficiency diseases is in a most rudimentary state, then it stands to reason that alteration of polygenic diseases or personality traits must be consigned to the rather distant future. To achieve these types of changes will require an intrinsic knowledge of the interactions among multiple sets of genes plus the ability to control gene expression, another area in which current knowledge is extremely lacking. There is no reason to

assume that some of the critical scientific knowledge will never be available, but gene transfer for various types of intellectual enhancement awaits quantum leaps in our knowledge of molecular genetics.

CONCLUSION

In viewing the end of the twentieth century, it is apparent that the past 50 years have represented the flowering of molecular genetics. There have been many important applications of the basic research in this area, and none is more potentially exciting than somatic cell gene therapy. While the early gains have been modest, there is every reason to assume that gene therapy will have a major impact on the practice of medicine within the next 20 to 25 years. To the credit of the scientific community and the public, there have been extensive public discussions of this technology, and the first national oversight program was a completely public process. When Marshall Nirenberg expressed his concerns in 1967, he alluded to the necessity of an informed society to make appropriate decisions about biochemical genetics. There have been systematic attempts to create that informed society and until now, there has been a pattern of responsibility in place. The challenges of the future are to protect the rights of the patient as additional disease states are subject to intervention, and to carefully assess the societal consequences of using gene transfer for nonmedical enhancement.

BIBLIOGRAPHY

1. H.I. Miller, *Hum. Gene Ther.* **1**, 3–4 (1990).
2. O.T. Avery, C.M. MacLeod, and M. McCarty, *J. Exp. Med.* **70**, 137–158 (1944).
3. J.D. Watson and F.H.C. Crick, *Nature (London)* **171**, 964–969 (1953).
4. H.F. Judson, *The Eighth Day of Creation: The Makers of the Revolution in Biology*, Simon & Schuster, New York, 1979.
5. D. Nathans and H.O. Smith, *Ann. Rev. Biochem.* **44**, 273–293 (1975).
6. S.N. Cohen et al., *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3240–3244 (1973).
7. D.J. Jolly, J.K. Yu, and T. Friedmann, *Methods Enzymol.* **149**, 10–25 (1987).
8. J.A. Wolff and J. Lederberg, *Hum. Gene Ther.* **5**, 469–480 (1994).
9. E.L. Tatum, *Perspect. Biol. Med.* **10**, 19–32 (1966).
10. J. Lederberg, *Proc. World Cong. Fertil. Steril., 6th, 1968*, 18–23 (1970).
11. A. Kornberg, in M. Burnet, ed., *Genes, Dreams, and Reality*, Basic Books, New York, 1971, p. 71.
12. H.V. Aposhian, *Perspect. Biol. Med.* **14**, 987–1008 (1970).
13. M.W. Nirenberg, *Science* **157**, 633 (1967).
14. R. Sinsheimer, *Eng. Sci.* **32**, 8–13 (1969).
15. S. Rogers and M. Moore, *J. Exp. Med.* **117**, 521–542 (1963).
16. H.G. Terheggen et al., *J. Exp. Med.* **119**, 1–3 (1975).
17. S. Rogers, in M. Lappe and R.S. Morison, eds., *Ethical and Scientific Issues Posed by Human Uses of Molecular Genetics*, New York Academy of Sciences, New York, 1976, pp. 66–70.
18. W.F. Anderson, in M. Hamilton, ed., *The New Genetics and the Future of Man*, Eerdmans, Grand Rapids, MI, 1972, pp. 109–124.
19. T. Friedmann and R. Roblin, *Science* **175**, 949–955 (1972).
20. P. Berg et al., *Science* **185**, 303 (1974).
21. Guidelines for Research Involving Recombinant DNA Molecules (NIH Guidelines), *Fed. Regist.* **41**, 27902 (1976).
22. M.J. Cline et al., *Nature (London)* **284**, 422–425 (1980).
23. N. Wade, *Science* **210**, 509–511 (1980).
24. N. Wade, *Science* **212**, 1253 (1981).
25. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life: The Social and Ethical Issues of Genetic Engineering with Human Beings*, U.S. Government Printing Office, Washington, DC, 1982.
26. National Institutes of Health (NIH), *Minutes of the Recombinant DNA Advisory Committee*, January 30, NIH, Washington, DC, 1989, pp. 6–14.
27. National Institutes of Health (NIH), *Minutes of the Recombinant DNA Advisory Committee*, October 6, NIH, Washington, DC, 1989, pp. 25–27.
28. National Institutes of Health (NIH), *Minutes of the Recombinant DNA Advisory Committee*, February 5, NIH, Washington, DC, 1990, pp. 5–14.
29. National Institutes of Health (NIH), *Minutes of the Recombinant DNA Advisory Committee*, March 30, NIH, Washington, DC, 1990, pp. 4–7.
30. National Institutes of Health (NIH), *Minutes of the Recombinant DNA Advisory Committee*, July 31, NIH, Washington, DC, 1990, pp. 3–17.
31. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology — Background Paper: Public Perceptions of Biotechnology*, U.S. Government Printing Office, Washington, DC, 1987.
32. March of Dimes Birth Defects Foundation, *Genetic Testing and Gene Therapy Survey: Questionnaire and Responses*, March of Dimes Birth Defects Found., White Plains, NY, 1992.
33. National Institutes of Health (NIH), *Minutes of the Recombinant DNA Advisory Committee*, February 10, NIH, Washington, DC, 1992, pp. 37–41.
34. National Institutes of Health (NIH), *Minutes of the Recombinant DNA Advisory Committee*, December 2, NIH, Washington, DC, 1993, pp. 51–53.
35. National Institutes of Health (NIH), *Minutes of the Recombinant DNA Advisory Committee*, September 12, NIH, Washington, DC, 1994, pp. 43–48.
36. Ad hoc Review Committee of the Recombinant DNA Advisory Committee, *Executive Summary of Findings and Recommendations*, 1995.
37. Guidelines for Research Involving Recombinant DNA Molecules (NIH Guidelines), *Fed. Regist.* **62**, 59032 (1997).
38. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997.
39. J. Katz, A.M. Capron, and E.S. Glass, *Experimentation with Human Beings: The Authority of the Investigator, Subject, Professions, and State in the Human Experimentation Process*, Russell Sage Foundation, New York, 1972.
40. A.C. Mastroianni, R. Faden, and D. Federman, *Women and Health Research: Ethical and Legal Issues of Including Women in Clinical Studies*, National Academy Press, Washington, DC, 1994.

41. N. Wivel and W.F. Anderson, in T. Friedmann, ed., *The Development of Human Gene Therapy*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY, 1999, pp. 1–19, 671–689.
42. Protection of Human Subjects (45 CFR 46), *Fed. Regist.* **48**, 9269 (1983).
43. L. Thompson, *Time*, June 7, 1993, pp. 50–53.
44. L. Jaroff, *Time*, May 31, 1993, pp. 56–57.
45. G. Kolata, *N.Y. Times*, December 13, 1994, p. C1.
46. Office of Recombinant DNA Activities, ORDA Rep., *Human Gene Therapy Protocols*, National Institutes of Health, Bethesda, MD, 1998.
47. A. Pollack, *N.Y. Times*, August 4, 1998, p. B1.
48. Royal Commission on New Reproductive Technologies, *Proceed with Care*, vol. 2, Minister of Government Services, Ottawa, Canada, 1993, p. 945.
49. United Kingdom, Committee on the Ethics of Gene Therapy, Report. HM Stationery Office, London, 1992, p. 22.
50. L. Li and R.M. Hoffman, *Nat. Med.* **1**, 705–706 (1995).

See other GENE THERAPY entries.

GENE THERAPY, LAW AND FDA ROLE IN REGULATION

RICHARD A. MERRILL
University of Virginia
Charlottesville, Virginia

GAIL H. JAVITT
Johns Hopkins School of Hygiene and
Public Health
Baltimore, Maryland

OUTLINE

INTRODUCTION

The discovery of methods to "recombine" DNA from different species, that is to transfer genes from one organism to another, has opened a new frontier for therapeutic medicine. It offers the prospect that some diseases and disorders — those known to be caused by specific and identifiable genetic defects — can be corrected and even prevented through skillful intervention in the body's genetic instructions. But scientists and investors who perceive the potential of this new technology have to confront not only daunting technical challenges and enormous biological uncertainties; they have to take account of societal regulation of new technology.

As this article describes, primary responsibility for such regulation at the federal level has fallen to the U.S. Food and Drug Administration (FDA), a unit of the Department of Health and Human Services (HHS). Congress first enacted national food and drug legislation nearly a century ago (1) to "protect consumers from dangerous products" by providing uniform federal regulation of therapeutic drugs (2). The FDA derives its current authority from several dozen laws passed by Congress over the ensuing 90 years. The most prominent of these is the Federal Food, Drug, and Cosmetic Act (FD&C Act), enacted in 1938 (3) and substantially revised in 1962 (4), 1976 (5), 1990 (6), and 1997 (7). This statute provides the basic framework for FDA regulation of medical products defined as "drugs" or "medical devices."

FDA also administers the Biologics Act of 1902 (8), now codified as part of the Public Health Service (PHS) Act (9). The Biologics Act gave federal officials authority over "biological products," which include "any virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, or analogous product" (10). Responsibility for administering this and another provision of the PHS Act, aimed at preventing the transmission of communicable disease, was originally vested in the National Institutes of Health (NIH) but it was transferred to the part of FDA now known as the Center for Biologic Evaluation and Research (CBER) in 1972 (11).

Thus FDA has for several decades been the federal agency exclusively responsible for regulating virtually all medical products. It usually exercises this authority by requiring the sponsor of a new treatment modality to demonstrate, through carefully controlled clinical trials, that it is safe and reliably produces beneficial effects before it can be made available to the medical profession generally. Most of the laws that FDA administers, however, were enacted before the era of biotechnology, and none were enacted with gene therapy specifically in mind. At the time researchers first began exploring the potential of gene therapy, it was therefore by no means clear that FDA would or should be the agency responsible

for its oversight or, if the agency were to have a role, that it possessed the appropriate tools for regulation.

At the outset another agency, also part of HHS, claimed a dominant interest in the subject. This was the NIH, which in 1974 established the Recombinant DNA Advisory Committee (the "RAC"), an interdisciplinary body, to review and at least tacitly approve federally funded research using the techniques of recombinant DNA technology (12). Because the early experiments did not involve human subjects, FDA did not initially seek to play a significant oversight role. When research began to move into the clinic, however, the RAC lost its initial regulatory monopoly.

FDA saw an important role for itself in regulating gene therapy, derived from its statutory responsibility for assuring the safety and effectiveness of commercially distributed therapeutic products. If the materials used in gene therapy fit the statutory definition of drug, device, or biological product — and, given the breadth of those definitions, it would be hard to assert that they did not — FDA was obligated to address them. And, because FDA has long overseen the clinical investigation of the medical products whose marketing it regulates, this has meant that the agency has major responsibility for monitoring the experimental applications of gene therapy as well. Reasoning in this way, FDA eventually made clear its intention to exert regulatory oversight and equipped itself, administratively and in personnel, to perform that role.

To many, FDA's assertion of responsibility appeared to duplicate, if not compete with, the role of the NIH RAC, and, more specifically, with the RAC's Gene Therapy Working Group. The first part of this article describes the uneasy relationship between the two bodies, a relationship that some saw as "rivalry" and others as natural evolution as research moved from the laboratory into the clinic. By the early 1990s FDA's claim to the primary oversight role had been established, while the RAC continued to conduct what many regarded as duplicative review of individual protocols. By 1997, however, FDA and NIH had agreed that FDA would be exclusively responsible for approving individual gene therapy protocols, and that the RAC would no longer review protocols for their technical merit. At the same time the RAC would continue its recognized role as the forum for public discussion of the social and ethical issues raised by novel applications of, or approaches to, gene therapy.

The second part of the article describes FDA's evolving regulatory policy and analyzes the requirements it has imposed. It is noteworthy that FDA has never promulgated formal regulations specifically addressing gene therapy products; all of its policy pronouncements have taken a less formal guise. This approach is consistent with FDA's position that gene therapy products are simply another type of drug, biologic, or device. But it also reflects a regulatory regime that is still a work in progress. The broad framework of FDA regulation of gene therapy will be familiar to any student of FDA drug law, but the decisions made within that framework betray the complexity and novelty of the technology involved.

EVOLUTION OF FDA REGULATION OF GENE THERAPY

1974–1984: Prehistory

In 1974 scientists announced that they had developed a method to recombine DNA from different species (e.g., two different types of microbes or viruses) to form new biological entities (13). Concerns about the potential impact of this new capability led to an event unprecedented in the history of science: Researchers themselves called for a moratorium on recombinant DNA experiments pending public review of its risks (13,14). The same year, the National Academy of Sciences established a Committee on Recombinant DNA Molecules to examine the risks associated with recombinant DNA research and to recommend specific precautionary measures (15). The NAS Committee's recommendations were published in the journal *Science* in July 1974 (16). The Committee recommended, among other things, that certain experiments be voluntarily deferred, and that the director of the NIH establish a committee to evaluate hypothetical risks, to develop procedures to minimize the spread of recombinant DNA molecules, and to recommend guidelines to be followed by investigators (16). Within four months the Department of Health Education and Welfare (now HHS) chartered the RAC and directed it to establish appropriate biological and physical containment practices and procedures for recombinant DNA research (12). In 1976 the RAC codified these practices as "guidelines," which had to be followed in all research conducted using NIH funds (17).

The scientists who in 1974 had called for a moratorium on recombinant DNA research had done so in part because of fears that a genetically modified organism inadvertently released into the environment could become an "Andromeda Strain" (13). The RAC's early preoccupation with containment methods accordingly reflected these concerns. While several authors called attention to the therapeutic, as well as the social and ethical, implications of genetic manipulation of the human genome (18), it was not until 1980 that gene therapy drew public attention. In that year an American physician, Dr. Martin Cline, chief of the division of hematology/oncology at the University of California at Los Angeles (UCLA), conducted unauthorized gene therapy experiments in Israel and Italy on two women suffering from beta thalassemia, a rare but often fatal genetic defect affecting the red blood cells (19). Cline had submitted his protocol to UCLA's Institutional Review Board (IRB) but neglected to wait for a decision. He later maintained that at the time he carried out the experiments he fully expected to receive approval (20). The UCLA IRB, however, ultimately rejected the protocol, citing insufficient animal studies (19). In addition Cline failed to inform authorities in Italy and Israel or the patients that the protocol entailed the first-ever purposeful insertion of recombinant DNA into humans (19).

Dr. Cline's experimental treatments did not improve or exacerbate the patients' condition but his own professional career was severely compromised. After investigating the incident, NIH issued a report censuring Cline (21), whose research it had supported. NIH canceled Cline's NIH

funding and attached a copy of its report to his subsequent applications for grant support (19).

Following this episode three major religious organizations — the U.S. Catholic Conference, the Synagogue Council of America, and the National Council of Churches — wrote to President Carter expressing concern about rapid advances in genetics and the absence of a federal oversight mechanism (22). The letter was referred to the recently formed President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. In November 1982, before a congressional hearing chaired by then-Representative Albert Gore, Jr., the President's Commission released its assessment, "Splicing Life" (23). The report saw a need for an oversight body to review gene therapy experiments, and also recommended the formation of a permanent federal bioethics commission (23). Speaking to the merits, the report took the position that somatic cell gene therapy — which does not affect the genetic material of reproductive cells — should be permitted to proceed, whereas germ-line gene therapy should not be undertaken without prior public debate (22).

1984–1989: Early Rivalry Between FDA and NIH

In 1983 the RAC formed the Working Group on Human Gene Therapy to study and respond to the report of the President's Commission. The working group comprised three laboratory scientists, three clinicians, three ethicists, three lawyers, two public policy specialists, and a representative from the public. It recommended that the RAC add to its purview experiments involving the "Deliberate transfer of recombinant DNA or DNA derived from recombinant DNA into human subjects" (24). It also drafted a document that it titled "Points to Consider," a nearly 4000 word instruction manual detailing the information that researchers seeking the RAC approval for gene therapy experiments in humans would have to submit (24).

The RAC's "Points to Consider" required researchers to address the scientific aspects of their protocol, including the research design, the anticipated methods of gene delivery, and the results of animal experimentation. In addition researchers were instructed to address a broad range of social concerns pertaining to gene therapy, signaling the RAC's intent to judge not only the scientific validity of human gene therapy protocols but also their social acceptability. The RAC's commitment to public review of all protocols was underscored by its expectation that the first proposals submitted for RAC review would contain no proprietary information or trade secrets, which would enable all aspects of the review to be open to the public (24). An initial draft of the "Points to Consider" document was unveiled on January 22, 1985, and was published in the Federal Register for public comment (24).

Although Congress enacted no legislation directing the formation of the RAC Working Group or authorizing NIH oversight of gene therapy protocols, the RAC assumed this role with substantial congressional as well as public support, and most notably the support of then-Senator Albert Gore, Jr. As one commentator noted, "In the absence of any other duly constituted body, the Working Group on Human Gene Therapy has become the locus for broad social discussion of [gene therapy] issues" (25).

FDA officials, however, were not sanguine about NIH's determination to take the lead regulatory role nor content with the approach it outlined for reviewing gene therapy protocols. NIH officials may have believed that FDA would not seek to participate in the regulation of gene therapy protocols, except for the individual contributions of Dr. Henry Miller as FDA liaison to the RAC, but if so, they proved to be mistaken (26).

On December 31, 1984, three weeks before the RAC's draft "Points to Consider" was published for public comment, FDA issued a policy statement that both dispelled doubt about its intention to regulate clinical trials of gene therapy products and that implied dissatisfaction with NIH's contemplated regulatory approach (27). The policy statement set forth FDA's position that existing laws conferred sufficient authority to regulate all the commercial applications of biotechnology within its jurisdiction. Furthermore, the statement announced FDA's view that gene therapy was not a fundamentally different therapeutic modality that required special scrutiny or new oversight mechanisms. According to the policy statement, "Nucleic acids used for human gene therapy trials will be subject to the same requirements as other biological drugs." The policy statement cautioned against the adoption of "[i]nconsistent or duplicative domestic regulation" of biotechnology that could "put U.S. producers at a competitive disadvantage." Yet, in its sole reference to NIH's oversight of gene therapy protocols, the statement acknowledged that "[i]t is possible that there will be some redundancy between the scientific reviews of these products performed by the National Institutes of Health and FDA" (27,28).

During 1985 it became apparent that a jurisdictional rivalry was brewing between NIH and FDA (26,29). The first gene therapy protocols were expected to be ready for clinical trials within the year (30). Perhaps in anticipation of their arrival, Dr. Henry Miller, Commissioner Frank Young's adviser on biotechnology issues and later head of FDA's Office of Biotechnology, publicly criticized the composition of the RAC Working Group, which he thought gave undue prominence to ethicists and lawyers at the expense of scientists and clinicians (26). Dr. Miller also took the position that since non-germ-line gene therapy was not a qualitatively unique form of therapy, physicians planning to initiate clinical trials with human gene therapy should simply file an investigational new drug (IND) application, permitting FDA to review the experimental gene and vector to be used (26). Submission of an IND to FDA is a standard prerequisite for clinical trials of new drugs.

In what some viewed as another FDA effort to undercut the RAC's hegemony, Commissioner Young strongly supported a proposal prepared in the Office of the Assistant Secretary for Health to create a federal Biotechnology Science Board (BSB) within HHS (31). The Board, which would have reported to the Assistant Secretary, would have had broad authority over research and development in genetic engineering. It would also have diminished NIH's authority over gene therapy (26,31). Specifically,

the proposal would have (*1*) barred NIH from publishing a final version of its "Points to Consider" document without BSB review and (*2*) precluded RAC review of clinical trials of genetically engineered products that were also under the jurisdiction of another regulatory agency (31). The proposal was submitted to the White House Office of Science and Technology Policy (OSTP) on August 1, 1985.

This early attempt by FDA to exert control over the regulation of gene therapy research was unsuccessful. The deputy director of OSTP, Bernadine Healy, favored maintaining a strong RAC (31). Critics viewed the BSB proposal as an attempt by FDA Commissioner Young — who was a supporter of research in the field and a leading candidate to become Assistant Secretary of Health — to become the Reagan administration's "biotech 'czar' " (31). Moreover on August 30, 1984, Senator Gore wrote letters to both Healy and Secretary of HHS Heckler, insisting that they immediately stop FDA's effort to "usurp" RAC's rope in overseeing human gene therapy experiments. Gore's endorsement of the RAC was clear, albeit lukewarm. It had, he wrote, "done an adequate job of addressing the scientific issues to date, and it appears capable of continuing to do so for the immediate future" (32). Gore also forecast the imminent passage of his legislation to establish a national commission on bioethics, which would "undoubtedly consider whether a formalized regulatory structure, like the BSB, for gene therapy will be needed" (32).

The Reagan administration ultimately rejected the BSB proposal in favor of a new interagency committee within the OSTP's Federal Coordinating Council for Science, Engineering, and Technology (FCCSET) (32). The committee would develop scientific policy recommendations for the agencies involved in recombinant DNA research, including NIH and FDA. Thereafter NIH and FDA reached an uneasy truce. The RAC continued to require researchers who contemplated gene therapy experiments to submit their protocols for review addressing the areas identified in the "Points to Consider," while FDA insisted that researchers seeking to conduct gene therapy trials using viral vectors must comply with FDA's IND regulations (31). NIH agreed to modify its "Points to Consider" to state explicitly that they applied only to "institutions receiving support for recombinant DNA research from the NIH" (33). While this limitation had been tacitly acknowledged previously, fear of public criticism and possible liability led unfunded academic and industrial sponsors of biotechnology research to voluntarily submit their projects for the RAC review (31).

The revised "Points to Consider" also included a footnote acknowledging that FDA "has jurisdiction over drug products intended for use in clinical trials for human somatic-cell gene therapy," and directing applicants to review FDA's Policy Statement. This version was published in the Federal Register on August 19, 1985 (33).

In February 1987 the RAC called attention to the continuing rivalry by modifying its "Points to Consider" explicitly to prohibit NIH-funded researchers from undertaking gene therapy clinical trials absent prior RAC review and approval, even if the experiment had been approved by another agency. FDA's Miller opposed this amendment on the ground that demanding separate RAC review would delay expeditious use of gene therapy in patients who might urgently need such treatment (34).

1989–1991: First Gene Therapy Protocols Authorized

The debate over which agency or agencies would regulate gene therapy experiments — and the appropriate parameters of such oversight — had preceded by several years the first authorized clinical trial of gene therapy in the United States. The first clinical experiments were expected as early as 1985, but technical difficulties delayed the first attempt to introduce foreign genes into patients (35). Finally, on May 22, 1989, NIH researchers did so (36). The protocol entailed removing tumor-infiltrating lymphocytes (TIL) from a cancer cell, inserting a bacterial gene, and giving the TILs back to the patient. The experiment was not intended to be therapeutic; rather, the bacterial genes were intended to serve as a "marker" that would permit scientists to trace the path of the tumor-fighting cells (37). Prior to conducting the experiment, the protocol had undergone 26 hours of formal review hearings, including review by 7 advisory committees (36,37). At the eleventh hour, Jeremy Rifkin — fierce opponent of biotechnology in all applications — further delayed commencement of the experiment by filing a lawsuit claiming that NIH had failed to follow proper procedures in approving the protocol (38).

Less than a year later NIH researchers Michael R. Blaese, W. French Anderson, and Kenneth Culver received approval from both the RAC and FDA to conduct the first clinical trial using gene therapy to treat a genetic disorder. On September 14, 1990, genetically modified white blood cells containing a gene for adenosine deaminase (ADA) were introduced into a four-year-old girl with severe combined immunodeficiency disorder (SCID), which is caused by a defect in the ADA gene (13). Children lacking a functional ADA gene cannot mount immune responses and usually die in early childhood. The researchers hoped that the new gene would begin to manufacture ADA in the patient's body and thereby restore immune function.

1991–1993: FDA Steps Up Regulatory Efforts

The first experiments helped quiet public fear and rekindled enthusiasm over the seemingly limitless possible applications of gene therapy to treat, for example, cancer, heart disease, high cholesterol, and AIDS (39). While the initial clinical trials did not unambiguously show effectiveness, the harms to patients that had been feared, such as illness caused by the viral vectors, did not occur. The regulatory environment for manufacturers of the viral vectors used in gene therapy, however, remained very uncertain. Maryland-based Genetic Therapy Inc. (GTI), which had supplied the viral vectors used in the initial gene therapy experiments under cooperative agreements with NIH, acknowledged in its May 1991 prospectus that "the precise regulatory requirements with which the company will have to comply are uncertain at this time due to the novelty of the human gene therapies currently under development" (40).

Anticipating a flurry of INDs from industry, FDA's CBER made available a draft Guidance Document

addressing somatic cell therapy and gene therapy. FDA emphasized that the document was not a binding regulation. Rather, it was intended to inform manufacturers engaged in the production and testing of products for these therapies about the issues that the agency believed should be considered and that, by implication, should be addressed in IND submissions (41). The guidance document recommended that sponsors of experiments address, among other issues, (*1*) quality control procedures, (*2*) procedures to prevent cell culture contamination by adventitious agents, (*3*) proper characterization of gene constructs, and (*4*) vector insertion methods. It also detailed the types of studies FDA believed should be performed to establish product safety.

In December 1992, FDA took another step to formalize its regulatory authority over anticipated gene therapy products. CBER announced its plan to establish, as of January 1993, a new Office of Therapeutics Research and Review (OTRR) that would oversee four new "lab-based" divisions, including a Division of Cellular and Gene Therapies (42). Creation of OTRR would allow CBER to remain "on the cutting edge" of "many very novel, very innovative types of therapy," according to then-Deputy Director of CBER, Janet Woodcock, M.D. (43). Woodcock acknowledged that there were then no marketed products under the purview of the new Division, but explained that FDA reviewers would be preparing "to deal with these burgeoning new types of therapeutic approaches" as they became ready for submission. Dr. Philip Noguchi was designated to head the division.

In October 1993, FDA issued another guidance document explaining how the agency's statutory authorities applied to human somatic cell therapy and gene therapy products. FDA stated that it was publishing the guidance document "in response to requests that the agency clarify its regulatory approach and provide guidance to manufacturers of products intended to be used in somatic cell therapy or gene therapy" (44). As the trade press observed, the document "essentially codifies current FDA practice to regulate . . . all gene therapies" (45).

Contemporaneously FDA Commissioner David Kessler took another step to explain the agency's evolving policy. Deploying a tactic that became familiar during his tenure, Kessler and several FDA officials within OTRR published an article in the *New England Journal of Medicine*. Their professed objectives were to "examine the regulation of somatic-cell and gene therapy by the Food and Drug Administration (FDA) in the context of the agency's traditional role in the development of biologic products and to stimulate discussion in areas in which policy is still being formulated." The tone and scope of the article suggested that another goal was to underscore FDA's centrality in the gene therapy approval process. The authors, however, professed no desire to supplant the RAC. Rather, they acknowledged that FDA and NIH had "important, complementary functions," and said that RAC review "ensures broad public discussion of the scientific evaluation of this new technology, particularly with regard to social and ethical concerns." FDA, in contrast, "focuses on the development of safe and effective biological products, from their first use in humans through their commercial distribution." Therefore "[p]roducts used in protocols subject to review by the Recombinant DNA Advisory Committee must also undergo FDA review" (46). This endorsement of the RAC's unique role as a forum for public debate of the societal and ethical issues raised by gene therapy proved prophetic.

1991–1994: The RAC Move Toward Streamlined Review

As FDA elaborated its regulatory approach, the RAC saw signs that its primary role in overseeing gene therapy research might be in jeopardy. At a July 31, 1991, meeting of the RAC, researcher Dr. W. French Anderson urged RAC members to combine the functions of the Human Gene Therapy Subcommittee with that of the full RAC. Lamenting that the lengthy and redundant review process for gene transfer protocols was a "difficult experience" for investigators, Anderson predicted that investigators with private sources of funding would seek to evade the RAC review and submit their protocols only to FDA for review to avoid the associated regulatory burdens (47). Anderson expressed the concern, however, that the loss of the public forum provided by the RAC might jeopardize public confidence in gene therapy research. He emphasized that, unlike the RAC, FDA conducted review of all submissions in private (48).

Soon after Anderson's cautionary warning about "redundancy," Viagene, a biotech company, seemed poised to prove his point. In May 1992, Viagene filed in IND with FDA for a Phase I (preliminary) gene transfer protocol aimed at preventing the onset of AIDS in HIV positive patients. The protocol had not been submitted to the RAC. Nevertheless, in June 1992, an FDA advisory committee recommended that FDA approve Viagene's clinical trial, representing the "first time that such an experiment has been treated as a routine drug trial rather than a foray into unknown territory that requires extraordinary safety and ethics review" (49). Commenting on this apparent sea change in the U.S. regulatory process, the British journal *Nature* opined that, "in the face of increasing acceptance of gene therapy, those who still wish for special reviews of everything from the basic biology to the theology of simple gene-transfer experiments appear to be losing ground to those who argue for business as usual" (49). *Nature* was premature in predicting the demise of the RAC review: The following year Viagene voluntarily submitted a related HIV gene therapy protocol to the RAC. The company explained that it expected that future testing of the gene therapy product would involve NIH-funded institutions, and it wished to have the RAC involved from the outset (50). Criticism from within the industry for its attempt to circumvent the RAC review process may also have played a role in its about-face (51).

In the face of (*1*) FDA's expanding role in overseeing gene therapy protocols, (*2*) the increasing number of protocols being submitted by both research institutions and industry, and (*3*) diminishing public concern over gene-based therapies, members of the RAC debated what the committee's oversight role should be. Some felt that parallel review by FDA and NIH was unnecessary, while others asserted that the RAC should focus on cross-cutting policy issues relating to gene therapy and on public education.

The committee members agreed that review of individual research protocols needed to be streamlined (52).

A July 1994 proposal by the National Task Force on AIDS Drug Development forced both the RAC and FDA to reconcile their overlapping regulatory efforts. Chaired by HHS Assistant Secretary for Health Philip Lee and including both FDA Commissioner Kessler and NIH Director Harold Varmus, the Task Force unanimously approved a proposal to streamline gene therapy protocol reviews. Under the proposed review process, all gene transfer protocols would be submitted directly *to* FDA and in the format required *by* FDA. If the RAC review were also deemed necessary, it would accept the same format (53).

Upon receipt of an application, FDA would begin its review. Simultaneously NIH's Office of Recombinant DNA Research (ORDA) would evaluate the protocol to decide whether RAC review was necessary. The Task Force proposal identified several factors that should be considered in making this determination, including whether the protocol (*1*) employed novel approaches, (*2*) involved a new disease, (*3*) involved unique applications of gene transfer, or (*4*) involved other issues requiring public review (53). Finally, the proposal recommended eliminating the RAC's Points to Consider document.

1994–1996: FDA Assumes Lead Role in Review of Gene Therapy Research

While both Kessler and Varmus agreed to the Task Force proposal, support for it was not universal. Criticism came from both extremes. Andrew Kimbrell, attorney for Jeremy Rifkin, declared that "[t]his is not the time for decreased reviews," and argued that the government should be doing more to follow up early experiments to discover possible late-developing harms from gene therapy (54). Kimbrell threatened a lawsuit if the RAC's regulatory role were diminished. A more eloquent defense of the RAC's continued relevance was offered by gene therapy pioneer W. French Anderson in his journal *Human Gene Therapy*. Anderson asserted that the "public's trust in this new experimental treatment is in large part the result of, and maintained by, the RAC's resolute role as a watchdog" (55).

Taking an opposing view, Henry Miller, now the former FDA Director of Biotechnology, maintained that the RAC review was an anachronism now that gene therapy experiments involved human subjects and potential products. He disdained the HHS effort to streamline reviews as "purely cosmetic" (54). Leonard Post, a former member of the RAC and an industry spokesman, supported Miller's view that RAC's review of gene therapy protocols unnecessarily duplicated FDA regulation (56).

Nor did members of the RAC favor the compromise. Although they reluctantly voted to approve in principle the consolidation of FDA and RAC reviews (57), they vehemently opposed deleting the "Points to Consider" from the RAC guidelines. In response to their criticism of certain aspects of the Task Force proposal, FDA's Noguchi offered a compromise, under which the RAC's "Points to Consider" would be retained and FDA, ORDA, and RAC would together determine the need for the RAC review of individual clinical protocols (58). NIH Director Varmus emphasized that he did "not intend to see the demise of RAC" but concluded that the increased number and diversity of gene therapy protocols requiring review necessitated coordination with FDA (58). He noted that numerous investigators had complained to the Task Force on AIDS Drug Development that the review process was too slow and was delaying the initiation of new approaches to the treatment of HIV. The RAC met only episodically while FDA received and reviewed INDs every day. Varmus also acknowledged receiving "a long series of complaints . . . from my colleagues in the field who claim that there [have] been undue delays in reviews of protocols in RAC." Admitting that the RAC's criteria for approval were unclear, Varmus directed the committee to establish an ad hoc review panel to examine its criteria for approving gene therapy protocols (59). This committee, which was headed by the Salk Institute's Inder Verma, Ph.D., a future RAC member, and became known as the "Verma Committee" (60).

Into 1995 the RAC and FDA wrestled with the details of coordinating review of experimental protocols, as advocates for industry and patient groups demanded quicker review (61). The irony of the situation was not lost on W. French Anderson, who wrote "this was the first time in history that anyone had wanted to go to the FDA because another federal review process was too slow!" (55). RAC member Alexander Capron found it "puzzling" that "AIDS activists want streamlining of gene therapy protocols, but others say we're going too fast" (61).

In March 1995 the members of the RAC unanimously approved a proposal for consolidated protocol review, which allowed simultaneous submission of protocols to FDA and NIH and incorporated elements of both the original Task Force proposal and the Noguchi compromise. Under the new process the RAC and FDA would receive simultaneous submissions of all protocols, which would include information required by RAC's "Points to Consider" document. FDA and ORDA would then decide which protocols merited RAC review. These would include protocols using new vectors or methods of gene delivery, targeting new diseases, employing unique applications of gene transfer, or raising ethical issues warranting public review (62). However, the committee members rejected a proposal by Viagene that would have further streamlined review. In a February 24, 1995, letter, Viagene Regulatory Affairs Director Sheryl Osborne recommended that the RAC not review expansions of previously approved Phase I gene therapy trials in Phase II and III, that sponsors be permitted to seek concurrent IRB and FDA/RAC approval, and that the NIH director be required to complete protocol review within 15 calendar days (62).

FDA's Noguchi initially praised the compromise, stating that it could cut industry's waiting time for approval from a minimum 8 months to as little as 45 days. He also announced FDA's plan to establish a registry to track every American who received gene therapy treatment (63). Within a few months, however, Noguchi was complaining that the RAC had failed to reduce review times and warning that continued FDA funding of the patient registry was in jeopardy because of the RAC's lack of cooperation. In a memorandum to the

RAC, Noguchi asserted that a "year after the mandate to streamline the ... process ... I am obligated to say that the RAC has not appeared to be as accommodating or committed to sharing tasks," such as assisting in development of the database. In addition he noted that "modest proposals to further streamline the RAC process, such as allowing concurrent review by the local IRBs have been unanimously disdained." Noguchi stated that he had "taken a lot of heat from my superiors for their impressions that FDA funding of the gene therapy registry has been for the NIH benefit, rather than for the FDA. Sadly, I reluctantly concur with that impression" (64).

The RAC's protocol review role was soon to be further curtailed, for a number of reasons. First, notwithstanding the official compromise between the RAC and FDA, industry spokespersons remained confused over when RAC review was and was not required (65). In addition the RAC's Verma Committee recommended that the RAC devote more attention to policy issues while delegating scientific review of most protocols to FDA (66). Furthermore proponents of biotechnology, like Henry Miller, continued to rail against the RAC's inefficiency and lack of technical expertise, and to argue that overregulation of gene therapy would inevitably delay development of new therapies (67). Miller criticized NIH Director Varmus for failing to eliminate the RAC (67). Pharmaceutical industry representatives concurred that the RAC review of gene therapy protocols was unnecessary. For example, Washington, DC, attorney Bruce Mackler opined that "[i]t is hard to identify in a quantifiable fashion a uniqueness of gene therapy risks and benefits that would call out for a particular regulation that differs from other biological products." Mackler contended that "the RAC's domain should be ... assuaging public fears and advancing scientific quality of knowledge" (68).

1996–1997: The RAC's Role in the Aftermath of Consolidation

During 1996 NIH Director Varmus considered three options: (*1*) terminate the RAC, (*2*) maintain consolidated FDA and NIH review, or (*3*) maintain the RAC as a vehicle for public accountability and access to information, while terminating its role in the review and approval of individual protocols. After much internal debate Varmus initially chose to pursue elimination of the RAC and transfer of its public policy functions to a new body to be developed within NIH. Following public protestations against elimination of the RAC, however, Varmus decided to retain the RAC and also to maintain simultaneous submission of gene therapy protocols to the RAC and FDA, even though only FDA would actually review individual protocols under most circumstances. Significantly, FDA would have exclusive authority to approve gene therapy protocols (after review and approval by local IRBs).

On July 8, 1996, NIH published a Notice of Intent announcing that NIH was considering amending its Guidelines to eliminate the RAC's review of most gene therapy protocols and to transfer *all* approval responsibility for such protocols to FDA (69). At the same time NIH stressed that its oversight of gene therapy would be "enhanced" through three new mechanisms: (*1*) the establishment of the Office of Recombinant DNA Activities Advisory Committee (OAC)—a 6 to 10 member interdisciplinary body that would "ensure public accountability for recombinant DNA research and relevant data"; (*2*) implementation of Gene Therapy Policy Conferences to "augment the quality and efficiency of public discussion of the scientific merit and the ethical issues relevant to gene therapy clinical trials"; and (*3*) "continuation of the publicly available, comprehensive NIH database of human gene transfer clinical trials," including adverse event reporting.

NIH explained its plan as simply another step in the normal "devolution" of NIH authority as scientists gained experience with gene therapy protocols and the concerns about the technology therefore shifted. Specifically, NIH characterized the current plan as a continuation of the 1995 decision to consolidate FDA and NIH protocol review:

> In 1995, a similar devolution of NIH oversight of human gene therapy occurred. By this time, the RAC had reviewed and approved 113 gene therapy protocols and over 1,000 patients had been enrolled in worldwide trials. The RAC, the scientific community, and the public had a substantial base of information regarding the use and safety of many of the vectors employed in, and target diseases addressed by, human gene therapy. Subsequent analyses revealed that the human health and environmental safety concerns expressed at the inception of gene therapy clinical trials had not materialized (70).

NIH described its further curtailment of RAC's protocol review function as resulting from the very success of the 1995 consolidation:

> Since the implementation of consolidated review in July 1995, only six of the 36 protocols submitted to ORDA required RAC review and approval; and five of those six protocols were already in the system before consolidated review. The consolidated review process proved to be so successful in eliminating the need for RAC review and approval, that NIH canceled both the March and June 1996 RAC meetings due to the lack of novel protocols requiring RAC attention (70).

Two features of the NIH notice are particularly noteworthy. First, NIH repeatedly stressed that the proposal should not be viewed as simply an "elimination" of the RAC but rather as a reallocation of NIH resources to avoid regulatory redundancy and permit NIH to carry out its role in leading public discussion: "Eliminating RAC protocol approval reduces duplication of effort with the FDA while enhancing the time and effort devoted to both ongoing anticipated gene therapy policy issues deserving of substantial public discussion" (70). In addition, refocusing NIH's efforts would assure NIH's ability to maintain public accountability over gene therapy research:

> NIH concludes that it is not the RAC per se that is critical for public accountability, but the system by which NIH continues to provide public discussion of the scientific, safety, and ethical/legal issues related to human gene therapy (70).

Second, NIH for the first time publicly acknowledged that only FDA possessed the statutory authority to review and approve gene therapy protocols: "The NIH Director has concluded that the current proposal ... is timely and

appropriate based on ... the duplication of review and approval by the NIH while the FDA holds the statutory authority" (70). Similarly, NIH acknowledged that, while the contemplated Gene Therapy Policy Conferences would enhance NIH's ability to provide advice on policy matters pertaining to gene therapy, "[t]he NIH cannot ... give the RAC, or any other NIH standing or ad hoc body, the authority to give policy advice or make recommendations to the FDA" (71).

In response to its Notice of Intent, NIH received 71 comments from individuals or groups reflecting the interests and concerns of academe, industry, patient advocacy groups, consumer advocacy organizations, professional scientific societies, ethicists, other federal agencies, NIH-funded investigators, past and present RAC members, and private citizens (72). According to NIH, of the 61 comments that addressed the proposal to terminate the RAC, 20 expressed support and 41 opposition (73).

The Biotechnology Industry Organization (BIO) wholeheartedly embraced the NIH proposal, agreeing that NIH review of individual gene therapy protocols had become redundant, and that NIH resources would be "better spent addressing truly novel gene therapies or those raising significant ethical issues such as in utero or germ line gene therapies" (74). Additionally BIO concurred with NIH's proposal to replace the RAC with the smaller OAC: "Since the number of protocols warranting full NIH-RAC review has dropped to close to zero, maintenance of a large panel is no longer necessary" (74). Similar support for the NIH proposal was conveyed in the comments of numerous industry members — ranging from established pharmaceutical companies such as Merck (75), Glaxo Wellcome (76), Parke Davis (77), and Baxter Healthcare Corporation (78) to entrepreneurial biotechnology outfits such as Cell Genesys (79), Genzyme (80), Auragen (81), and IntraImmune Therapies, Inc. (82). Henry Miller applauded the proposal to eliminate NIH oversight of gene therapy protocols as "long overdue" but criticized NIH's proposal to establish the OAC and to maintain a database of gene therapy clinical trials as a step in the wrong direction:

> In the future, NIH should approach gene therapy in a way no different from other kinds of techniques and treatments, except as medical or scientific considerations dictate. Circumstances do not now require retention of ORDA, the creation of a new advisory committee, or the maintenance of a gene therapy database (83).

Opponents of eliminating the RAC — including current and former RAC members, bioethicists, academe, consumer and patient advocacy groups, and some members of the public — were equally fervent in their insistence that the RAC continue to review individual research protocols. They argued, in essence, that the RAC had developed significant expertise in considering the social and ethical consequences of gene therapy experiments and that it had gained the public's trust in carrying out this role. Elimination of the RAC, they feared, would undermine that trust and jeopardize public accountability for or acceptance of gene therapy experiments. They rejected the OAC as a poor substitute that would have neither the experience nor reputation to lead public debate of the social and ethical issues surrounding gene therapy (84). Their concerns were eloquently expressed by a science teacher from Illinois:

> Gene therapy intimately affects the future of the human race and the technology must not be allowed to proceed only in the name of good science. We must work hard to prevent the technology from driving the ethical, moral, and legal issues. While the advisory panel is no guarantee that "what's best for humanity" will drive genetics research and therapy, it does provide a necessary oversight function and should not be disbanded (85).

In response to the number and fervor of comments advocating the retention of the RAC, Varmus retreated, stating that he had "underestimated the historical purpose and significance" of the RAC (86). Varmus offered a "compromise" proposal under which the RAC's size would be reduced but several of its functions maintained (86). However, Varmus contemplated that under the compromise the RAC would no longer participate in reviewing or approving individual gene therapy protocols (86).

On October 31, 1997, NIH announced its final decision. In its revised Guidelines, NIH reduced the RAC from 25 to 15 members while retaining its interdisciplinary composition. The RAC's functions were amended to include (*1*) identifying novel human gene transfer experiments deserving of public discussion by the full RAC, (*2*) transmitting to the NIH director specific comments or recommendations concerning specific experiments or categories of experiments, and (*3*) identifying novel social and ethical issues relevant to specific human applications of gene transfer and recommending appropriate modifications to NIH Guidelines to provide guidance in preparing relevant informed consent documents and in designing and submitting human gene transfer clinical trials (72,87).

The revised Guidelines provided that NIH would relinquish all approval responsibilities for gene therapy experiments to FDA. However, no human gene therapy experiment could proceed if the protocol had not been simultaneously submitted to NIH and FDA. Submissions to NIH would need to comply with NIH's "Points to Consider" (Appendix M), while submissions to FDA would be required to comport with FDA's regulations pertaining to the content and format of IND submissions (88). In addition investigators who had received FDA approval would still be required to report any adverse events to both agencies (89).

The Guidelines provided that submissions to NIH would be for "registration purposes," and would "ensure continued public access to relevant human gene transfer information conducted in compliance with the NIH Guidelines" (72,90). If NIH/ORDA determined that a protocol possessed novel features or raised novel concerns that should be discussed by the full RAC, the principal investigator would be notified. In determining whether the full RAC discussion was warranted, NIH/ORDA reviewers would examine "the scientific rationale, scientific content (relative to other proposals reviewed by RAC), whether the preliminary in vitro and in vivo safety data were obtained in appropriate models and are sufficient, and whether

questions related to relevant social and ethical issues have been resolved" (90). Otherwise, FDA alone would review the protocol. Whether or not the RAC reviewed the protocol, FDA would be solely responsible for determining whether or not to grant approval.

NIH acknowledged that the Guidelines were obligatory for those institutions and investigators receiving NIH funding or entities collaborating with NIH-funded institutions (90). However, it encouraged continued voluntary compliance by all other entities conducting gene therapy research (87).

ANALYSIS OF FDA'S EVOLVING APPROACH TO GENE THERAPY REGULATION

With the October 31, 1997, announcement, NIH formally and finally acknowledged FDA's exclusive statutory authority to approve gene therapy protocols and, by implication, to regulate the development and marketing of gene therapy-derived products, whether privately supported or federally funded. This had been FDA's position from the mid-1980s. FDA's position was that gene therapy protocols represented "business as usual" and that FDA would therefore review gene therapy research protocols — and eventually marketing applications — within the same framework it applied to other medical products.

This part reviews the formal pronouncements that FDA issued describing and justifying this position. It concludes with a brief discussion of the agency's experience applying its "off the shelf" principles of regulation to individual research protocols. The reader should be cautioned, however, that our text does not attempt to provide a full account of what the sponsor of a research protocol will experience in dealing with the agency. This experience varies so much with the technology, the evidence, the agency's current workload, and the personalities of both sponsor and reviewer that no broad generalizations can be reliable. The discussion does not even attempt to describe how FDA will review applications for approval to commercialize gene therapy products. No such applications have come before the agency, and it would be foolish to speculate how it will process the first of these inherently precedent-setting technologies.

Between 1984 and 1998, FDA issued five separate policy documents outlining the agency's plans for regulating gene therapy. These documents provided progressively more explicit guidance for sponsors of gene therapy research regarding the statutory authorities on which FDA relied to regulate their activities, as well as the procedures that should be followed and the clinical data that must be assembled to demonstrate the safety and effectiveness of gene therapy products. From the outset these documents reflected a consistent regulatory approach, one that affirmed FDA's existing legal authorities were broad enough to encompass and flexible enough to fit this emerging technology. There has, in FDA's view, been no need to develop a new regulatory paradigm.

1984 Policy Statement

FDA first described its approach to regulating products derived from biotechnology generally, including gene therapy, on December 31, 1984. The agency's "Statement of Policy for Regulating Biotechnology Products" (27) (Policy Statement) was published for public comment in the Federal Register two months after the RAC Working Group reviewed a draft "Points to Consider in the Design and Submission of Human Somatic-Cell Gene Therapy Protocols," and three weeks before a draft of the RAC "Points to Consider" was published in the Federal Register for public comment. FDA's Policy Statement can therefore be viewed as the agency's response to — and at least mild dissent from — the NIH approach.

The very title of the Policy Statement is revealing; it signaled FDA's (and the Reagan administration's) view, which would be reiterated in subsequent policy documents, that gene therapy should be viewed as merely one example of a much larger class of therapeutic modalities derived from biotechnology, for which FDA already possessed adequate statutory authority and scientific expertise. The introduction to the statement explained:

> A small but important and expanding fraction of the products the Food and Drug Administration (FDA) regulates represents the fruits of new technological achievements. ... It is also noteworthy that technological advancement in a given area may give rise to very diverse product classes, some or all of which may be under FDA's regulatory jurisdiction. For example, new developments in recombinant DNA research can yield products as divergent as food additives, drugs, biologics, and medical devices (27).

The Policy Statement acknowledged that "there are no statutory provisions or regulations that address biotechnology directly" (27), but asserted that the agency's existing regulatory authorities were broad enough to extend to these products and, moreover, that such extension was appropriate:

> The Agency possesses extensive experience with the administrative and regulatory regimens described as applied to the products of biotechnological processes, new and old, and proposes no new procedures or requirements for regulated industry or individuals (27).

Consistent with FDA's existing approach to regulating other medical products, the Policy Statement announced that "the administrative review of products using biotechnology is based on the intended use of each product on a case-by-case basis" (27). FDA thus preserved its discretion to choose among available statutory authorities and to assign administrative responsibility as it judged appropriate.

The statement then proceeded to summarize the basic Act requirements applicable to manufacturers of human drugs:

- Under the FD&C Act, a "drug" is defined to include articles (*1*) "intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease in man," or (*2*) "intended to affect the structure or any function of the body of man" (91). The Act defines a "new drug" as a drug that is not "generally recognized by qualified scientific experts as safe and effective for

its proposed use" (92). Thus the manufacturer of a new drug must establish its safety and effectiveness before marketing. And, to do this, the manufacturer must submit to FDA clinical data from investigations of the drug in human subjects.

- The "sponsor" of a clinical investigation — the entity responsible for the clinical trials — must first file an IND with FDA and obtain the agency's approval to conduct the study or studies described in it (27,93). (This submission follows approval by a research facility's IRB, whose review typically encompasses both ethical and scientific merit.) The IND must contain information sufficient to demonstrate the propriety of testing the drug in human subjects. This would include, for example, drug composition, manufacturing and controls data; results of animal testing, training, and experience of the investigators; and a plan for clinical investigation (27,94). Furthermore the IND sponsor must ensure that informed consent will be obtained from the human subjects who participate in the studies and that the rights and safety of the human subjects will be protected (27,95).
- At the conclusion of three phases of clinical investigation, the manufacturer of a new drug must, if it wishes to distribute the product commercially, submit a New Drug Application (NDA) to FDA. An NDA must contain information including (*1*) full reports of investigations, as well as the results of clinical investigations, demonstrating the drug's safety and effectiveness, (*2*) a list of components of the drug and a statement of the drug's quantitative composition, and (*3*) a description of the methods used in, and the facilities and controls used for, the manufacturing, processing, and packaging of the drug (27,96).

After outlining the framework for regulation of new human drugs, the FDA Policy Statement compared the NDA process to that required for biological products (27,97). FDA's authority to regulate biological products derives chiefly from section 351 of the PHS Act, which defines a "biological product" as:

> any virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, or analogous product . . . applicable to the prevention, treatment, or cure of diseases or injuries of man. . . . (98)

The Policy Statement explained that biological products "are regulated similarly to new drugs during the IND phase" (97). Approval for marketing biological products, however, "is granted by license, which is only issued upon demonstration that both the manufacturing establishment and the product meet standards designed to ensure safety, purity, potency, and efficacy" (97). To obtain approval, the manufacturer must submit a Product License Application (PLA) and include information demonstrating that both the manufacturing facility and the product meet FDA requirements (97). The facility must also pass a prelicensing inspection and be separately licensed (97).

The Policy Statement emphasized that manufacturers of both new drugs and biological products must comply with "good manufacturing practice" (GMP) regulations, which specify requirements for, among other areas, (*1*) manufacturing facilities, (*2*) personnel training, and (*3*) processing methods (97). The Policy Statement also reviewed FDA's quite different statutory authority to regulate medical devices, which do not achieve their intended effect through chemical action in or on the body or through metabolism (99).

After completing its survey of FDA's general requirements for drugs, biological products, and devices, the Policy Statement addressed their application to specific products, including products that are "genetically engineered." Here the agency spoke in vague terms. It reiterated that its approach to regulating these products was "product specific" rather than "technology specific" (27,97,100). Taking genetically altered viruses used as vaccines as an example, the agency declared:

> The composition, concentration, subtype, immunogenicity, reactivity, and nonpathogenicity of the vaccine preparation are all considerations in the final review, whatever the techniques employed in "engineering" the virus (100).

The 1984 Policy Statement addressed gene therapy in just a single sentence: "Nucleic acids used for gene therapy trials will be subject to the same requirements as other biological drugs" (100). This brevity was consistent with agency's position that FDA regulation "must be based on the rational and scientific evaluation of products, and not on *a priori* assumptions about certain process" (100). In other words, that gene therapy might or would entail the use of genetically engineered products to achieve a therapeutic effect would not alter FDA's regulatory approach. The agency would place such a product into the statutory framework that best fit the product's intended use and mode of action.

The Policy Statement was similarly terse in describing the role of the RAC Working Group in reviewing gene therapy protocols: "It is possible that there will be some redundancy between the scientific reviews of these products performed by the National Institutes of Health and FDA" (100).

Certain features of FDA's initial Policy Statement bear emphasis. The document surely cannot be described as a "how to" manual for developers of biotechnology products generally, much less for sponsors of gene therapy protocols. It provided no concrete guidance for sponsors of gene therapy products, save that they would be required to submit and gain approval for an IND before undertaking clinical experiments. Indeed, the Policy Statement did not specify whether products used in gene therapy would be regulated as "drugs," "devices," or "biological products," but instead left open the option to regulate them under any one — and perhaps some combination — of these categories. Of course, given FDA's premise — that gene therapies fell within the agency's customary jurisdiction — this lack of specific guidance was not surprising.

The Policy Statement did not specify which biological component or components used in gene therapy would be considered to be the "drug" or "biological product." The agency thus left open the possibility that it would regulate

the genetic material itself, the vector used to deliver the genetic material, some other component used in the gene transfer process, or all of these.

Also notable was the FDA's oblique criticism of the NIH role in regulating gene therapy. To be sure, the regulatory issues addressed in the Policy Statement overlapped in several respects with those identified by the RAC's "Points to Consider," although these points specified the relevant considerations in substantially more detail. Like FDA's Policy Statement, the RAC's "Points to Consider" emphasized the need for sponsors of clinical trials to address the study's design, results of animal studies, investigator and personnel qualifications, and the adequacy of the laboratory facilities. However, the issues FDA addressed were a subset of a substantially larger range of issues that NIH required sponsors to address — some of which FDA's Policy Statement implied were not necessary and perhaps not even appropriate for its consideration.

For example, under "Description of Proposal," NIH asked investigators to address (*1*) why the disease in question was appropriate for gene therapy and what alternative therapies existed, (*2*) what "equity issues" would likely arise in the selection of patients and how they would be addressed, and (*3*) whether the "innovative character" of gene therapy would be discussed with patients (24,101). In addition, under the heading "Social Issues," NIH asked investigators to address (*1*) the steps that would be taken to ensure that accurate information was made available to the public regarding concerns raised by the study and (*2*) whether the investigator intended to protect the products or procedures used in the proposed study under patent or trade secret laws and, if so, what steps would be taken to permit full communication among investigators and clinicians concerning research methods and results (101). The initial draft of the "Points to Consider" stated that the RAC and its working group would also consider (*1*) whether the proposed study would likely affect the reproductive cells of the patient, (*2*) whether the proposed study was an "extension of existing methods of health care" or represented a "distinct departure from present treatments of disease," and (*3*) whether somatic cell gene therapy would likely lead to germ-line therapy (i.e., therapy affecting reproductive cells and therefore future generations, enhancement of human capabilities through genetic means, or government-sponsored eugenic programs (101). However, the final version, adopted on September 29, 1986, deleted the section addressing germ-line and eugenic implications of gene therapy (102).

FDA's Policy Statement reflected the agency's position that, from its perspective, medical applications of biotechnology generally and gene therapy specifically represented "business and usual." By contrast, NIH's "Points to Consider" signaled the RAC Working Group's belief that gene therapy was a fundamentally new approach to therapy, which posed unique social and ethical questions as well as complex scientific challenges.

Following review of public comments, FDA in June 1986 published a substantially unchanged final version of the 1984 Policy Statement. For the next five years FDA issued no other formal statements regarding gene therapy or its regulation.

FDA's 1991 "Points to Consider"

On August 27, 1991, almost a year after the first clinical experiment employing gene therapy, FDA issued a document entitled *Points to Consider in Human Somatic Cell Therapy and Gene Therapy* ("Points to Consider") (44,103). Unlike its 1984 Policy Statement, which addressed a wide range of applications of biotechnology that could be subject to regulation, FDA's 1991 document focused specifically on gene therapy products (as well as somatic cell therapy products). FDA stated that the purpose of issuing its "Points to Consider" was to provide information to "manufacturers engaged in the production and testing of products for these therapies (103, p. 3). Compared to the 1984 Policy Statement, the FDA "Points to Consider" provided concrete guidance to manufacturers regarding the procedures they should be using to develop their products and the type of supporting data the agency would expect. FDA stressed that its "Points to Consider" did not constitute binding regulations but merely "represent issues that the Center for Biologic Evaluation and Research (CBER) staff believes should be considered at this time" (103, p. 3).

FDA's "Points to Consider" set forth two definitions. Somatic cell therapy was defined as "the administration to humans of . . . living cells which have been manipulated or processed ex vivo" (103, p. 3). Gene therapy was defined as "a medical intervention based on modification of the genetic material of living cells." The agency continued:

> Cells may be modified ex vivo for subsequent administration to humans, or may be altered in vivo by gene therapy given directly to the subject. When the genetic manipulation is performed ex vivo on cells which are then administered to the patient, this is also a form of somatic cell therapy. The genetic manipulation may be intended to have a therapeutic or prophylactic effect, or may provide a way of marking cells for later identification (103, p. 3).

The agency observed that initial approaches to gene therapy "have involved the alteration and administration of somatic cells," but it forecast that "future techniques may include approaches such as the direct administration to patients of retroviral vectors or other forms of genetic material" (103, pp. 3–4).

The FDA document then identified several subjects that sponsors of gene therapy and somatic cell therapy procedures should consider during product development, as well as information that should be presented to the agency. Sponsors are expected to address (*1*) quality control, (*2*) development and characterization of cell populations for administration, (*3*) preclinical testing, (*4*) lot-to-lot manufacturing control and release testing, and (*5*) clinical trials (103). FDA emphasized that it is essential to characterize the gene sequence being introduced into cells and the vector used to do this, and specified that "each distinct vector is considered a different product and should be fully characterized and tested for safety" (103, p. 12). The agency also stated that manufacturers should address the methods used to

insert the vector into cells and the implications of the method used. Specifically, manufacturers should be able to demonstrate that the introduced genes were integrating at the correct location in the chromosome, and that they were functioning appropriately once integrated (i.e., were expressing the correct gene product in the correct quantity) (103).

FDA noted that clinical trials using gene therapy raise "some novel concerns due to the nature of the therapeutic agents" (103, p. 19). For example, use of viral vectors "may in special cases require testing of clinical personnel or household contacts to confirm lack of infectious spread," (103, pp. 19–20). In addition "the product of the inserted gene must be considered as a potential source of immune reactions" (103, p. 20); thus sponsors might in some cases need to document whether such reactions were observed and whether they altered the safety or therapeutic effectiveness of the product (103, p. 20).

While FDA's "Points to Consider" provided details not contained in the 1984 Policy Statement, both documents reflected the same regulatory approach. The "Points to Consider" concluded with the observation that somatic cell and gene therapy raise issues common to all biological products as well as issues unique to these types of therapies. It reflected FDA's confidence in its ability to address these issues, and any new issues that the future might reveal, within the framework of existing legislation (103).

1993 Guidance

Shortly after establishing the new Division of Cellular and Gene Therapy, FDA published in the Federal Register a notice entitled "Application of Current Statutory Authorities to Human Somatic Cell Therapy Products and Gene Therapy Products" (44). This document, though it largely reiterated information contained in the 1984 and 1991 documents, said it was being published "in response to requests that the agency clarify its regulatory approach and provide guidance to manufacturers of products intended to be used in somatic cell therapy or gene therapy" (44).

Like its 1984 Policy Statement, FDA's 1993 self-described "guidance" identified the FD&C Act and the PHS Act as the legal sources of its authority to regulate gene therapy. It also reiterated that clinical trials of biological products — including gene therapy — to gather data on safety and effectiveness must be conducted under an IND. In addition the document stated that a product could be subject to regulation under both statutes concurrently: "Products considered to be biological products subject to the provisions of section 351 of the PHS Act are simultaneously also drugs or devices subject to the applicable provisions under the Act" (44). While FDA asserted that products regulated as biological products must also meet requirements applicable to drugs or devices, it also emphasized that sponsors were not required to submit multiple applications for marketing approval; for any product the agency would require only a biologic application (PLA), new drug application (NDA), or a device application (PMA) (104).

The 1993 guidance offered additional but still general information about FDA's statutory authorities over drugs, medical devices, and biological products — including, for example, its authority to regulate labeling, its requirements for licensing and inspection of manufacturing facilities, and its enforcement powers.

The 1993 guidance for the first time provided concrete examples of treatments that the agency would regulate as gene therapy products, either by the mechanism of a PLA or an NDA:

> Final products containing the genetic material intended for gene therapy are regulated as biological products requiring PLA's (e.g., viral vectors containing genetic material to be transferred, ex vivo transduced cells and analogous products) or as drugs requiring NDA's (e.g., synthetic products) regardless of whether they are intended for use in vivo or ex vivo. Gene therapy products that are licensed biological products will be approved as biological products intended for further manufacture if they are intended to be used ex vivo during the manufacture of genetically altered cells.
>
> Examples include the following: (*1*) A synthetic polynucleotide sequence intended to alter a specific genetic sequence in human somatic cells after systemic administration is regulated as a drug requiring an NDA; (*2*) a retroviral vector containing the adenosine deaminase (ADA) gene, intended to be administered intravenously to the patient, is regulated as a biological product requiring a PLA, and (*3*) a retroviral vector containing the ADA gene and intended to modify cells ex vivo is regulated as a biological product intended for further manufacture requiring a PLA (105).

1996 Addendum

In 1996 FDA issued a draft Addendum to its 1991 "Points to Consider" document (106,107). The Addendum focused on the regulatory requirements for viral vectors used in gene therapy. FDA explained that since the 1991 document was issued, "the range of proposals has expanded to include additional classes of vectors and also the in vivo use of vectors (direct vector administration to patients)" (106). Accordingly it was issuing the Addendum to "provide manufacturers with current information regarding current regulatory concerns for production, testing, and administration of recombinant vectors for gene therapy" (107, p. 2). The Addendum instructed manufacturers regarding proper characterization of the gene sequence of the vector, proper maintenance of the cells used to produce the vectors, and appropriate methods for testing vectors for purity, potency, and safety. The Addendum also addressed issues applicable to specific types of vectors.

All three of FDA's previous publications discussed the authority to regulate gene therapy products generally, offering virtually no information about its views on any particular facet or application of the technology. The 1996 Addendum was the first to address the requirements pertaining to a specific component of gene therapy, the gene delivery system. Apparently FDA no longer considered it necessary to repeat its claims to regulatory authority, assuming perhaps that any doubts on that point had been resolved by acquiescence and the passage of time. Now the agency could focus exclusively on what sponsors must do to satisfy their legal obligations.

1998 Guidance

In March 1998 FDA issued a guidance document concerning human somatic cell therapy and gene therapy (108). FDA stated that the guidance document was intended to update and replace the 1991 "Points to Consider" document "with new information intended to provide manufacturers with current information regarding regulatory concerns for production, quality control testing, and administration of recombinant vectors for gene therapy; and of preclinical testing for both cellular therapies and vectors" (108, p. 1).

STATUS OF PROTOCOLS

A recent compilation from NIH's registry of gene therapy protocols reveals that, between 1989 and 1999, a total of 280 gene transfer protocols were submitted to NIH and FDA for review (109–111). These include both protocols intended to provide therapeutic benefit and those in which the protocol's intent is not therapeutic (e.g., "marker" gene experiments). Of the 280, 107 were subject to review by the full RAC, approval by the NIH director, and IND approval by FDA — this review process is no longer in effect (109–111). An additional seven protocols were subject to the accelerated RAC review, NIH/ORDA approval, and FDA IND approval — this review process also is no longer in effect (109–111). Ten protocols were subject to FDA IND approval and full RAC discussion — this form of review is currently in effect for some protocols (109–111). The NIH list does not reflect whether protocols approved by NIH were subsequently approved by FDA.

An additional 151 protocols have been screened by NIH/ORDA and referred for exclusive FDA review (109–111). Submission to NIH/ORDA has been required for the purpose of data monitoring and adverse event reporting. As of February 10, 1999, four protocols are awaiting NIH/ORDA's determination of their review requirements (109–111). Finally, one protocol was submitted for sole FDA review and was also voluntarily submitted to NIH/ORDA, presumably by an individual or entity not funded by NIH (109–111).

In early 1997 FDA and NIH approved the first gene therapy trial to be conducted in healthy volunteers. The protocol provides for the injection of healthy individuals with a viral vector in order to measure their immune response (112). The data can be used to establish a baseline that will help evaluate the vector's therapeutic effect in sick people. Use of healthy volunteers is standard in early phase drug development studies (112).

The early sponsors of gene therapy protocols were NIH researchers or investigators at medical research institutions. Between 1989 and 1991, a total of 11 protocols were approved by NIH, all of which were sponsored by researchers at NIH or in academe (109). Beginning in 1992, the caseload began to include more industry sponsors and the pace of submissions rose dramatically. In 1992 and 1993, 44 protocols were approved by NIH, 8 of which were identified as having industry sponsors (109). Between 1994 and 1996, 27 out of 115 protocols received, reviewed, or approved by NIH were sponsored by industry. (During this time period NIH increasingly began referring protocols for sole FDA review, and NIH's publicly available list of protocols does not reflect the fate of the protocols thereafter) (109). Of 110 protocols received by NIH/ORDA between 1997 and 1999, 44 were sponsored by industry (109). It should also be noted that private firms may, and often do, participate in and contribute financially to government or university-sponsored protocols without being listed as sponsors.

As the private sector has assumed a growing role in financial sponsorship of gene therapy research, the types of protocols submitted for review have changed. Although gene therapy was originally conceived as a means of treating or preventing rare diseases caused by "monogenic" defects, namely defects in a single gene — ADA being a prime example — only 36 experimental protocols for such diseases have been approved (110). Most of the industry-sponsored protocols have been aimed at widespread diseases such as cancer and heart disease, or at AIDS. While these diseases may have a genetic component or be susceptible of treatment through gene therapy techniques, they certainly are not the "classical" genetic diseases for which researchers first contemplated gene therapy (51).

As both researchers and FDA become more expert in and comfortable with the techniques of gene therapy, new, and perhaps more difficult questions are emerging. Initially, FDA's concerns focused on the safety of research subjects, such as how would cells be removed from patients? would these cells become contaminated while outside the patients? how would reintroduction of gene-altered cells be achieved (51)? Increasingly, however, FDA's concerns have shifted from "can you" to "should you." NIH recently sponsored its first Gene Therapy Policy Conference to discuss the use of gene therapy for "enhancement," namely for use in non-life-threatening diseases or conditions, such as baldness (113). Public discussion of this issue, FDA's Noguchi emphasizes, is warranted because "the concept of gene therapy is being pushed to lots of different diseases, some of which are more social than physiological" (51).

According to Dr. Noguchi, the RAC is seen as having a crucial role in ensuring that the social and ethical issues raised by particular applications of gene therapy are fully debated and debated in public. Unlike FDA, which must conduct most of its review of gene therapy protocols in private to protect the trade secret information they embody or contain, the RAC is "well constituted for public discussion" (51).

CONCLUSION

The past two decades have seen remarkable advancements in gene therapy research. Given the pace with which experimental protocols are being pursued, it would not be surprising to see the first approved gene therapy product before the official close of the twentieth century. The rapid pace of research has also posed profound challenges for those federal agencies charged with the task of fashioning social controls on scientific technology. As the story

of FDA's regulation of gene therapy reveals, the very first challenge was to determine which agency — NIH or FDA — would take the lead in designing and implementing these social controls. More recently the NIH has formally recognized FDA's statutory authority to review and approve individual gene therapy protocols, and FDA has in turn acknowledged the important role played by the NIH RAC in leading the public discussion on the social and ethical implications of gene therapy. FDA and NIH, working together, thus seem well prepared to confront the challenges that will soon emerge when gene therapy makes the leap from the laboratory to the pharmacy.

ACKNOWLEDGMENTS

The authors are grateful to Marilynn Whitney for her assistance in preparing this article and to Peter Barton Hutt for helpful comments on an early draft.

BIBLIOGRAPHY

1. Federal Food and Drugs Act of 1906, *U.S. Statutes at Large* **34**, 768 (1906).
2. *United States v Sullivan*, 332 US 689, 696 (1948).
3. *U.S. Statutes at Large* **52**, 1040 (1938), codified at *U.S. Code*, vol. 21, secs. 301 et seq.
4. Drug Amendments of 1962, *U.S. Statutes at Large* **76**, 780 (1962).
5. Medical Device Amendments of 1976, *U.S. Statutes at Large* **90**, 539 (1976).
6. Safe Medical Devices Act of 1990, *U.S. Statutes at Large* **104**, 4511 (1990).
7. Food and Drug Administration Modernization Act of 1997, *U.S. Statutes at Large* **111**, 2296 (1997).
8. *U.S. Statutes at Large* **32**, 728 (1902).
9. *U.S. Statutes at Large* **58**, 682 (1944), codified at *U.S. Code*, vol. 42, secs. 201 et seq.
10. *U.S. Statutes at Large* **58**, 702 (1944), as amended, codified at *U.S. Code*, vol. 42, sec. 262(a).
11. Public Health Service and Food and Drug Administration, *Fed. Regist.* **37**, 12865 (1972).
12. National Institutes of Health, *Fed. Regist.* **39**, 39306 (1974).
13. J. Lyon and P. Gorner, *Altered Fates*, Norton, New York, 1996, p. 61.
14. S.G. Michaud, *Business Week*, September 19, 1977, p. 18.
15. National Institutes of Health, *Fed. Regist.* **61**, 35774 (1996).
16. Committee on Recombinant DNA Molecules, *Science* **185**, 303 (1974).
17. National Institutes of Health, *Fed. Regist.* **41**, 27902 (1976).
18. J. McWethy, *U.S. News and World Report*, July 12, 1976.
    J. Goodfield, *Playing God: Genetic Engineering and the Manipulation of Life*, Random House, New York, 1977.
19. J.C. Fletcher, *Hum. Gene Ther.* **1**, 55, 60–61 (1990).
20. W. Froelich, *San Diego Union-Tribune*, January 1, 1985.
21. National Institutes of Health, *Report of the NIH Ad Hoc Committee on the UCLA Report Concerning Certain Research Activities of Dr. Martin J. Cline* (memo. May 21), NIH, Washington, DC, 1981, p. 8.
22. R.M. Cook-Deegan, *Hum. Gene Ther.* **1**, 163–170 (1990).
23. President's Commission for the Study of Ethical Problems in Medicine and Biomedical And Behavioral Research, *Splicing Life: The Social and Ethical Issues of Genetic Engineering with Human Beings*, U.S. Government Printing Office, Washington, DC, 1982.
24. National Institutes of Health, *Fed. Regist.* **50**, 2940 (1985).
25. B.J. Culliton, *Science* **227**, 493 (1985).
26. B.J. Culliton, *Science* **229**, 736 (1985).
27. Food and Drug Administration, *Fed. Regist.* **49**, 50878 (1984).
28. Food and Drug Administration, *Fed. Regist.* **49**, 50880 (1984).
29. B.J. Culliton, *Sci. News*, August 31, 1985.
30. S. Squires, *Washington Post Health*, October 30, 1985.
31. R. Rhein, *Chem. Week*, October 27, 1985.
32. R. Rhein, *Chem. Eng.*, November 11, 1985.
33. National Institutes of Health, *Fed. Regist.* **50**, 33462–33463 (1985).
34. Human gene therapy research should have NIH review, even if funded by another federal agency, RAC decides. *F-D-C Rep. — The Blue Sheet*, February 4, 1987.
35. R. Steinbrook, *Los Angeles Times*, April 13, 1986.
36. S. Squires, Sally, *Washington Post*, May 30, 1989.
37. P. Gorner, *Chicago Tribune*, May 21, 1989.
38. L. Roberts, *Science* **243**, 734 (1989).
39. R.M. Henig, *N.Y. Times Mag.*, March 31 (1991).
40. Genetic Therapy, Inc. goes public as fourth gene therapy collaborative trial with NIH is set to begin. *F-D-C Rep. — The Blue Sheet*, July 3, 1991.
41. Food and Drug Administration, *Fed. Regist.* **56**, 61022 (1991).
42. Food and Drug Administration, *Health News Daily*, December 11, 1992.
43. Food and Drug Administration, *F-D-C Rep. — The Pink Sheet*, December 14, 1992.
44. Food and Drug Administration, *Fed. Regist.* **58**, 53248 (1993).
45. Synthesized gene therapy products will fall under FDA drug review regulations, agency states. *F-D-C Rep. — The Blue Sheet*, October 20, 1993.
46. D.A. Kessler et al., *N. Eng. J. Med.* **329**, 1169–1173 (1993).
47. Cancer vaccine/TIL gene therapy protocols proposed by NCI's Steven Rosenberg okayed by RAC subcmte. *F-D-C Rep. — The Blue Sheet*, July 31, 1991.
48. J. Palca, *Science* **253**, 624 (1991).
49. C. Anderson, *Nature (London)* **357**, 615 (1992).
50. Viagene's Phase I in vivo HIV gene transfer protocol receives unanimous RAC approval June 7. *F-D-C Rep. — The Blue Sheet*, June 9, 1993.
51. Interview by authors with Dr. Philip Noguchi, November 24, 1997.
52. S. Jenks, *J. Nat. Cancer Inst.* **85**, 1544–1546 (1993).
53. FDA and NIH will consolidate gene transfer protocol review by December; HHS' Lee calls for progress report at October AIDS Task Force Meeting. *F-D-C Rep. — The Pink Sheet*, July 25, 1994.
54. E. Marshall, *Science* **265**, 599 (1994).
55. W. French Anderson, *Hum. Gene Ther.* **5**, 1309–1310 (1994).
56. L.E. Post, *Hum. Gene Ther.* **5**, 1311–1312 (1994).
57. S. Sternberg, *BioWorld Today*, September 14, 1994.

58. NIH 'Points-to-consider' retained in new gene therapy NIH/FDA consolidated review proposal from FDA's Noguchi; RAC will stay intact, NIH's varmus assures. *F-D-C Rep. — The Pink Sheet*, September 19, 1994.
59. RAC protocol approval criteria should be assessed by ad hoc review group, NIH Director Varmus recommends, noting complaints about confusing procedures. *F-D-C Rep. — The Pink Sheet*, September 19, 1994.
60. RAC termination called 'Timely and appropriate' in notice submitted by NIH; Sen. Pryor concerned. *F-D-C Rep. — The Blue Sheet*, July 3, 1996.
61. NIH and FDA RAC balks at regulatory reforms. *Cancer Res. Wkly.*, December 19, 1994.
62. National Institutes of Health, *Fed. Regist.* **60**, 20726 (1995).

    NIH committee approves combined FDA/NIH submissions of gene therapy protocols. *Gene Ther. Wkly.*, April 17, 1995.
63. L. Neergaard, *The Ledger (Lakeland, FL)*, May 25, 1995.
64. FDA/NIH gene therapy database future funding may be threatened, FDA's Noguchi warns at RAC meeting. *F-D-C Rep. — The Blue Sheet*, June 14, 1995.
65. So you want to develop gene therapy products? What lies ahead in regulatory hurdles. *Bio Venture View*, July–August 1995.
66. FDA should conduct most reviews of gene therapy protocols, while NIH RAC should continue as policy body — Ad hoc groups; Database development supported. *F-D-C Rep. — The Pink Sheet*, September 4, 1995.
67. H.I. Miller, *Hum. Gene Ther.* **6**, 1361–1362 (1995).
68. Gene therapy should be regulated within traditional biological framework, asserts Fenwick & West's Mackler; Calls current RAC role counterproductive. *F-D-C Rep. — The Pink Sheet*, January 22, 1996.
69. National Institutes of Health, *Fed. Regist.* **61**, 35774 (1996).
70. National Institutes of Health, *Fed. Regist.* **61**, 35775 (1996).
71. National Institutes of Health, *Fed. Regist.* **61**, 35776 (1996).
72. National Institutes of Health, *Fed. Regist.* **62**, 59032 (1997).
73. National Institutes of Health, *Fed. Regist.* **61**, 59726–59727 (1996).
74. Letter from Alan Goldhammer, Ph.D., BIO, to Office of Recombinant DNA Activities, August 2, 1996.
75. Letter from Bonnie J. Goldmann, M.D., Merck Research Laboratories, to ORDA, August 7, 1996.
76. Letter from M. Lynn Pritchard, Ph.D., Glaxo Wellcome Inc., to Dr. Harold Varmus, NIH, August 6, 1996.
77. Letter from Leonard E. Post, Parke-Davis, to Office of Recombinant DNA Activities, July 31, 1996.
78. Letter from Robert C. Moen, M.D., Ph.D., Baxter Healthcare Corporation, to Office of Recombinant DNA Activities, August 5, 1996.
79. Letter from Bridget P. Binko, Cell Genesys, to Office of Recombinant DNA Activities, August 5, 1996.
80. Letter from Alexander E. Kuta, Ph.D., Genzyme, to Office of Recombinant DNA Activities, August 7, 1996.
81. Letter from James G. Timmins, Auragen, to Office of Recombinant DNA Activities, July 29, 1996.
82. Letter from Susan D. Jones, Ph.D., IntraImmune Therapies, Inc., to Office of Recombinant DNA Activities, July 25, 1996.
83. Letter from Henry I. Miller, M.D., Hoover Institution, to Office of Recombinant DNA Activities, August 7, 1996.
84. Comments of Charles R. McCarthy, LeRoy Walters, John C. Fletcher, Foundation on Economic Trends, Council for Responsible Genetics, Alliance of Genetic Support Groups, Friends of the Earth, Genetic Therapy, Inc.
85. Letter from David K. Vanderberg to Dr. Harold Varmus, NIH, June 25, 1996.
86. J.L. Fox, *Nat. Biotechnol.* **15**, 11–12 (1997).
87. National Institutes of Health, *Fed. Regist.* **62**, 59042 (1997).
88. *Code of Federal Regulations*, vol. 21, p. 312.
89. National Institutes of Health, *Fed. Regist.* **62**, 59046 (1997).
90. National Institutes of Health, *Fed. Regist.* **62**, 59034–59035 (1997).
91. *U.S. Code*, vol. 21, secs. 321(g)(B) and (C).
92. *U.S. Code*, vol. 21, sec. 321(p)(1).
93. *Code of Federal Regulations*, vol. 21, sec. 312.20.
94. *Code of Federal Regulations*, vol. 21, secs. 312.23–312.31.
95. *Code of Federal Regulations*, vol. 21, sec. 312.23(a)(iv).
96. *Code of Federal Regulations*, vol. 21, sec. 314.50.
97. National Institutes of Health, *Fed. Regist.* **49**, 50878–50879 (1984).
98. *U.S. Code*, vol. 42, sec. 262(a).
99. *U.S. Code*, vol. 21, sec. 321(h)(3).
100. Food and Drug Administration, *Fed. Regist.* **49**, 50880 (1984).
101. National Institutes of Health, *Fed. Regist.* **50**, 2942–2944 (1985).
102. National Institutes of Health, *Fed. Regist.* **54**, 10956 (1989).
103. CBER, *Points to Consider in Human Somatic Cell Therapy and Gene Therapy*, Center for Biologic Evaluation and Research, 1991.
104. Food and Drug Administration, *Fed. Regist.* **58**, 53249 (1993).
105. Food and Drug Administration. *Fed. Regist.* **58**, 53251 (1993).
106. Food and Drug Administration, *Fed. Regist.* **61**, 5786 (1996).
107. Food and Drug Administration, *Fed. Regist.* (1991).
108. CBER, *Guidance for Industry: FDA Guidance for Human Somatic Cell Therapy and Gene Therapy*, CBER, 1998.
109. Office of Recombinant DNA Activities, *Human Gene Therapy Protocols* (last modified February 10, 1999). Available at: <http://www.nih.gov/od/orda/protocol.pdf>
110. Office of Recombinant DNA Activities, *Protocol List Table*. Available at: <http://www.nih.gov/od/orda/prottab.pdf> The protocol list contains 284 protocols, however, only 280 are included in the accompanying table. According to information received from NIH/ORDA, the table does not include the four protocols because they are incomplete submissions.
111. Telephone call with Gene Rosenthal, May 6, 1999.
112. J.L. Fox, *Nat. Biotechnol.* **15**, 314 (1997).
113. National Institutes of Health, *Fed. Regist.* **62**, 44386 (1997).

See other GENE THERAPY entries; see also HUMAN SUBJECTS RESEARCH, LAW, FDA RULES.

GENE THERAPY, LAW, RECOMBINANT DNA ADVISORY COMMITTEE (RAC).

See RECOMBINANT DNA, POLICY, ASILOMAR CONFERENCE.

GENE THERAPY: OVERVIEW

LeRoy Walters
Kennedy Institute of Ethics, Georgetown University
Washington, District of Columbia

OUTLINE

TERMINOLOGY

The phrase "gene therapy" was used for the first time in a published article in 1970 (1,2). However, functional equivalents to this phrase had been used in academic discussions from the early 1960s on. Among the synonymous phrases were "genetic surgery," "nanosurgery," "euphenics," "genetic engineering," "gene replacements," "directed mutation," "directed genetic change," "designed genetic change," "algeny," "programming cells with genetic messages," and "genetic therapy" (3–14). The phrase "genetic(al) engineering" had been employed at least as early as the Sixth International Congress of Genetics held in 1932 in Ithaca, New York (15), and some authors continued to use the phrase "genetic engineering" even after the advent of molecular biology in the early 1950s.

As the preceding paragraph suggests, there was during the 1960s a striking diversity in the terminology used by commentators on the new techniques that molecular biology might make possible. This diversity is reflected in several publications and reminiscences from that important decade. At a Ciba Foundation symposium held in London on November 26–30, 1962, Hermann J. Muller suggested that it would be preferable to secure gametes from genetically fit individuals rather than to use what he called "nano-needles" to "cause prespecified changes in [those individuals]" (16,17). In opposition to Muller, Joshua Lederberg advocated the use of molecular techniques in the manipulation of germ cells to achieve "the direct control of nucleotide sequences in human chromosomes, coupled with recognition, selection and integration of the desired gene" (18, p. 265). At an April 1963 symposium sponsored by Ohio Wesleyan University, Salvador Luria, Edward Tatum, and Muller favored the phrase "genetic surgery," although Luria also discussed the "removal, addition, and replacement of genes" (19, p. 10), while Tatum at times employed the alternative terminology of "directed gene mutation" and "genetic engineering" (20, p. 22).

As the decade progressed, additional academic disciplines became involved in the discussion of genetic intervention, and more precise descriptions of the actual techniques that might be employed in gene transfer began to appear. At a symposium held at Gustavus Adolphus College in January 1965, theologian Paul Ramsey and biologist Tatum both considered the theme, "Genetics and the Future of Man" (21). Ramsey described one of two possible approaches to genetic control in the following terms:

> The first [method] is some direct attack upon the deleterious mutated gene, either by what is called "genetic surgery," "micro-surgery" or "nano-surgery" [see Ref. 22] or by the introduction of some anti-mutagent chemical that will cause the gene to mutate back or will eliminate it from among the causes of genetic effects (23, pp. 9–10).

In his presentation at the same symposium Tatum described three methods for achieving what he called "*manipulation* [his italics] of genetic change." The second method identified by Tatum sounds like a blueprint for what in the 1990s came to be called ex vivo gene therapy.

> Another potential future approach to directed mutation is via the synthesis in the laboratory of a desired molecule of DNA. This tailored molecule, or any desired DNA molecule if it can be isolated from an organism or cell, can probably be amplified by already known enzymatic replication processes to any needed quantity. This new or isolated gene can then hopefully be introduced into mammalian cells in culture, as in bacterial transformation.
>
> If the rare desired transformed cell can be selected and cultured, the new cells so derived could conceivably be transplanted into a living organism, there to correct a defective function of the original host cells (24, p. 58).

By the late 1960s biologists were describing what would later be called somatic-cell gene therapy in slightly different terms. For example, in a 1968 essay, Lederberg suggested that

> an attempt could then be made to transform liver cells of male offspring of haemophilic ancestry by the introduction of carefully fractionated DNA carrying the normal alleles of the mutant haemophilia gene. This experiment would appear to be entirely analogous to the typical attempts at transforming bacterial forms. However, it is not clear whether one should regard this as a pure example of genetic engineering, since the practical outcome would probably be best achieved by influencing the nuclear constitution of somatic tissues rather than by direct tackling of the germ line (26).

In the same year, 1968, Robert Sinsheimer gave a lecture at the Fordham Chapter of the Society of the Sigma Xi. This lecture, published in the spring of 1969, drew a clear distinction between somatic-cell and germ-line approaches to what he called "designed genetic change" (27). Sinsheimer's specific suggestion for somatic-cell genetic change was that the almost-available technique could be used for the treatment of diabetes.

> If we could obtain a virus analogous to simian virus 40, able to persist within altered cells and, let us say, carrying an expressable gene for proinsulin [the precursor to insulin] in lieu of a normal viral gene, we might indeed be able to provide a genetic alternative to the daily injection of insulin (27, p. 140).

The most widely cited essay during this era was almost certainly Bernard Davis's paper on "Threat and Promise in Genetic Engineering," first given at a December 1969 symposium, then revised and published a year later in *Science* under the title, "Prospects for Genetic Intervention in Man" (28). Davis drew important distinctions between polygenic (including many behavioral) and monogenic traits and between somatic-cell and germ-cell alteration. He also pointed to the technical obstacles standing in the way of even the simplest types of somatic-cell gene alterations. However, Davis also noted a potential advantage of a genetic approach to therapy.

> Such a one-shot cure of a hereditary disease, if possible, would clearly be a major improvement over the current practice of continually supplying a missing gene product, such as insulin (28, p. 1280).

By 1970 there was a consensus among biologist-commentators like Lederberg, Davis, and H. Vasken Aposhian that the phrase "genetic engineering" was unduly alarming to the public (28–30). As noted above, the phrase "gene therapy" began to be used in the published literature during the same year. By the time of two major symposia on this subject in May 1971 the alternative formulations were virtually always "genetic therapy" and "gene therapy" (31,32). Beginning with two articles by Aposhian in 1970 (1,2), one finds an additional three articles that use the phrase "gene therapy" in their titles in 1971 (33–35), and two additional articles plus the proceedings of an NIH symposium in 1972 (36–38). Perhaps the most decisive publication in this series was Theodore Friedmann and Richard Roblin's widely cited *Science* article from March 1972 entitled "Gene Therapy for Human Genetic Disease" (39). (The phrase "gene therapy" was added to the *Bioethics Thesaurus of Bioethicsline* in 1980, to the Library of Congress Subject Headings in 1985, to the National Library of Medicine's Medical Subject Headings in 1989, and to the Dewey Decimal Classification with the 21st edition of 1996.)

It is worth pausing to note that the transition from "genetic engineering," often associated with the evolution of the human species and with voluntary programs of positive eugenics, to the more modest goals of "gene therapy" paralleled a shift in emphasis from classical genetics to microbiology and molecular biology. H.J. Muller and those who sympathized with his views were concerned about an increasing load of mutations in the human gene pool and accented the difficulty of modifying polygenic traits through molecular techniques. Muller himself argued that the basic unit of genetic improvement was the human sperm or egg cell, derived from a willing donor who possessed many desirable characteristics. In contrast, the microbiologists and molecular biologists who wrote on human genetic intervention in the 1960s and early 1970s extrapolated from laboratory research involving bacteria and bacterial and animal viruses. Some commentators, for example, J. Lederberg, were willing to consider targeted genetic changes in both germ-line and somatic cells, but the general trend of the discussion in the 1960s was clearly toward a focus on somatic cells and on the effort to alleviate diseases like hemophilia or diabetes. By the early 1970s the language of "genetic engineering" had been left behind by most biomedical scientists, and germ-line effects were taking second place to the conception of "gene therapy" as a new kind of drug, or biologic, or transplant. This paradigm shift occurred even before the exquisite techniques of recombinant DNA research were widely available.

The consensus in favor of the phrase "gene therapy" (with the word "human" sometimes included as a prefix) remained intact from 1972 through approximately 1998. Journals called *Human Gene Therapy* and *Gene Therapy* were inaugurated to track developments in the field, and the conceptual distinction between somatic-cell gene therapy and more ambitious programs of genetic intervention became firmly established among both academics and members of the general public. However, in a searching article published in 1998, Larry Churchill and his colleagues criticized the use of the phrase "gene therapy" for implying benefit to patients when, at least until that point, little therapeutic success had been achieved in gene therapy trials (40). Churchill and associates argued that the terms "gene therapy" and "gene therapy research" should be deleted from all federal documents describing this new technique and should be replaced by more neutral and modest terminology like "gene transfer research" (40). In their view, this alternative language "more accurately conveys the experimental practice that is currently at issue" (40, p. 45). In the remainder of this article the phrase "gene transfer" will be employed instead of "gene therapy" whenever possible.

MAJOR MODES OF HUMAN GENETIC INTERVENTION

The foregoing discussion has highlighted one of the central distinctions in scientific and ethical discussions of genetic intervention — the distinction between genetic alterations that affect the recipient of the gene transfer alone and alterations that will be passed on to the recipients' descendants through the germ line. Implicit in the early discussions and in the use of the word "therapy" was a second distinction, namely the distinction between the treatment or prevention of disease, on the one hand, and the enhancement of human capabilities or characteristics, on the other. The possible types of enhancements have been classically categorized by Muller and subsequent commentators as physical, intellectual, and moral enhancements (41,42). Thus one can conceive of two-by-two matrix (Table 1) that depicts these two distinctions (43,44).

More recent commentators, especially Matthew Bacchetta and Gerd Richter, have argued that this simple two-dimensional framework is inadequate to deal with another type of potential genetic intervention, namely a deliberate alteration of the mitochondrial DNA that is present in the cytoplasm of mammalian cells (45). In the

Table 1. Gene Therapy Techniques

	Somatic	Germ Line
Prevention, treatment, or cure of disease	1	1
Enhancement of capabilities or characteristics	3	4

Source: Ref. 44, p. xvii.

view of Bacchetta and Richter, the inclusion of genetic interventions to treat or prevent mitochondrial disease yields a three-dimensional matrix, involving the six possibilities shown in Figure 1 (45).

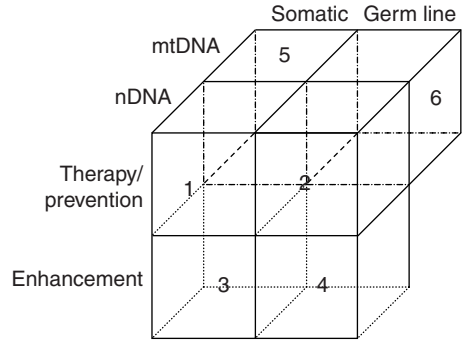

Figure 1. Dimensions of genetic interventions in the human genome: cube 1, somatic gene therapy within nDNA; cube 2, germ-line gene therapy within nDNA; cube 3, enhancement in somatic cells within nDNA; cube 4, enhancement in germ-line cells within nDNA; cube 5, somatic mitochondrial gene therapy; and cube 6, mitochondrial germ-line therapy. *Source*: Ref. 45, p. 456.

EARLY ATTEMPTS AT HUMAN GENE TRANSFER

In 1980 there was clearly an experimental attempt made to perform gene transfer in two subjects who were afflicted with beta-thalassemia. However, three leading accounts of the history of this topic also discuss an event that had occurred a decade earlier, the effort by Stanfield Rogers to infect three German sisters aged 5, 2, and a few months old with the Shope papilloma virus. Rogers, a biochemist and physician at Oak Ridge National Laboratory, hoped that the virus would carry a gene into the sisters' bodies that would counteract the effects of an inherited disease that resulted in toxic levels of an amino acid, arginine, in the children's livers (25,46,47). Rogers later acknowledged that the intervention did not help any of the three girls, although a delay in administration of the virus to the youngest of the three may have rendered the preparation inactive (48–50). The experiment was reported in the *New York Times* on September 21, 1970, by Harold Schmeck (51). At a May 1971 symposium (52) and in a 1972 article published in *Science* (53) Rogers was severely criticized for having undertaken this clinical trial with insufficient preclinical research.

A second early study of gene transfer bore an uncanny resemblance to Rogers's experiment. On May 30, 1979, Martin Cline, a hematologist at the UCLA Medical Center, and three colleagues submitted a clinical protocol to the local institutional review board (IRB) — called the Human Subject Protection Committee (47). What Cline and his colleagues proposed to do was to perform gene transfer experiments in subjects who had hemoglobin disorders like sickle cell anemia or thalassemia. The precise strategy that the researchers planned to follow was to introduce functioning beta-globin genes into bone marrow cells that had been removed from the subject's body. Their hope was to be able to confer a selective advantage on the genetically modified cells so that, when the cells were reintroduced into the body of the subject, they would outgrow the native, deficient bone marrow cells. After modifying the protocol to exclude the use of recombinant DNA and after having waited for more than a year for approval by the local IRB, Cline took matters into his own hands. In the summer of 1980 he flew to Israel and Italy; in each country he performed a gene transfer experiment on a subject who was afflicted with β-thalassemia (47). Once again, the press broke the story about the experiment. On October 8, 1980, Paul Jacobs published a story in the *Los Angeles Times* entitled "Pioneer Genetic Implants Revealed" (54). In response, National Institutes of Health (NIH) Director Donald Fredrickson appointed an ad hoc committee to review Cline's action. After the committee's final report in May 1981, Cline was punished by both UCLA and NIH for having violated both federal regulations that protect human subjects and the NIH "Guidelines for Research with Recombinant DNA Molecules" (47).

DEVELOPMENT OF A REVIEW PROCESS FOR GENE TRANSFER RESEARCH IN THE UNITED STATES

The lively discussion of human genetic intervention that had occurred between 1962 and 1972 came to an almost-abrupt halt in 1972. During the remainder of the 1970s two other topics in genetics and molecular dominated ethical discussion — genetic testing and screening and recombinant DNA research (44). Laboratory research with recombinant DNA, in particular, was brought to the attention of the public by scientists concerned that their work could harm either laboratory workers or the public's health, and the Asilomar meeting held in February 1975 marks a decisive moment in the recombinant DNA debate. Out of the discussion process that led to Asilomar there emerged a public-oversight body for this important field of research, the NIH Recombinant DNA Advisory Committee (RAC). By the late 1970s the concerns of scientists, policy makers, and the public about the potential biohazards had been allayed by additional research and further discussion.

In 1980 public attention began to return to the topic of "genetic engineering," which had so abruptly disappeared from sight in the early 1970s. The resuscitation of the topic began in Europe, where in January of 1980 Mr. B. Elmquist, a Danish member of the Legal Affairs Committee for the Parliamentary Assembly of the Council of Europe, requested that the assembly pass a "recommendation on the protection of humanity against genetic engineering" (44, p. 146). Mr. Elmquist's request initiated a two-year discussion process that

led to a parliamentary hearing in May 1981 and to Recommendation 934 (1982) in January 1982. This recommendation, framed in terms of the postwar human-rights tradition, asserted the right of every human being to "inherit a genetic pattern that has not been artificially changed" (55). On the other side of the Atlantic, a June 1980 letter from three religious leaders to President Carter asserted that "We are moving into a new era of fundamental danger triggered by the rapid growth of genetic engineering" (56, p. 95). In response, an already-existing presidential commission on bioethics decided to conduct a formal study of genetic engineering as applied to human beings. News of Cline's unauthorized experiments, which broke later in the same year, and two widely cited essays on "gene therapy" in the *New England Journal of Medicine* (57,58), also contributed to a renewed awareness of a topic that had been actively debated in the 1960s but almost forgotten during most of the 1970s.

The report of the presidential commission, entitled *Splicing Life*, and a congressional hearing on "Human Genetic Engineering" held in November 1982 led gradually but almost inexorably to the establishment of a public-oversight system for human gene transfer research in the United States (44). The NIH advisory committee that had nurtured recombinant DNA research through its early years, the Recombinant DNA Advisory Committee (RAC), now turned its attention to human gene transfer. Through an interdisciplinary working group that was later renamed the Human Gene Therapy Subcommittee, RAC devised guidelines called the "Points to Consider" for a new arena of human-subjects research. These guidelines were reactive in the sense that Martin Cline had already performed a human gene transfer experiment in 1980. However, they were also proactive: the "Points to Consider" were essentially complete in 1985, yet the first gene-marking proposal was not formally submitted to the RAC until 1988. Two years later the first gene-transfer experiments that aimed to be therapeutic were proposed and publicly debated (46,47).

EARLY YEARS OF PUBLICLY APPROVED GENE TRANSFER RESEARCH

On September 14, 1990, the era of publicly approved gene transfer research with a therapeutic aim began. NIH researchers W. French Anderson, R. Michael Blaese, Kenneth Culver, and their colleagues administered genetically modified T-lymphocytes to a girl named Ashanti DeSilva, who had just turned four. Ashanti was suffering from a rare genetic disease called severe combined immune deficiency, which was caused by a defect in a single gene, the gene that produces an enzyme called adenosine deaminase (ADA). A synthetic enzyme had initially helped strengthen Ashanti's weakened immune system, but the beneficial effects of the enzyme therapy had gradually diminished. The researchers, RAC members, members of the press covering the story, and the general public all hoped that this new mode of somatic-cell genetic intervention would prove to be beneficial (46,47).

During the next four-and-a-half years the field of human gene transfer research grew at a rather steady pace in the United States and commenced in several additional countries. By early 1995 the number of gene therapy protocols approved by RAC stood at 82, and the number of gene-marking studies had reached 25 (59). The number of subjects involved in these 107 early trials was small: Only 597 participants took part in the studies. Of the early gene therapy protocols, all but three were Phase I studies, aimed primarily at determining whether the gene transfer procedure would be toxic to patients. Two studies were categorized as Phase I/II studies, and only one study, approved in March 1995, was a Phase II study — a preliminary evaluation of efficacy (59). The gradual growth of activity in human gene transfer research is indicated in Table 2.

Table 2. Number of Gene Transfer Protocols Submitted by Year, 1990 to 1995

1990	1991	1992	1993	1994	1995
2	9	23	31	33	44

In the early studies of gene transfer in the United States several types of diseases were targeted, and multiple vectors were employed. More than 60 percent of the gene therapy studies (51/82) sought to combat cancers of various kinds. Twenty studies were directed toward monogenic diseases, including cystic fibrosis (11), Gaucher disease (3), and severe combined immune deficiency (1). The remaining 11 protocols sought to combat HIV infection or AIDS (9), peripheral artery disease (1), and rheumatoid arthritis (1). The vast majority of the initial 107 gene transfer studies in the United States employed retroviral vectors (76/107). Adenovirus vectors began to be employed in 1993 and were used in 15 of the initial 107 studies. In 1994 the first adeno-associated viral vector was proposed and approved. Non-viral delivery vehicles included liposomes (12/107), plasmid DNA (2), and particle mediation (1) (59).

GENE-TRANSFER RESEARCH AND OVERSIGHT POLICIES IN OTHER NATIONS

For gene transfer research conducted outside the United States the earliest published information began to appear in the journal *Human Gene Therapy* in December of 1992. In the earliest annual summary of gene transfer research the following three non-U.S. protocols were listed (60):

- Fudan University and Changhai Hospital, Shanghai, China: Hemophilia B
- Centre Leon Berard, Lyon, France: cancer (a gene-marking study)
- San Raffaele Scientific Institute, Milan, Italy: severe combined immune deficiency

By the end of 1993 two additional protocols had been initiated in the Netherlands (61). At the end of 1994, the number of non-U.S. gene transfer protocols had increased to 13. New countries added in the 1994 summary included the United Kingdom, Germany, and Sweden (62). In 1995, or approximately the same time as the U.S. audits

discussed below, the number of non-U.S. protocols stood at 17, with Canada and Poland joining the list as new sites in 1995 (63). At mid-year in 1996, a more comprehensive report compiled by Tony Marcel and J. David Grausz cited the following additional countries in which gene transfer studies either were being planned or had already been undertaken: Switzerland, Egypt, Spain, Australia, Finland, Japan, and Israel (64).

Oversight bodies similar to RAC or FDA regulate the conduct of human gene transfer research in several European countries and in Japan. In the United Kingdom the Gene Therapy Advisory Committee (GTAC) reviews all gene transfer protocols on the basis of guidance that the committee has prepared for researchers (65). In 1996 France passed a special law for cell and gene therapies that requires these biologics to be reviewed by the French Medicines Agency (66). Japan's oversight system for human gene transfer includes two national review bodies, one of which meets publicly and the other of which convenes in private (66).

AUDITS OF 1995 AND THE CRISIS OF 1999 IN THE UNITED STATES

The June 1995 RAC report from which the information in the preceding section was derived reached three primary conclusions about the field of human gene transfer. First, gene-marking studies had advanced the science of bone marrow transplantation by allowing researchers to distinguish between cells that had been removed from a patient's body and purged and cells that had not been removed. Second, there was little evidence of toxicity in the early gene transfer studies. And third, "It is clearly too early . . . to assess the therapeutic efficacy of gene therapy or even to predict its promise" (60, p. 1789).

In December 1995 a committee co-chaired by RAC member Arno Motulsky and Stuart Orkin delivered a more somber verdict to NIH Director Harold Varmus and the NIH Director's Advisory Committee. According to the Orkin-Motulsky committee,

> While the expectations and the promise of gene therapy are great, clinical efficacy has not been definitively established at this time in any gene therapy protocol, despite anecdotal claims of success and the initiation of more than 100 Recombinant DNA Advisory Committee (RAC) approved protocols (67).

The Orkin-Motulsky committee stopped short of recommending a moratorium on clinical trials of human gene transfer, but the committee clearly thought that the balance between preclinical and clinical research should be radically shifted.

In retrospect, one can discern a clear, though gradual downward trajectory in human gene transfer research in the U.S. from the end of 1995 through late 1999. There were several elements in this decline. The early hope for therapeutic success continued to be frustrated, and the 1995 verdicts of the RAC and the Orkin-Motulsky committee remained valid through 1996, 1997, and 1998. During this same period gene transfer research became much less visible in the United States because of changes in the oversight system. The existing system of review and approval or disapproval of all human gene transfer protocols by both RAC and the Food and Drug Administration (FDA) gave way first to a system of more selective review and approval or disapproval by RAC, then to sole regulatory authority by FDA, with informal advice from RAC on protocols that raised novel issues. Researchers were asked to continue submitting reports of serious adverse events both to RAC and FDA, as well as to provide an annual report to the RAC staff on the progress of their research. However, the RAC staff lacked both the number of people and the database capabilities required to tabulate and analyze the reports that were submitted. There was, in addition, increasing evidence that not all of the required information was being provided to the RAC and its staff by researchers and companies. As a result of these multiple factors, no annual reports on progress (or the lack of progress) in human gene transfer research were produced by the RAC and its staff between 1995 and 1999.

The crisis of 1999 for human gene transfer research in the United States began when RAC noted at its September meeting that one researcher and one company had labeled adverse event reports as "proprietary information" and had asked RAC not to discuss the events publicly. Even as RAC responded by insisting that adverse event reports be a matter of public record, a second event occurred that would change the field of gene transfer research for the foreseeable future. On September 17, 1999, Jesse Gelsinger, an eighteen-year-old young man who had ornithine transcarbamylase (OTC) deficiency, died as a direct result of having received gene transfer with an adenoviral vector. The tragedy was compounded by the fact that Mr. Gelsinger's disease had been relatively well controlled through all of 1999 by a combination of drugs and diet (68,69). To their credit, the team of researchers conducting this study at the University of Pennsylvania promptly reported Mr. Gelsinger's death to both NIH and FDA.

The crisis of 1999 continued as FDA placed several gene transfer studies on clinical hold. In addition a vigorous attempt by the RAC staff and FDA to gather all serious adverse events that had occurred in human gene transfer trials, especially those using adenoviral vectors, revealed that less than 10 percent of these events had been reported to RAC and its staff in a timely fashion. At the December 1999 RAC meeting, a working group on adenoviral vectors criticized gene transfer researchers for their lack of standardization in calculating doses of vector and in assuring that the properties of the vector had not changed before it was administered to human subjects. At the same meeting FDA alleged that the University of Pennsylvania research group had committed several violations of FDA regulations in its conduct of the OTC deficiency gene transfer trial. These oral criticisms were followed in January 2000 by FDA's release of "inspectional observations" detailing FDA's charges. The Penn researchers responded to the charges in February, acknowledging some technical errors but denying that those errors were causally related to Mr. Gelsinger's death (70–73).

A February 2000 Senate hearing chaired by Senator William First explored the current oversight system for gene transfer research in the United States. At the hearing Mr. Gelsinger's father, Paul Gelsinger, asserted that he and his son had not been told important information from the preclinical studies that preceded the start of the human gene transfer study for OTC deficiency. More specifically, the deaths of monkeys in preclinical studies employing earlier generations of adenoviral vectors had not been disclosed. Mr. Gelsinger also reported that he and his son had been led to believe that the son's participation in the gene transfer study was likely to be clinically beneficial, even though the study was a Phase I trial. Other witnesses, including the present author, commented on weaknesses in the current oversight system for gene transfer research in the United States and on steps that are being undertaken by NIH and FDA in an effort to remedy those weaknesses.

PROSPECTS FOR THE FUTURE

As of late 1999, approximately 320 U.S. gene therapy (as distinct from gene-marking) trials had been registered with NIH and RAC. For studies being conducted in other countries, the most recent figures available suggest that at least 36 gene transfer trials have been initiated or completed in 11 countries (74). This latter number is almost surely a gross underestimate. Many companies seem to prefer reporting gene transfer trials only on a confidential basis to the regulatory agencies that oversee those trials. The earlier generalizations about the lack of demonstrated clinical efficacy, at least in published articles, continue to hold, although there are rumors of initial success with severe combined immune deficiency in several children and in hemophilia with adults. The 1995 conclusion about the lack of toxicity in gene transfer trials has now had to be reversed, in the light of Mr. Gelsinger's death and several other reported laboratory and clinical toxicities.

It is fair to say that in late 1999 and early 2000 the field of human gene transfer research was undergoing an agonizing reappraisal. Success in this use of these techniques may lie ahead, but it will not be as easy as all hoped and many believed in 1990, when the first approved study was initiated. The precise role for gene transfer, cell transfer (often called "cell therapy"), organ and tissue transplantation (including xenotransplantation), biologics, and drugs in the armamentarium of the future remains to be clarified. The future success of gene transfer in the treatment of disease cannot be guaranteed. What can definitely be achieved, however, is the creation of a transparent, accountable oversight system that assures that the human subjects who make this research possible will be dealt with honestly and with the highest measure of respect.

ACKNOWLEDGMENTS

I wish to thank my colleague, Linda A. Powell, for helpful editorial assistance. I am indebted to my colleague, Martina Darragh, for information on the first recorded usage of the phrase "gene therapy."

BIBLIOGRAPHY

1. H.V. Aposhian, *Perspect. Biol. Med.* **14**, 98–108 (1970).
2. J.V. Osterman, A. Waddell, and H.V. Aposhian, *Proc. Nat. Acad. Sci.* **67**, 37–40 (1970).
3. H.J. Muller and J. Lederberg, in G. Wolstenholme, ed., *Man and His Future*, J. & A. Churchill, London, 1963, pp. 255, 265.
4. S.E. Luria, E.L. Tatum, and H.J. Muller, in T.M. Sonnenborn, ed., *The Control of Human Heredity and Evolution*, Macmillan, New York, 1965, pp. 1, 9–11, 17, 29, 40, 42–43, 101, 104, 108–109, 119, 121, 123, 125–127.
5. E.L. Tatum and P. Ramsey, in J.D. Roslansky, ed., *Genetics and the Future of Man*, Appleton-Century-Crofts, New York, 1966, pp. 53–54, 57–58, 118–119, 162–153, 155.
6. R.D. Hotchkiss, *J. Hered.* **56**, 197–202 (1965).
7. J. Lederberg, *Bull. Atomic Scientists* **22**, 4–11 (1966).
8. E.L. Tatum, *Perspect. Biol. Med.* **10**, 19–32 (1966).
9. M.W. Nirenberg, *Science* **157**, 633 (1967).
10. J. Lederberg, *Science* **158**, 313 (1967).
11. R.L. Sinsheimer, *Am. Sci.* **57**, 134–142 (1969).
12. B.D. Davis, in P.N. Williams, ed., *Ethical Issues in Biology and Medicine*, Schenkman Publishing, Cambridge, MA, 1973, pp. 17–32.
13. B.D. Davis, *Science* **170**, 1279–1283 (1970).
14. J.A. Wolff and J. Lederberg, *Hum. Gene Ther.* **5**, 469–480 (1994).
15. H.D. Goodale, in D.F. Jones, ed., *Proc. Sixth International Congress of Genetics*, vol. 2, Brooklyn Botanic Gardens, Ithaca, NY, 1932, pp. 65–66.
16. H.J. Muller, in G. Wolstenholme, ed., *Man and His Future*, J. & A. Churchill, London, 1963, p. 255.
17. H.J. Muller, *Perspect. Biol. Med.* **3**, 1–43 (1959).
18. J. Lederberg, in G. Wolstenholme, ed., *Man and His Future*, J. & A. Churchill, London, 1963, p. 265.
19. S.E. Luria, in T.M. Sonnenborn, ed., *The Control of Human Heredity and Evolution*, Macmillan, New York, 1965, p. 10.
20. E.L. Tatum, in T.M. Sonnenborn, ed., *The Control of Human Heredity and Evolutions*, Macmillan, New York, 1965, p. 22.
21. J.D. Roslansky, ed., *Genetics and the Future of Man*, Appleton-Century-Crofts, New York, 1966.
22. H.J. Muller, in T.M. Sonnenborn, ed., *The Control of Human Heredity and Evolution*, Macmillan, New York, 1965.
23. P. Ramsey, in J.D. Roslansky, ed., *Genetics and the Future of Man*, Appleton-Century-Crofts, New York, 1966, p. 118.
    Reprinted in P. Ramsey, *Fabricated Man: The Ethics of Genetic Control*, Yale University Press, New Haven, CT, 1970, pp. 9–10.
24. E.L. Tatum, in J.D. Roslansky, ed., *Genetics and the Future of Man*, Appleton-Century-Crofts, New York, p. 58.
25. J.A. Wolff and J. Lederberg, *Hum. Gene Ther.* **5**, 471 (1994).
26. J. Lederberg, in *Proc. Sixth World Congress on Fertility and Sterility*, May 20–27, 1968, Israel Academy of Sciences and Humanities, Jerusalem, 1970, pp. 18–23.
27. R.L. Sinsheimer, *Am. Scientist* **57**, 134–142 (1969).
28. B.D. Davis, *Science* **170**, 1279–1283 (1970).
29. J. Lederberg, *N.Y. Times*, September 26, 28 (1970).
30. H.V. Aposhian, *Perspect. Biol. Med.* **14**, 106–107 (1970).

31. E. Freese, ed., *The Prospects of Gene Therapy*, Fogarty International Center, National Institutes of Health, Bethesda, MD, 1972.
32. M.P. Hamilton, ed., *The New Genetics and the Future of Man*, Eerdmans, Grand Rapids, MI, 1972, pp. 107–175.
33. S. Rogers, *Res. Commun. Chem. Pathol. Pharmacol.* **2**, 587–600 (1971).
34. J.V. Osterman et al., *An. N.Y. Acad. Sci.* **179**, 514–519 (1971).
35. P.K. Qasba et al., *Proc. Nat. Acad. Sci.* **68**, 2345–2349 (1971).
36. T. Friedmann and R. Roblin, *Science* **175**, 949–955 (1972).
37. E. Freese, *Science* **175**, 1024–1025 (1972).
38. H.V. Aposhian et al., *Fed. Proc.* **31**, 1310–1314 (1972).
39. T. Friedmann and R. Roblin, *Science* **175**, 949–955 (1972).
40. L.R. Churchill et al., *J. Law, Med. Ethics* **26**, 38–47 (1998).
41. *Nature* **144**, 521–522 (1939).
42. H.J. Muller, *Perspect. Biol. Med.* **3**, 22, 37–38 (1959).
43. L. Walters, *Nature* **320**, 225–227 (1986).
44. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997.
45. M.D. Bacchetta and G. Richter, *Camb. Q. Healthcare Ethics* **5**, 450–457 (1996).
46. J. Lyon and P. Gorner, *Altered Fates: Gene Therapy and the Retooling of Human Life*, W.W. Norton, New York, 1995.
47. L. Thompson, *Correcting the Code: Inventing the Genetic Cure for the Human Body*, Simon & Schuster, New York, 1994.
48. S. Rogers, in M. Lappé and R.S. Morrison, eds., *An. N.Y. Acad. Sci.* **265**, 66–71 (1976).
49. S. Rogers, in S. de Grouchy, ed., *Proc. 4th Int. Congr. Hum. Genet.* (Paris, September 6–11, 1971), Excerpta Medica, Amsterdam, 1971, pp. 36–40.
50. H.G. Terheggen et al., *Z. Kinderheilkunde* **119**, 1–3 (1975).
51. H.M. Schmeck, *N.Y. Times*, September 21, 28 (1970).
52. P. Ramsey, in M.P. Hamilton, ed., *The New Genetics and the Future of Man*, Eerdmans, Grand Rapids, MI, 1972, pp. 162–163.
53. T. Friedmann and R. Roblin, *Science* **175**, 949–955 (1972).
54. *L.A. Times*, October 8, 1, 26 (1980).
55. Council of Europe, Parliamentary Assembly, *Texts Adopted by the Assembly*, 33rd Ordinary Session, Third Part, January 25–29 the Council, Strasbourg (1982).
56. United States, President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life*, U.S. Government Printing Office, Washington, DC, November 1982.
57. W.F. Anderson and J.C. Fletcher, *N. Engl. J. Med.* **303**, 1293–1297 (1980).
58. K.E. Mercola and M.J. Cline, *N. Engl. J. Med.* **303**, 1297–1300 (1980).
59. G. Ross et al., *Hum. Gene Ther.* **7**, 1781–1790 (1996).
60. *Hum. Gene Ther.* **3**, 729 (1992).
61. *Hum. Gene Ther.* **4**, 856 (1993).
62. *Hum. Gene Ther.* **5**, 1550–1551 (1994).
63. *Hum. Gene Ther.* **6**, 1676–1678 (1995).
64. T. Marcel and J.D. Grausz, *Hum. Gene Ther.* **7**, 2025–2046 (1996).
65. Gene Therapy Advisory Committee, U.K. Department of Health, *Gene Ther.* **1**, 1–20 (1994).
66. O. Cohen-Haguenauer, *Current Opinion Biotechnol.* **8**, 361–369 (1997).
67. Stuart H. Orkin and Arno G. Motulsky, Co-Chairs, "Report and Recommendations of the Panel to Assess the NIH Investment in Research on Gene Therapy," December 7, 1995. Unpublished report available from the Office of the Director, National Institutes of Health, Bethesda, MD.
68. R. Weiss and D. Nelson, *Washington Post*, September 29, A1, A21 (1999).
69. N. Wade, *N.Y. Times*, September 29, A20 (1999).
70. R. Weiss and D. Nelson, *Washington Post*, January 22, A1, A12 (2000).
71. S.G. Stolberg, *N.Y. Times*, January 22, A1, A10 (2000).
72. R. Weiss and D. Nelson, *Washington Post*, February 15, A3 (2000).
73. S.G. Stolberg, *N.Y. Times*, February 15, A16 (2000).
74. *Hum. Gene Ther.* **10**, 3119–3123 (1999).

See other GENE THERAPY entries.

GENETIC COUNSELING

ANNE L. MATTHEWS
Case Western Reserve University
Cleveland, Ohio

OUTLINE

INTRODUCTION

Genetic counseling is relatively new as a recognized health profession. The need for genetic counseling services as a discipline unto itself arose out of technological advances in human genetics and their applications to clinical medicine and health care. The profession's continued growth and maturity is a reflection of the continuing explosion of new knowledge in the discipline of human and medical genetics and of its impact on health care.

History

The concept of genetic counseling can be considered eons old. Genetic advice has been provided to families by their relatives, friends, and neighbors since humans began communicating with one another. The Bible, other religious writings, and civil laws have provided families for centuries with rules and laws regulating reproductive behavior based on heredity. However, the term "genetic counseling" as currently utilized is attributed to Sheldon Reed, a Ph.D. geneticist, who published his philosophy of genetic counseling in his text entitled *Counseling In Medical Genetics* (1,2). Reed defined the concept of genetic counseling as "a kind of social work done for the benefit of the whole family entirely without eugenic connotations" (3, p. 335). He stated that the primary function of counseling was to provide people with an understanding of the genetic problems they had in their families. For Reed, in order for one to provide genetic counseling the individual had to have some knowledge of human genetics, have a "deep respect for the sensitivities, attitudes, and reactions of the client," and have a "desire to teach, and to teach the truth to the full extent that was known" (1, pp. 11–12).

Providers of genetic information during the first 40 years of the twentieth century were primarily doctorally trained geneticists in academic institutions. Their major interests focused on laboratory approaches and population genetics to answer questions regarding evolution and how to decrease the presence of disease genes in the gene pool. When consulted, genetics experts provided information dealing with the genetic contribution of heredity and the prevention of genetic diseases and birth defects. Their counseling strategy was one of advice and recommendations. This fit well with the basic tenets of eugenics that were so prevalent in academic and scientific circles during this time period (4). It was not until the later part of the 1940s and 1950s that medicine began to demonstrate an interest in human genetics. Clinics dealing with genetic questions began to appear, two of the first were the Hereditary Clinic at the University of Michigan and the Dight Institute of Human Genetics at the University of Minnesota (5). However, much of this information continued to be presented with the overall intent of decreasing the incidence of genetic disease and birth defects (6).

By the 1960s technological advances in a number of areas — including cytogenetics, biochemical, and molecular genetics, population screening of newborns for phenylketonuira, and the technological advances making prenatal diagnosis a reality (7) — had a tremendous impact on physicians' interests in the field of human genetics (5). Medical genetics was becoming a recognized area of practice as physicians began to replace the laboratory-based geneticist in providing genetic information to patients and families. And while the emphasis of counseling remained preventative in nature, by assisting families in a process of making informed, rational decisions regarding reproductive choices, the need to respond to the psychosocial aspects of genetic disorders was beginning to emerge as an appropriate goal of genetic counseling (8).

In the 1970s the movement in medicine and health care as a whole shifted toward greater patient autonomy (9). In the genetic counseling arena, effectiveness of counseling was no longer solely based on whether or not reproductive plans were altered or information and recurrence risks recollected. It was also reflected by the importance of patients and families making reproductive decisions based on personal values and family circumstances. This philosophical shift, from prevention of genetic disease to concern for the patient's or counselee's total well-being, brought medical geneticists closer to Reed's definition of genetic counseling. Moreover this shift was coincident with the initiation of the first master's level genetic counseling program, a team approach to providing genetic services, and the general public's request for access to such services. This philosophy was also legitimized when the Ad hoc Committee on Genetic Counseling of the American Society of Human Genetics published their definition of genetic counseling that acknowledged the importance of the psychological dimensions of the counseling process and the role of master's level trained personnel who were advocates of this approach (10). The definition that was adopted by the American Society of Human Genetics in 1975 continues to be accepted as the definitive definition of genetic counseling:

> Genetic counseling is a communication process which deals with the human problems associated with the occurrence, or the risk of occurrence, of a genetic disorder in a family. This process involves an attempt by one or more appropriately trained persons to help the individual or family (*1*) comprehend the medical facts, including the diagnosis, the probable course of the disorder, and the available management; (*2*) appreciate the way heredity contributes to the disorder, and the risk of recurrence in specified relatives; (*3*) understand the options for dealing with the risk of recurrence; (*4*) choose the course of action which seems appropriate to them in view of their risk and the family goals and act in accordance with that decision; and (*5*) make the best possible adjustment to the disorder in an affected family member and/or to the risk of recurrence of that disorder (10, p. 240).

THE GENETIC COUNSELING PROCESS

To understand the role of the genetics team, and more specifically genetic counselors, it is helpful to understand what the process of genetic counseling encompasses. In simple terms, genetic counseling can be thought of as an information exchange provided in a team approach milieu (11). Based on information gleamed from the patient or counselee (terms used interchangeably) and working in concert with team members, the genetic counselor provides information and psychological support regarding specific concerns or questions of the family. The process often begins with the identification of individuals within the family affected with birth defects or genetic disorders or with the identification of those who may be at increased risk for a variety of inherited conditions. The process consists of a number of different steps, often posed as questions by families. The first such question is usually: "What is it?" or "Am I at increased risk because of my genetic make-up?" The next step in the

process is to attempt to answer the questions of "What caused it?" and "What can be done about it?" Ideally this aspect of the process involves establishing an accurate medical diagnosis, which in turn forms the basis for genetic counseling. While the genetic evaluation consists of many of the same components found in any medical evaluation, the emphasis is often quite different. In particular, medical information regarding extended family members is often a prerequisite to providing accurate information and genetic counseling. After the establishment of a specific diagnosis, a discussion of the prognosis, treatment, and management options based on the most current information available can ensue. Counseling next addresses the question of "Will it happen again?" Recurrence risks can be discussed and families provided with a variety of options regarding future reproductive choices.

Equally important, the genetic counseling process assists individuals and families in coping with the emotional burdens and adjustments required where the person or a family member is at risk for or affected with a genetic condition.

Data Gathering

In a genetics evaluation the information obtained from families provides important and often defining data needed to reach a specific diagnosis. Major categories of information elicited include prenatal, perinatal, medical, developmental, family, and psychosocial histories. The prenatal and perinatal histories provide an overview of fetal and newborn well-being. Documentation of such information as a potential teratogenic exposure, maternal disease and acute illness, or fetal growth and behavior, such as reflected in fetal movement, may provide important clues to identifying a specific diagnosis and etiology. These specific histories are often helpful in differentiating between prenatal etiologies of abnormalities and those resulting from birth injury (12).

Medical and developmental histories provide the genetic counselor with information regarding the natural history of the disorder in the affected individual. They also provide important information regarding variability, namely how a specific etiology affects what is seen in a specific patient (the phenotype). The medical and developmental histories can also provide the genetics team with direction in delineating a list of differential diagnoses that could help establish a specific diagnosis.

Obtaining a reliable family history can be extremely helpful in clarifying an etiology, diagnosis, and/or risk of recurrence in a family. The genetic counselor constructs a detailed and extensive pedigree that outlines in pictorial format three to four generations in a family. The genetic counselor obtains specific information about family members based on his/her knowledge of the variability of birth defects and the often multiple expressions of a genetic disorder. Families may not always be aware of how variable the expression of a genetic disorder can be, and thus may not be aware that other family members are also affected. For example, a child with a cleft palate, a heart defect, and learning disabilities is diagnosed with velocardiofacial syndrome (a genetic disorder that can have numerous effects including abnormalities of the palate, heart, and other organ systems as well as a characteristic facial appearance) (13). By obtaining a detailed family history and asking questions that would guide the family to describe specific findings, two other family members are identified (and the diagnosis is confirmed on physical examination and/or genetic testing as also having features of velocardiofacial syndrome). Based on this information, the genetic counselor can provide the family with specific information regarding this diagnosis including prognosis, treatment and management options, and specific recurrence risks to future offspring or other family members.

The physical examination part of the genetic evaluation differs somewhat from a routine medical examination. Performed by the physician geneticist, it is aimed not only at detecting major and minor malformations (dysmorphic features) but also at describing a pattern or constellation of findings that may provide clues in determining a diagnosis (14). For example, a congenital heart defect may be an isolated finding. However, if other features are noted such as poor muscle tone, upslanting eyes, and single palmer creases (simian creases), one might suspect a diagnosis of Down syndrome. Moreover such detailed descriptions are helpful in determining whether or not a physical finding is of significance or represents a normal variation or familial finding. Laboratory data collection is another important aspect of the data gathering. If a genetic disorder is suspected, diagnostic studies may be useful in delineating a specific diagnosis. Such laboratory tests as chromosome analysis, molecular DNA or biochemical studies, radiographs, or organ imaging may be appropriate. Additionally the patient may need to be referred to other specialties such as neurology or ophthalmology or for developmental studies.

Lastly, families may be asked for permission to take photographs. In many cases the adage "A picture is worth a thousand words" is well deserved, as photographs provide accurate descriptions of what has been noted on the physical examination. This also allows the genetic counselor to share the patient's findings more accurately with other clinicians in order to get assistance in reaching a specific diagnosis.

Counseling

Once all of the information has been gathered, and the genetics team has had the opportunity to consider possible diagnoses and review pertinent literature and databases, discussion with the client and/or family can begin. If a diagnosis has been identified, the geneticist or genetic counselor discusses the etiology and its genetic basis (what caused it), the medical and developmental implications of the diagnosis (what is it); the prognosis, treatment, and/or management options (what can be done about it); as well as the recurrence risks and availability of prenatal testing (will it happen again). The counselor also identifies and discusses the psychosocial impact of the disorder on the affected individual as well as on other family members, what needs the family may have, and what resources and support groups may be available to help the family or patient to take the next step forward in the process of learning and coping with new information.

If more testing is needed before a specific diagnosis can be made, the team provides the family with their impressions and/or an explanation of the diagnoses being considered. The genetic counselor discusses the specific testing or other specialty evaluations that the genetics team has recommended to confirm or rule out a diagnosis(es) and provide the family with information regarding costs, referrals, insurance issues, and possible date(s) when results may be expected. The team then discusses any interventions that may be useful such as a referral for physical and occupational therapy. And finally, the counselor sets up a plan with the family regarding how best to relay results of recommended tests as well as how best to provide follow-up information to the family and their health care providers.

Finally, if no diagnosis can be made, the ensuing discussion with the family includes the level of suspicion that the condition is genetic and the possibility of further evaluation and new testing in the future. The counselor provides the family with a range of recurrence risks depending on the possible genetic etiologies for similar findings, and with options for monitoring future pregnancies, if appropriate. Discussion also includes any available options regarding management or therapy for symptoms. When no specific diagnosis can be reached, a frustrating experience for both counselor and counselee, the emphasis shifts to helping the family address the impact of the lack of a diagnosis with supportive follow-up counseling for the family as needed.

Follow-up

Perhaps one of the most important, although often unacknowledged, aspects of the genetic counseling process is that of follow-up. As mentioned above, a major aspect of the genetic counseling process is to help the patient and/or family incorporate the information component of the session with the psychosocial, emotional impact of that information. Kessler notes that the "psychological responses of counselees are not only a normal, but often a necessary step in comprehending, integrating and coping with a medical diagnosis or the content material of genetic counseling" (15, p. 19). Following any counseling session, a written summary of the information provided and recommendations made is sent to the family. This written summary provides the family with an opportunity to review and refresh their memories of all the information that they received, as well as allowing them to share the information with other family members if they so desire. Specific recommendations can be reviewed at a later date, or the importance of additional counseling can be reemphasized if reproductive plans or circumstances change. The summary may also identify any misinformation or misconceptions that need to be addressed (9).

The genetic counselor will often follow up with phone contact to answer remaining questions or clarify issues for the family. If further testing or evaluations are pending, the counselor will arrange for the family to return to the clinic to discuss any additional information that has been obtained. These contacts also allow the counselor some ongoing assessment of how the family is assimilating information and coping and adjusting to the condition. In many situations the counselor is able to provide the family with anticipatory guidance and support regarding a number of different issues. However, when situations arise that are beyond their therapeutic expertise, the counselor will refer the client or family to the appropriate health care provider. The counselor may also help families to identify support groups or other resources, such as financial or educational resources, that would be appropriate for their particular situation. In many cases the genetic counselor may have an ongoing relationship with the client and family as new information or family circumstances arise.

GENETIC COUNSELING SETTINGS

Genetic counseling services are provided in a number of different settings. These include prenatal diagnostic clinics, diagnostic or general genetics clinics, and specialty and management clinics. Genetic services may also be available as part of a consultative service in some hospitals. In each setting, while specific aspects may vary and different components of the genetic counseling process may be emphasized, the overall approach and goals of the genetic counseling encounter remain constant.

Prenatal Setting

In the prenatal setting, pregnant women are usually referred for genetic counseling services based on one of two scenarios: the ideal situation when counseling occurs prior to prenatal screening or procedures, and the more common scenario when an abnormality is noted on ultrasound or maternal serum screening results change a woman's risk for abnormalities. In either scenario, genetic counseling can assist pregnant couples by providing a forum for discussion and detailed information on which informed decisions about further procedures can be made.

In the ideal scenario, one in which a referral is made prior to screening or testing because of a known risk factor (i.e., pregnant woman over the age of 35 or a positive history of a genetic disorder), there is usually a two-step process involving a counseling session followed by the actual prenatal testing procedure. Patients meet with a genetic counselor who obtains information regarding their indication for referral, their pregnancy, and their extended family history in the form of a pedigree. The counselor reviews the couple's risk for having a fetus affected with a birth defect or genetic disorder based on all of this information. At this juncture the couple may decide to have further diagnostic testing initiated or may decide not to pursue further evaluation. It is incumbent on the genetic counselor to provide the appropriate support in this decision-making process in a nondirective (i.e., noncoercive) manner (16) so that couples feel comfortable with the process and counseling does not become a mandate for subsequent procedures. Walker notes that "there is still a widespread perception that the sole purpose of prenatal diagnosis is to identify anomalous fetuses so that their birth can be prevented by pregnancy termination" (9, p. 604). There are a number of good reasons beyond termination why prenatal testing may be

appropriate for couples, including reassurance about the well-being of the fetus, anticipatory management based on a fetal diagnosis such as where or how the delivery should occur (in a tertiary hospital setting vs. local facility), the opportunity to plan for treatment in the neonatal setting, and, equally important to many patients, the opportunity to adjust, plan ahead, and gather support and resources prior to the baby's birth (17). If the patient elects to have one of the diagnostic procedures, such as chorionic villus sampling, amniocentesis, or chordocentesis, the procedure is reviewed and discussed in detail. Associated risks of the procedure and follow-up are also described. After the test results are received, which are usually available within 10 days to 3 weeks depending on what studies are being conducted, the referring physician is notified. If the results are abnormal, the patient often returns to the genetic counselor to discuss the test results, what the meaning of such results may have for the present pregnancy, and what options may be available to the family such as termination or continuation of pregnancy and management at delivery. Whatever the decision, the counselor continually acknowledges the impact of such findings and subsequent actions on the patient and the family. Continued acknowledgment and validation regarding couples' decisions should remain a primary concern for the counselor (17). Even if the pregnancy is terminated after a diagnosis of a disorder in the fetus, the genetic counselor provides continued support and referral resources.

The more common prenatal diagnosis scenario occurs when the patient or couple is referred for genetic counseling following ultrasound of the pregnancy where an abnormality is noted or the maternal serum screening test is abnormal. In this all too common situation, the genetic counselor must provide information, counseling, and support in a milieu of anxiety and fear to a couple or patient that did not previously know of any increased risk and is now facing unanticipated decisions (9). The counselor will need to focus the discussion on what is known and not known and what other procedures may be helpful in better delineating the abnormality or problem. This process is conducted within a time-limited framework, where decisions regarding possible options available to the couple must be made quickly amid confusion, stress, and high anxiety.

General Genetics Setting

The general genetics clinic, once thought of as the realm of pediatrics, now often includes a growing number of adult patients as more knowledge is gained regarding late-onset genetic disorders or as affected individuals live longer with excellent medical care. In general, the initial visit to the genetics clinic is preceded by some form of intake process, during which the genetic counselor or clinic staff contacts the family to briefly discuss their concerns, gain some preliminary information, and provide the family with some overview of what they can expect during the visit. Additionally medical records and past testing results are requested prior to the visit.

At the appointment, families are again asked about their understanding of why they were referred, what their concerns are, or what they hope to gain from the visit. Previously obtained information is confirmed and additional prenatal, perinatal, developmental, medical, and psychosocial histories are elicited as needed. A detailed four-generation pedigree is also constructed. If appropriate, the medical geneticist then performs a physical examination, paying particular attention to minor findings and variations. At this point, testing and other evaluations may be ordered or recommended to confirm or rule out a suspected diagnosis. If a specific diagnosis is made, the genetics team provides information regarding the diagnosis, prognosis, treatment, and management, risk of recurrence, and available reproductive options or testing. Available resources and referrals to support groups may be made. Other family members at risk for developing the disorder or for having an affected offspring are identified. It may also be suggested that other family members may wish to seek genetic counseling to discuss the same or similar issues. In most clinic settings, a written summary of the counseling session is then sent to the family and health care providers for their records.

Sometimes, despite the best of efforts and for a number of different reasons, no diagnosis can be established or test results are inconclusive. If the patient is a child, a family may be asked to return to the clinic at some point in the future as, not rarely, a child will "grow into" a diagnosis (12). This also allows the genetics team to observe growth and development over time, since they may provide clues to a diagnosis. In any case, the only options available to the genetic counselor may be reviewing the information that is available regarding possible categories of diagnosis, recurrence risks, and future options. This unfortunate scenario is as frustrating for the genetic counselor as it is for the families. Nevertheless, the genetic counselor will help the family address the issue and impact of the lack of a diagnosis and provide support and follow-up for the family as needed.

Specialty and Management Clinic Settings

Specialty and management clinics, in which the medical geneticist and genetic counselor are integral members of a multidisciplinary team, are designed to assist families and affected individuals deal with a myriad of problems associated with a specific disorder or constellation of findings that can arise over a lifetime. Such specialty or management clinics may deal with a specific genetic disorder such as hemophilia or Marfan syndrome, or with a grouping of disorders or diagnoses such as craniofacial clinic or neuromuscular clinic. The clinic team works closely with the families and their primary care professionals to provide specialized health care services and treatments as well as address such issues as financial concerns and special needs such as equipment. The genetic counselor can provide information and counseling to parents of affected children regarding the genetics of the diagnosis as well as discuss recurrence risks and future reproductive options. These same issues can be discussed with affected adults and their spouses and can continue into the next generation with the children born to these families.

Two of the fastest growing areas of specialty clinics are the hereditary cancer clinics and those that deal with late-onset disorders, such as Huntington disease or Alzheimer's disease. These clinics have seen tremendous growth in the last few years due to the technological advances made in human genetics and the discovery of specific genes responsible for some inherited cancers and certain neurological disorders. In both cancer and late-onset disorders, families are actively seeking information regarding susceptibility or presymptomatic testing (18). In the case of cancer, genetic testing for specific inherited cancers has sparked great interest from individuals and families although only 15 percent of cancer cases are known to have an inherited susceptibility gene (19). For many patients, test results may influence medical management. In late-onset disorders such as Huntington disease, presymptomatic testing (i.e., the ability to identify the gene causing the disorder years before any symptoms appear) has led people to seek testing in order to provide themselves with information on which to make choices about career paths, lifestyle changes, or reproductive decisions. These genetic counseling sessions may be extremely complex as they need to cover a wide range of concerns, not only about the disorder in question, but about the impact of testing and test results on the patient and family, including such issues as confidentiality, discrimination, and possible loss of life or health insurance (20).

In summary, genetic counseling is a complex and multifaceted process that draws on a field of professional expertise that involves diagnosis, provision of information, and counseling of individuals about their genetic makeup and their chances of being affected with a genetic disorder or having offspring with birth defects or genetic disorders. It is usually best accomplished as a team approach by genetics professionals including genetic counselors, medical geneticists, Ph.D. trained geneticists, and other specialized health care professionals such as genetic nurse specialists and social workers.

GENETIC COUNSELORS

Having defined the specific roles a genetic counselor carries out, we turn our attention to the process of acquisition and official recognition of these skills, given the genetic counselor's unique relationship between the counselee and their family. Genetic counselors are health professionals with specialized graduate training and experience in the areas of medical genetics and counseling (21). They work as members of the health care team, providing information and support to families with or at risk for genetic disorders and birth defects. In addition to providing expertise regarding genetic disorders, genetic counselors also provide supportive counseling, serve as patient advocates, and act as educational and resource professionals for their patients, other health care providers, and the general public. Many counselors are also actively engaged in research activities related to the field of medical genetics and genetic counseling. Bartells and colleagues state that "genetic counselors work at the intersection between the information produced by scientists and the hopes, dreams, and fears of clients whose lives could dramatically change as a result of receiving that information" (22, p. ix).

Genetic Counseling Training Programs

Sarah Lawrence College developed the first master's level genetic counseling program in 1969. Program developers, drawing on the expertise of several well-known geneticists, developed a genetic counseling program that balanced theoretical coursework in human and medical genetics with counseling theory and clinical experience in genetics centers providing genetic counseling services (23). This program based its counseling curriculum on a client-centered approach in which students learned interviewing skills, the concept of nondirective counseling, and "employing empathic responses against the background of unconditional positive regard" (23, p. 20).

Curriculum

In continuing efforts to define the role of the genetic counselor as a major provider of genetic services and to standardize the minimum educational requirements for a genetic counselor, directors of genetic counseling programs and other genetics professionals met three times between the years of 1974 and 1979 to address standards of genetic counseling education. The conclusions and recommendations were outlined by Dumars and colleagues in 1979 in *Genetic Associates: Their Training, Role and Function* (24). It was the consensus of the program directors that genetic counseling programs needed to maintain a dialogue among themselves and with other genetics professionals to assure the quality and effectiveness of the professionals being trained (25). Currently, there are 26 genetic counseling programs recognized by the American Board of Genetic Counseling, three international programs, and three graduate nursing programs that have a clinical nurse specialist tract in genetic counseling (26).

Most genetic counseling programs are two year curricula that include theoretical coursework, laboratory exposure, research experience, and extensive clinical training. The curricula have expanded considerably in both depth and complexity from the initial recommendations outlined in 1979. Coursework usually includes human and medical genetics, cytogenetics, biochemical genetics and molecular genetics as well as principles of quantitative and population genetics. Most programs balance the basic science and medical genetics courses with courses that stress the psychosocial aspects of genetic counseling, including principles of genetic counseling and the counseling process, interviewing techniques, ethical, legal, social, and ethnocultural issues pertaining to genetic evaluation, screening, testing, and counseling. Clinical training covers all areas provided by genetic services. Students usually receive extensive exposure in prenatal diagnosis, general genetics and metabolic clinics including both children and adults, specialty clinics such as cystic fibrosis or craniofacial clinic, and such clinics as hereditary cancer and neurodegenerative disorders. Students are expected to take an active role in most counseling sessions under the supervision of a certified genetic counselor and/or medical geneticist in order

for the counseling session to be used toward formulating a logbook of cases (27). Today's genetic counseling student may have well over 1000 hours of clinical practicum (28), far greater than the 1979 "optimum curricula" recommendations that suggested 400 hours of supervised clinical placement in a minimum of two settings (25).

National Society of Genetic Counselors

The genetic counseling profession continued to establish its presence as a knowledgeable member of the genetics health care team with the formation and incorporation of its professional society, the National Society of Genetic Counselors (NSGC) in 1979 (21). NSGC instilled the sense of professionalism needed by this emerging group of health care professionals. Today NSGC has over 1700 members. Its mission is promoting "the genetic counseling profession as a recognized and integral part of health care delivery, education and public policy" (21). NSGC promotes the professional interests of genetic counselors and provides a network for professional communications, local and national continuing education opportunities, and the discussion of issues relevant to human genetics and the genetic counseling profession (21). In 1991 NSGC adopted and then published its Code of Ethics (29). This publication provided NSGC with a framework within which its members could function and practice. Moreover as NSGC continues to grow, its presence and influence are becoming more widespread regarding programs and policy in medical genetics and genetic counseling. NSGC has representatives on a number of policy-making groups and works closely with such groups as the American Society of Human Genetics, the American College of Medical Genetics, the American Board of Genetic Counseling, the American Board of Medical Genetics, the International Society of Nurses in Genetics, and the National Coalition for Health Professionals Education in Genetics.

Certification and Accreditation

Another important milestone in the professional development of genetic counselors occurred with the initiation of a certification examination process in 1981 by the American Board of Medical Genetics (25). Initially physicians, Ph.D. trained geneticists, and genetic counselors who met specific credential requirements could apply to sit for the certification examinations within five categories. Certification, or being eligible to sit for certification examinations, gave employers some degree of confidence that health professionals hired to provide genetic services possess a minimum level of knowledge in human and medical genetics and genetic counseling. Most employers require that genetic counselors be certified or eligible to sit for the certification examinations as prerequisites for employment.

In 1990 the genetic counseling subspecialty section of the certification examinations came under the auspices of its own board, the American Board of Genetic Counseling (ABGC). Current practice requires that applicants applying to sit for certification examinations have graduated from an accredited genetic counseling program. The applicant must also demonstrate that they have acquired a wide range of clinical experiences documented by a logbook of cases supervised and signed by a board certified genetic counselor or medical geneticist. Currently over 1300 genetic counselors have achieved board certification.

The most recent achievement in the professional evolution and development of the genetic counseling field has been the establishment of an accreditation process for master's programs in genetic counseling. Following a consensus meeting that included genetic counseling program directors, members of ABGC and experts in a number of different fields, the ABGC published its criteria for accreditation including the practice-based competencies as defined by the consensus meeting (30). Programs must demonstrate adequate institutional support and facilities, adequate leadership and management, including board certified genetics professionals, and a curriculum that provides educational experiences, including theoretical courses, clinical training, and supplementary activities that would provide graduates with the necessary knowledge and skills to perform, accurately and reliably, the functions of a genetic counselor (27).

Continuing Education. With the explosion of new information in human genetics, genetic counselors are finding it necessary to continue to expand their knowledge base and keep abreast of new developments in the field. This continuing education effort has taken the form of requiring certified genetic counselors to obtain documentation of continuing educational activities through continuing education units (CEUs). While NSGC developed a continuing education model in the early 1980s to ensure quality educational programs, it was not until 1996 that courses and conferences sponsored by NSGC had to meet established continuing education criteria (31). Certified genetic counselors may demonstrate their continuing education by either sitting for a re-certification examination or obtaining a minimum of 250 approved contact hours during a period of 10 years. NSGC is continuing its efforts to broaden the types of activities by which counselors can obtain the appropriate educational experiences.

Employment

As noted previously, genetic counselors usually work as members of the health care team in a number of settings. Once every two years, NSGC conducts a status survey of its membership to obtain information about employment, professional roles, and activities (32). In the survey conducted in spring 1998, almost half (47 percent) of the counselors reported working in a university medical center and 24 percent reported working in a private hospital or medical facility. The remaining counselors were scattered between other categories such as HMO's (7 percent), diagnostic laboratories (6 percent), federal/state/county offices (5 percent), and self-employed (1 percent). Almost three-quarters of counselors (73 percent) stated they worked in two or more specialty areas. These areas included prenatal genetic counseling (70 percent), followed by pediatric genetics (45 percent), and cancer genetics (34 percent). Of potential interest to new graduates of genetic counseling training programs, the 1998 survey noted that it took less than two months for

members to obtain genetic counseling positions following graduation (32).

Issues in Genetic Counseling

While there are a number of issues that genetic counselors encounter such as dealing with nonpaternity or ambiguous findings on genetic test results (covered elsewhere in this text), two issues that continue to be important to the profession are nondirective counseling and financial reimbursement of genetic counseling services.

Nondirective Counseling. From Sheldon Reed's first publication defining genetic counseling as a "a kind of social work done for the benefit of the whole family entirely without eugenic connotations" (3, p. 335), genetic counseling has been equated with the concept of nondirectiveness. Nondirectiveness appealed to the genetics community as a way to distance itself from the eugenics movement associated with Nazi Germany. Genetic counseling embraced Carl Rogers's client-centered counseling approach. Rogers felt that counselors need to provide a warm, accepting environment free from pressure or coercion for clients to reach a successful self-acceptance and self-understanding (33). Nondirectiveness is understood to mean nonprescriptive. Fine defines nondirectiveness as a genetic counseling strategy that supports autonomous decision making by clients (34). NSGC Code of Ethics states, "Therefore, genetic counselors strive to: ... Enable their clients to make informed independent decisions, free of coercion, by providing or illuminating the necessary facts and clarifying the alternatives and anticipated consequences" (35). Genetic counselors therefore are facilitators and advocates of informed decision making, with the goal of having the counselee make a decision based solely on his or her own values and beliefs.

However, consensus regarding the terms directiveness and nondirectiveness is difficult to find among genetics professionals and in the literature (36). Kessler notes that depending on how one defines the term nondirective will determine whether or not it can be achieved (37). White notes that nondirectiveness is often equated with value-neutrality, which may "either imply that the counseling approach as a whole does not represent any values or moral positions, or it may refer to value-free communication, representing an ideal in which concepts and facts are expressed in impartial terms" (38). A number of authors have argued that counseling is never value-neutral. The types of information provided or not provided, the tone of voice, body language, all convey counselor values (39,40,37). Singer elaborates on this theme by noting that many of the decisions that patients make take place in an atmosphere of crisis and that the issues are often highly emotional (41). Counselors relay information that is often highly technical, while most counselees are likely to have limited knowledge of the biological and statistical issues that arise; they are a vulnerable population. Thus a counselor, who has a duty to provide all the information that clients need in order to make informed decisions, must decide what information the counselee needs and how to present the information. In this sense the genetic counselor utilizes her expertise to decide what and how much information the counselee needs to make the best possible decision for her. Brunger and Lippman (42) would agree. They conclude that genetic counseling is not a "one-size-fits all" endeavor; rather, it is information that is tailored to specific counselees in specific situations. For some, this would be considered a directive approach. However, authors such as Kessler or Singer would suggest that the counselor is facilitating the goal of genetic counseling by providing information upon which counselees can make autonomous, independent, and informed decisions (41).

One of the often quoted mechanisms for deciding whether or not a counselor is being directive is to ask whether or not counselors answer questions such as "What would you do in my situation?" Michie and colleagues (43) analyzed 131 transcripts of genetic counseling sessions and quantified directiveness based on how often counselors provided advice (what was best for the client), evaluation (provided views about an aspect of the counselee's situation), and reinforcement (provided statements that affirmed or rejected a counselee's behavior, thoughts or emotions). The authors found a mean of 5.7 advice statements per counseling session. Moreover none of the genetic counselors in these sessions rated their counseling as "not at all" directive. Interestingly there was no significant association found between any of the measures of directiveness and such outcome measures as satisfaction with information, nor any information on whether the counselees' expectations were met or on the amount of anxiety and concern (43). Bernhardt concluded that the study found that the practice of genetic counseling "is not characterized ... as uniformly nondirective" (43, p. 40), and that it provided "data to substantiate the long held impression that nondirective genetic counseling is impossible to achieve" (44, p. 17). Kessler interpreted Michie and colleague's data differently based on a definition of directiveness that emphasizes coercion. He noted that if there is an attempt by the counselor, through deception, threat, or coercion, to undermine the individual's autonomy and compromise his or her ability to make an autonomous decision, that can be defined as directiveness (37). Moreover Kessler points out that one does not have to answer the "What would you do" question to be directive.

In a study by Bartels et al. (45), 781 members of NSGC were surveyed to assess how they defined nondirectiveness and what they actually did in practice. Of the 383 respondents, 96 percent reported viewing nondirectiveness as very important; however, 72 percent stated they were sometimes directive. Bartels concluded that although nondirectiveness was a goal of genetic counseling, it was not the only goal. They found that counselors made important distinctions between concerns for directiveness about the decision-making process and directiveness about decision outcomes. Counselors noted that they should take responsibility for directing a counseling session such as clarifying counselee expectations and questions, sharing genetic information, and facilitating understanding and communication. Counselors felt that recommending genetic or medical testing was consistent with informed consent, and that principle outweighed

being nondirective (45). However, counselors were more conflicted regarding decision outcomes such as having counselors share personal biases or values with their patients. The study participants described the need to be very careful in what and how these values would be presented. However, once counselees made a decision, genetic counselors felt that the decision should reflect the counselee's values and not the counselor's values.

While the literature regarding nondirective counseling remains unsettled, research suggests that nondirectiveness is not the only guiding principle employed by genetic counselors. In their efforts to provide and facilitate autonomous, independent, and informed decision making by counselees, genetic counselors strive to maintain a delicate balance between a nondirective stance and enhanced counselee understanding.

Reimbursement. The issues of financial reimbursement of genetic counseling services is also of major concern to genetic counselors. Currently there is a paucity of literature that speaks to reimbursement of services provided by counselors. In most instances, reimbursement for genetic counseling services provided by a genetic counselor is under the control of the physician provider. While nonphysicians can bill insurers for their services, they may or may not be reimbursed as some insurers do not recognize nonphysicians as health care providers. Moreover Medicare and Medicaid require that professional services of a nonphysician must be rendered under the physician's direct supervision (Medicare Part B Carriers Manual). Currently the Economics of Genetic Services Committee of the American College of Medical Genetics is working with the American Medical Association CPT Editorial Committee in an effort to obtain billing codes that are more representative of genetic services (46).

NSGC has also actively pursued licensure as a possible avenue to improving the reimbursement of services provided by genetic counselors. Currently the only state considering licensure of genetic counselors is California (47). A continuing and frustrating challenge for genetic counselors is to seek avenues of reimbursement within the context of a managed care environment. Another challenge remains how to educate other health care providers, hospital administrators, insurers, and the public to the skills and expertise that genetic counselors have to offer.

OTHER HEALTH PROFESSIONALS

Medical genetics is fast becoming central to the delivery of health care and preventive services. This process is in large part due to the progress in the Human Genome Project which promises, and in some cases has brought to fruition, wide-ranging applications in the diagnosis, treatment, and prevention of human diseases (48). Gene testing for a number of different disorders is now available to large segments of the population. For example, gene testing for breast and colon cancer susceptibility is available for individuals with a positive family history (49), an NIH Consensus Development Conference (50) has advocated offering DNA-based carrier testing for cystic fibrosis to pregnant couples, and population screening is being considered for hemochromatosis (51). These advances, coupled with inquiries regarding genetic testing from patients, have made it increasingly necessary for primary care providers to be informed about genetic information and testing. Ideally genetic services are best provided by geneticists and genetic counselors who have had extensive training and exposure to the complex genetic principles and issues regarding genetic testing, screening, and counseling. Although genetic counselors remain the "gold standard" for the provision of genetic services, there is a growing demand that other health care providers be able to provide some of these services. Nevertheless, it is evident from a number of studies that neither physicians nor nurses, the primary sources of health care, have received the training and knowledge needed to deal with the proliferation of new genetic information. Two recent surveys, one of physicians and one of nurses, demonstrated that while the majority of health care professionals stated they were providing some genetic related services (at least occasionally), provider knowledge of genetics was fragmented and uneven (52,53). Giardiello (54), in a study of the use and interpretation of gene testing for Familial Adenomatous Polyposis (FAP), found that one-third of physicians ordering the tests were misinterpreting the results and did not fully understand their meaning. The authors concluded that the "use of genetic counseling before testing would be expected to eliminate many ... errors" (54, p. 826).

With the move toward managed health care and a focus on primary care, it is the physicians and nurses who will now be ordering and interpreting many genetic tests, providing genetic counseling, and assuming responsibility for obtaining informed consent and protecting client privacy. In an effort to rectify some of the knowledge deficit, a number of organizations have come together to ensure that health professionals are prepared for this era of genetic technologies and testing. The National Coalition for Health Professional Education in Genetics (NCHPEG), a group of professionals spearheaded by the National Institute for Human Genome Research, the American Medical Association, and the American Nurses Association, has brought together genetics and other health care professionals to ensure that physicians and nurses have the knowledge, skills, and resources to integrate new genetic knowledge and technologies into the prevention, diagnosis, and management of disease (55). NCHPEG has established a number of goals for developing and implementing a comprehensive genetics education initiative. Its goals include persuading health care organizations to establish genetics education as a top priority; creating mechanisms for collaboration between genetics professionals, other health care professionals, consumers, industry, and government; and identifying existing and future genetics education activities.

Primary health care providers may assume a number of different roles in the care they give to individuals who have or are at risk for a genetic disorder. There are major areas and roles identified by both physician (56) and nursing genetics professionals (53), in which all primary care providers should be knowledgeable. These include

(*1*) identifying individuals and families who would benefit from genetic services and counseling; (*2*) knowing how to obtain genetic services once a risk factor for a genetic disorder has been identified, (*3*) being able to interpret risk factors and information in genetic test results and explaining the information to patients and their families, and (*4*) having a foundation to read and interpret the medical literature in order to provide competent guidance to patients and families regarding their questions about genetic testing, gene therapy, and genetic disorders. Additionally primary care providers can help families learn about what to expect in terms of a genetic evaluation and the genetic counseling; they can coordinate care for the family by assisting in the arrangements for diagnostic testing, referrals for further evaluations, and appointments with other specialists if appropriate; and they can reinforce, interpret, and clarify information obtained during a genetic counseling evaluation. Certainly the primary care provider could continue to assess family dynamics, coping strategies, and other psychosocial responses and serve as a resource person and educator to his/her patients and families regarding birth defects and genetic disorders.

These roles and responsibilities, however, should not be construed as a substitute for referring patients and families for genetic services and counseling. Ideally the primary care provider will work in partnership with the genetic counselor and recognize that a major component in the genetic counseling process is the opportunity to generate a dialogue that will result in a patient being able to make an informed decision. This is particularly necessary in presymptomatic or susceptibility genetic testing, such as Huntington disease and BRCA1 and BRCA2 testing for breast cancer, or in screening programs such as cystic fibrosis (50). The dialogue becomes part of an informed decision-making process rather than merely the transmission of information (57). Counseling skills required for such interactions must combine respect for a patient's right to make an autonomous decision with an appropriate level of support to facilitate the decision-making process. Genetic counseling provides this component and blends the informational, educational aspect of the counseling session with a dialogue on the benefits and risks of a genetic test. Rarely does a primary care provider have the time or specific training to be responsible for this process.

SUMMARY

The advent of recombinant DNA research and the initiation of the Human Genome Project has produced a revolution in human genetics. Every week scientists announce the identification and chromosomal location of another gene associated with a human genetic disorder. With these discoveries comes the ability to test for these disorders, both diagnostically and predictively. For some people it may mean the development of new treatment strategies. It is imperative that health professionals, in particular primary care providers, be informed about human genetics, genetic testing, and the ethical, legal, and social issues that accompany these areas of research and practice. Primary care providers are on the front line of providing health care services to families. Thus they need to be able to identify, assess, counsel, and refer patients, clients, and families with or at risk for genetic disorders and birth defects.

Finally, as new tests for carriers are developed and the diagnosis of genetic conditions and genetic susceptibility to disease continues to grow, the number of individuals utilizing genetic services will also increase. As the number of individuals and families seeking services increases, the need for genetic counselors to provide quality genetic counseling will also grow. Genetic counselors, utilizing their expertise and skill within the genetic counseling process, will be needed to explain facts, educate patients regarding inheritance and recurrence, provide counseling about benefits and risk, aid in decision making, and help set public policy. Well-trained, board-certified genetic counselors will continue to expand their presence in a number of settings: in hospitals and clinics to counsel families who are affected or may be at risk for genetic disorders; in diagnostic labs as resources for physicians and their patients; and in government agencies to design genetics education programs for health care providers, shape public health policy, and develop more effective ways of communicating the many new findings to employers, insurers, and the general public.

BIBLIOGRAPHY

1. S.C. Reed, *Counseling in Medical Genetics*, W.B. Saunders, Philadelphia, PA, 1955.
2. S.C. Reed, in H.G. Hammons, ed., *Hereditary Counseling: A Symposium sponsored by the American Eugenics Society*, Hoeber & Harper, New York, 1959.
3. S.C. Reed, *Soc. Biol.* **21**, 332–339 (1975).
4. R.H. Kenen, *Soc. Sci. Med.* **18**, 541–549 (1984).
5. F.C. Fraser, *Birth Defects Original Article Series, National Foundation March of Dimes* **15**, 5–15 (1979).
6. C. Stern, *Principles of Human Genetics*, 3rd ed., W.H. Freeman, San Francisco, CA, 1973.
7. N.L. Nadler, *Pediatrics* **42**, 912–918 (1968).
8. J.R. Sorenson and A.J. Culbert, *Birth Defects Original Article Series, National Foundation March of Dimes* **15**, 85–102 (1979).
9. A.P. Walker, in D.L. Rimoin, J.M. Connor, and R.E. Pyeritz, eds., *Emery and Rimoin's Principles and Practice of Medical Genetics*, 3rd ed., Churchill Livingstone, New York, 1996, pp. 595–618.
10. C.J. Epstein et al., *Am. J. Hum. Genet.* **27**, 240–242 (1975).
11. J.G. Hall, in R.E. Stevenson, J.G. Hall, and R.M. Goodman, eds., *Human Malformations and Related Anomalies*, vol. 1, Oxford University Press, New York, 1993.
12. E. Sujansky and A.L. Matthews, in G.B. Merenstein and S.L. Gardner, eds., *Handbook of Neonatal Intensive Care*, 4th ed., Mosby-Year Book, St. Louis, MO, 1998, pp. 604–624.
13. J. Leana-Cox et al., *Am. J. Med. Genet.* **65**, 309–316 (1996).
14. J.M. Aase, *Diagnostic Dysmorphology*, Plenum Medical Book Company, New York, 1990.
15. S. Kessler, in S. Kessler, ed., *Genetic Counseling: Psychological Dimensions*, Academic Press, New York, 1979.
16. S. Kessler, *Am. J. Med. Genet.* **72**, 164–171 (1997).

17. A.L. Matthews, *Birth Defects Original Article Series, National Foundation March of Dimes* **26**, 168–175 (1990).
18. J.L. Benkendorf et al., *Am. J. Med. Genet.* **73**, 296–303 (1997).
19. K.A. Schneider, *Counseling about Cancer: Strategies for Genetic Counselors*, Graphic Illusions, Dennisport, MA, 1994.
20. E. Kodish, G.L. Weisner, M.J. Mehlman, and T.H. Murray, *J. Am. Med. Assoc.* **279**, 179–181 (1998).
21. National Society of Genetic Counselors, *Genetic Counseling as a Profession*, NSGC Executive Office, Wallingford, PA, 1999.
22. D.M. Bartels, in D.M. Bartels, B.S. LeRoy, and A.L. Caplan, eds., *Prescribing Our Future: Ethical Challenges in Genetic Counseling*, Aldine de Gruyter, New York, 1993.
23. J.H. Marks, in D.M. Bartels, B.S. LeRoy, and A.C. Caplan, eds., *Prescribing Our Future: Ethical Challenges in Genetic Counseling*, Aldine de Gruyter, New York, 1993.
24. K.W. Dumars et al., eds., *Genetic Associates: Their Training, Role and Function: A Conference Report*, U.S. Department of Health, Education and Welfare, Washington, DC, 1979.
25. J.A. Scott, A.P. Walker, D.L. Eunpu, and L. Djurdjinovic, *Am. J. Med. Genet.* **42**, 191–199 (1987).
26. ABGC News Corner, American Board of Genetic Counseling, Spring 2000, <http://www.faseb.org/genetics/abgc/abgcmenu.htm>
27. American Board of Genetic Counseling, Requirements for graduate programs in genetic counseling seeking accreditation by the American Board of Genetic Counseling, American Board of Genetic Counseling, Inc. Bethesda, MD, 1996.
28. A.C.M. Smith, *J. Genet. Counsel.* **2**, 197–211 (1993).
29. J.L. Bendendorf et al., *J. Genet. Counsel.* **1**, 31–43 (1992).
30. B.A. Fine, D.L. Baker, M.B. Fiddler, and ABGC Consensus Development Consortium, *J. Genet. Counsel.* **5**, 113–121 (1996).
31. D.L. Baker, J.L. Schuette, and W.R. Uhlmann, *A Guide to Genetic Counseling*, Wiley, New York, 1998.
32. K.A. Schneider and K.J. Kalkbrenner, *Perspect. Genet. Counsel.* **20**(3), S1–S8 (1998).
33. C. Rogers, *Client-Centered Therapy: Its Current Practice, Implications, and Theory*, Houghton Mifflin, Boston, MA, 1951.
34. B.A. Fine, in D.M. Bartels, B.S. LeRoy, and A.C. Caplan, eds., *Prescribing Our Future: Ethical Challenges in Genetic Counseling*, Aldine de Gruyter, New York, 1993.
35. National Society of Genetic Counselors, *J. Genet. Counsel.* **1**, 41–43 (1991).
36. S. Kessler, *J. Genet. Counsel.* **1**, 9–17 (1992).
37. S. Kessler, *Am. J. Hum. Genet.* **61**, 466–467 (1997).
38. M.T. White, *J. Genet. Counsel.* **6**, 297–313 (1997).
39. A. Clarke, *Lancet* **338**, 998–1001 (1991).
40. A. Lippman and B.S. Wilfond, *Am. J. Hum. Genet.* **51**, 936–937 (1992).
41. G.H.S. Singer, in B. Gert et al., eds., *Morality and the New Genetics: A Guide for Students and Health Care Professionals*, Jones and Bartlett, Boston, MA, 1996.
42. F. Brunger and A. Lippman, *J. Genet. Counsel.* **4**, 151–167 (1995).
43. S. Michie, F. Bron, M. Bobrow, and T.M. Marteau, *Am. J. Hum. Genet.* **60**, 40–47 (1997).
44. B.A. Bernhardt, *Am. J. Hum. Genet.* **60**, 17–20 (1997).
45. D.M. Bartels, B.S. LeRoy, P. McCarthy, and A.L. Caplan, *Am. J. Med. Genet.* **72**, 172–179 (1997).
46. D.L. Doyle, Personal communication, September 1999.
47. A.P. Walker, Personal communication, October 1999.
48. E.P. Hoffman, *Am. J. Hum. Genet.* **54**, 129–136 (1994).
49. S.J. Laken, G.M. Petersen, and S.B. Gruber, *Nat. Genet.* **17**, 79–83 (1997).
50. Genetic Testing for Cystic Fibrosis, NIH consensus statement, No. 15, April 14–16, 1997, pp. 1–37.
51. T. Cox, *Nat. Genet.* **13**, 386–388 (1996).
52. K.A. Hofman et al., *Acad. Med.* **68**, 625–632 (1993).
53. C. Scanlon and W. Fibison, *Managing Genetic Information: Implications for Nursing Practice*, American Nurses Publishing, Washington, DC, 1995.
54. F.M. Giardiello et al., *N. Engl. J. Med.* **336**, 823–827 (1997).
55. National Coalition for Health Professional Education in Genetics, Planning meeting, July 23, 1996.
56. J. Stephenson, *J. Am. Med. Assoc.* **279**, 735–736 (1998).
57. G. Geller et al., *J. Am. Med. Assoc.* **277**, 1467–1474 (1997).

See other entries GENETIC INFORMATION, ETHICS, FAMILY ISSUES; see also REPRODUCTION, ETHICS entries.

GENETIC DETERMINISM, GENETIC REDUCTIONISM, AND GENETIC ESSENTIALISM

ROBERT WACHBROIT
University of Maryland
College Park, Maryland

OUTLINE

INTRODUCTION

In the public debate over the significance of genetic discoveries, the terms "genetic determinism," "genetic reductionism," and "genetic essentialism" are often used interchangeably. This is unfortunate because these terms refer to quite distinct claims about the importance of genes, claims that vary greatly in their plausibility, their implications, and their popular acceptance. Thus, for example, few people would accept genetic determinism, but many would subscribe to doctrines that are characteristic of some version of genetic reductionism or essentialism. The purpose of this article is to clarify what each of these terms mean, and how they differ from each other. The first, I will show, is a claim about causation; the second, about explanation; and the third, about identity.

Being clear about these terms is not just a matter of lexographic tidiness; it is critical for avoiding serious misunderstanding. The significance of this misunderstanding can be illustrated by contrast with another common confusion in the area of genetics; between genes and alleles.

Articles and discussions about genetics often confuse genes and alleles, using "genes" when the proper term should be "alleles." But this muddle is usually harmless. The knowledgeable and attentive reader can usually understand which term the writer intends even if they happen to be misused at the time. No significant issue is raised by this common conflation.

The situation is quite different with regard to the above "isms." Many writers reject what they regard as an inflated importance attached to genetics by rejecting genetic determinism, genetic reductionism, or genetic essentialism. These three positions are, however, quite different ways of articulating what the importance of genetics consists in. Not only does one not entail either of the other two, but also, as we will see below, each raises different significant issues.

The underlying assumption of this article is that the semantic structure of these "isms" should be taken seriously: Genetic determinism is a kind of determinism. Its implications, significance, considerations for and against, should be shaped by the history of the dispute over determinism itself. Otherwise, the expression is misleading. Similar assumptions hold for genetic reductionism and genetic essentialism. If they are not a kind of reductionism and a kind of essentialism, respectively, then we need to be told what they mean.

Finally, let me repeat, the purpose of this article is to clarify succinctly the differences between these "isms." The aim is not to provide a comprehensive survey of the many issues and the extensive literature surrounding these terms. If interested, the reader is advised to look to the bibliography.

DETERMINISM: GENETIC AND OTHERWISE

Determinism is a thesis about universal causation — every event has a cause sufficient for its occurrence. Thus identical states of the universe would have identical outcomes (1). The phrase "genetic determinism" would, strictly speaking, mean that every event has a genetic cause that is sufficient for that event's occurring. No one takes genetic determinism in this way. It is usually understood to be restricted to a specific class of events or properties — such as organism's physical and mental traits. Thus genetic determinism is the thesis that an organism's physical and mental traits are entirely the causal result of its genes.

So understood, the general consensus is that genetic determinism is false. This view is embedded in the standard distinction in genetics between an organism's genotype — the combination of the organism's genes — and its phenotype — its set of mental and physical traits. Genes alone do not yield any traits; various environmental factors must be present for the trait to arise. Regardless of how close the tie is between a certain gene and trait, there is always an environment where the presence of the gene does not result in the presence of the trait, including hostile environments in which the organism does not survive long enough to develop the trait.

Nevertheless, noting that genetic determinism is false does not mean determinism is false as well. One could reject genetic determinism and still hold that the combination of genes and environment determines traits. Or one could reject determinism even for the genes plus environment combination, noting the role of various stochastic processes in the development of traits (2, chap. 2). As we will see, this difference affects the significance of claiming that genetic determinism is false.

The significance of genetic determinism lies in its implications for predictability and manipulability. If genetic determinism is true, then it should be theoretically possible to predict people's traits solely from a knowledge of their particular genes. Furthermore the only way to alter an individual's particular trait would be to manipulate the relevant genes. If the rejection of genetic determinism is a rejection of determinism, then predictability and manipulability are also rejected. But if genetic determinism is rejected in favor of a "genes plus environment" determinism, then the rejection is much less significant. Prediction is still theoretically possible, though practically much more difficult given the complex variety of environmental factors that need to be considered. And manipulability becomes theoretically easier since there are now more options available.

Despite the widespread rejection of genetic determinism, we should acknowledge the occasions when researchers will talk about a gene "determining" a trait — for example, they might say that having a certain structure on chromosome 4 determines that the individual will develop Huntington's disease. This expression of genetic determinism rests on a special restriction on the events being considered. All circumstances in which the individual dies before he manifests the disease are being excluded. Furthermore we are confining attention to only known circumstances — no one is denying that researchers might some day discover that currently unknown circumstance in which we are able to prevent or cure that disease. Once these restrictions are understood, such talk does not really amount to an endorsement of genetic determinism.

REDUCTIONISM: GENETIC AND OTHERWISE

Reductionism is one of the main topics of modern philosophy of science, and during that time it has undergone considerable sophistication and complexity in its formulation and arguments. For a standard introduction, see Ref. 3. For an advanced discussion of this topic as applied to genetics, see Refs. 4–6. Fortunately, for our purposes, many of these details can be ignored. The core claim of reductionism — a thesis about the relation between two kinds of things (X and Y) — is that X's are nothing but Y's. Different ways of specifying the meaning of the relation "is nothing but" as well as the different sorts of things at issue yield different types of reductionism. For example, taking the "is nothing but" relation to be that of composition, we get a class — sometimes called "ontological" — of reductionist positions. Illustrations from the physical sciences provide some of the least controversial examples, such as "physical objects are nothing but swarms of atoms" and "water is nothing but hydrogen and oxygen."

One of the most familiar reductionist positions in this class is variously called "materialism" or "physicalism." It is roughly the view that everything is nothing but atoms, parts of atoms, and their interactions. This reductionism is a central principle of modern biology. Biologists reject vitalism and instead believe that biological objects are composed entirely of the stuff of physical objects — molecules, free-floating atoms, electrons and the like.

Genetic reductionism, understood ontologically, is the position that organisms consist of nothing but genes. It should be clear that no one is a genetic reductionist in this sense since the position is so obviously false. We are clearly composed of more than just DNA molecules. Indeed, the occasional expression such as "we are nothing but our genes" cannot be seriously meant literally. It must be understood as figurative way of referring to a different sense of reductionism, to be described below. Thus, while all biologists, including geneticists, are reductionists, in the sense of being materialists, no one is a genetic reductionist, understood ontologically. There is therefore no need to discuss it further.

The other major class of reductionist positions are those where the "is nothing but" relation is understood in terms of explanation and Y is understood as a theory. X can itself be a theory, where Y is a more fundamental or basic theory than X. Or X can be a representation or description of one type of phenomena, and Y is a theory typically about phenomena of a different, more fundamental realm. As opposed to ontological, this class of reductionist positions is sometimes called "epistemological" or "methodological." Fairly uncontroversial examples from the physical sciences would be "chemistry is nothing but physics" and "heat is nothing but molecular motion." In the first case, the thesis is that chemical theories can be reduced to physical theories: For example, the correctness of explanations in chemistry can be explained by physics. In the second case, the thesis is that heat, a phenomenon characterized by thermodynamics, can be explained by the more fundamental theory of statistical mechanics. In the end, both types of cases are similar: Chemical explanations can be replaced, at least in principle, by physical explanations; thermodynamic explanation can be replaced, at least in principle, by statistical mechanical explanations. Their differences are more in the methods of achieving reduction. In the first case, reduction is achieved by translating or associating the language of one theory into the language of a more fundamental theory; in the second case, reduction is achieved by directly explaining the phenomenon by using a fundamental theory.

Genetic reductionism, understood in this sense, can refer to either the thesis that biological explanations can be translated into or reformulated as genetic explanations or the thesis that biological phenomena can be explained by genetics. As we will see, unlike genetic determinism, genetic reductionism is an open question. It depends on how the science of genetics develops. Given the present state of the science, genetic reductionism is at best a working hypothesis, although some scientists claim that it is not a particularly promising hypothesis. In any case, whether genetic reductionism is true or false is an empirical matter.

Assessing the truth of genetic reductionism plainly turns on our understanding of what constitutes a genetic explanation. A precise characterization of genetic explanation is challenging, especially since there is considerable dispute over what constitutes an explanation (particularly explanations in biology!), but the intuitive idea is fairly clear. An explanation of a phenomenon is genetic just in case all the nongenetic activities or interactions that are part of the phenomenon are relegated to the background or are deemed to have minor importance. Thus a genetic explanation does not need to deny the existence of any nongenetic activity. Consider, for example, the genetic explanation for the probability of inheriting a certain (i.e., Mendelian) trait. One begins by determining whether to associate the trait with a dominant or recessive gene, and then one calculates the probability that the gene will be transmitted to offspring. The role of environment in the development of the trait is not denied; it is relegated to a background condition, deemed not a salient factor in the cases to be explained.

Similarly nongenetic explanations, such as environmental explanations, need not deny the existence of genes or genetic processes. A sociological explanation of family dynamics need not deny that genetic processes typically play a role in the existence of families. Such processes are deemed background conditions.

Not surprisingly, just as there is a controversy over what is an explanation, there is a controversy over what constitutes a background condition. Some have suggested objective criteria — for example, if it makes a comparatively small contribution to the phenomena, it is a background condition. An illustration of this can be found in explanations of differences. Suppose that we wanted to explain the differences in health between monozygotic twins. One might argue that since there is little genetic difference between the twins, genetic factors will be background conditions. So the explanation of differences in health will not be a genetic explanation.

Others have suggested subjective or pragmatic criteria — for example, factors that are not salient for the context of inquiry are background conditions. An explanation is here seen as primarily a response to a particular question, which characterizes, explicitly or implicitly, certain factors as salient. For example, a traffic accident is the result of how the car functions and how the driver functions. Whether we explain the accident in terms of the driver's behavior or the car's functioning can depend on which is, from the context of inquiry, deemed normal. Normality — either in behavior or in functioning — is usually relegated to a background condition.

We need not pursue any further the details of the concept of a background condition — or of explanation or of theory — in order to discuss the significance of genetic reductionism. As a scientific hypothesis about the power and scope of genetic explanations, it is plainly important for scientific research. But, as we will see, it also has social and ethical significance.

From the standpoint of science, genetic reductionism is a thesis about the direction scientific research should take. If genetic reductionism is true, investigations that focus on the genetic aspects of biological phenomena are

more likely to come up with interesting or important results. The patterns that emerge are more likely to reflect fundamental principles of biology than accidental or derivative correlations. Moreover such research is more likely to discover unifying themes in that divers biological phenomena might be explicable by the same type of (genetic) explanation. Indeed, genetic reductionism indicates what constitutes progress in biology — being able to account for more and more biological phenomena in terms of genetics.

The social significance of genetic reductionism is closely connected to these scientific implications. Genetic reductionism suggests that the "real" explanation of a biological phenomenon is a genetic explanation. Even if a nongenetic explanation of a particular biological phenomenon were available, it would not get to the heart of the matter. Nongenetic factors play a relatively minor role. It is a short step from this to holding that the nongenetic is of little importance or value. The significance of this becomes especially clear in the case of explaining human behavior. If many of our standard ways of explaining human behavior — explanations in terms of character or in terms of motivation or in terms of intentions or in terms of beliefs, desires, and circumstances — are held to be not "real" explanations or are held to be genetic explanations in disguise, then it can seem that character, intention, circumstances, and so forth, are unimportant or are themselves reducible to genetics. Indeed, some writers reject genetic reductionism because they maintain that nongenetic factors can be important.

At this point it might be worth summarizing some of the differences between genetic determinism and genetic reductionism. Their clearest divergence is with interactionism — the view that both genetic and nongenetic factors play a causal role in most biological phenomena. Interactionism is incompatible with genetic determinism but can be compatible with genetic reductionism. Recall that a genetic reductionist can acknowledge nongenetic causal factors as long as they amount to no more than background conditions of proper scientific explanations of biological phenomena. Thus the falsity of genetic determinism does not entail the falsity of genetic reductionism. (Whether the converse is true — whether genetic determinism entails genetic reductionism — turns on issues concerning the relation between causation and explanation, which we cannot pursue here.)

ESSENTIALISM: GENETIC AND OTHERWISE

The general doctrine of essentialism is linked to a particular conception of what change and identity consists in. It begins with the question, What is the difference between an object changing and an object ceasing to exist and being replaced by a different object? According to essentialism, an object has two kinds of properties — essential properties and accidental properties. An alteration in an accidental property results in the object being changed; an alteration in an essential property results in the object ceasing to exist. For example, the temperature of gold is an accidental property; change the temperature, and the result is still gold, only warmer or colder. In contrast, the atomic number of gold is an essential property; change it by even one digit and the result is no longer gold.

Genetic essentialism is the view that the genetic properties of an organism are essential. Given the multitude of genes, genetic essentialism fans out into a spectrum of specific views. At the extreme is the view that every gene — and hence every genetic property — is essential; if even one gene in the organism changes, say, from some environmental damage, the result is a different organism. Indeed, there is no such thing as a harmless error in gene replication on this view. An error means that the organism has ceased to exist and a new one, albeit quite similar, is now in its place. Moreover gene therapy must be seen as logically impossible. A procedure that alters the gene of an individual could not be therapy for that individual, since the alteration would destroy the individual and replace him with another. Hardly anyone is a genetic essentialist of this sort.

The situation becomes more complicated once we consider changes in certain genes or in many genes. How many and which phenotypic traits of an organism would have to be different — and to what degree — in order for us to regard the result as no longer the same organism? Some might hold that changing the genes that determine the organism's sex results in a different organism. Others might hold that in the case of persons, altering the genes associated with mental capacities and abilities can result in a different person. And others might hold that only those genetic changes that constitute a different species result in a different organism. Where we draw the line — when (genetic) change becomes replacement — is the central controversy regarding genetic essentialism. It also indicates the conceptual limits of gene therapy since altering an essential property can never be therapeutic (7).

This last point has been especially important in discussions of the alleged paradox of "genetic harm," which refers to a harm that is the result of an abnormal gene. The paradox arises if the abnormal gene is an essential gene. In that case, removing the harm, altering the essential gene, would not be altering the organism — presumably for the better — but rather replacing that organism with a different one. It would seem then that treating genetic harms when they involve essential genes could never be beneficial to the individual.

Genetic essentialism makes no claim about causation and so is distinct from genetic determinism, as we would expect from general discussions in philosophy regarding essentialism and determinism. Essential properties need not be deterministic (i.e., sufficient causes) and deterministic properties need not be essential. But see Ref. 8 which treats genetic essentialism as incorporating theses of determinism and reductionism.

Nor need genetic essentialism have anything to do with genetic reductionism of either type. A genetic reductionist of the ontological type — one who believes that organisms are composed of nothing but genes — need not be committed to any view regarding which or how many genes are essential to the identity of the organism. This claim is in line with the larger view that materialism doesn't entail any view about which material is essential. A genetic

reductionist of the epistemological type — one who believes that biological phenomena can be entirely captured in genetic explanations — need not be committed to any particular genetic essentialism. Essential properties do play a critical role in the explanation of change — how an organism can remain the same even when some of its properties are different — but it plays no role in scientific explanations. In short, a genetic reductionist need not have any view about which genes, if any, are essential.

CONCLUSION

Advances in genetics have highlighted the role genes play in various biological phenomena. This increased attention to genetics has led to various assertions and denials of the importance of genes.

Genetic determinism, genetic reductionism, and genetic essentialism are three different ways of stating what that importance consists in. The primary purpose of this article has been to articulate these differences. Any useful discussion of where genes are important and where they are not, if it invokes any of these "isms," must be clear about what exactly is meant.

BIBLIOGRAPHY

1. John Earman, *A Primer on Determinism*, Reidel Publishing Dordrecht, The Netherlands, 1986.
2. R. Lewontin, *Biology as Ideology: The Doctrine of DNA*, Harper Perennial, New York, 1992.
3. E. Nagel, *The Structure of Science*, Chapter 11, Hackett Publishing Company, Indianapolis, IN, 1979.
4. P. Kitcher, *Philosophical Review* **93**, 335–373 (1984).
5. S. Sarkar, *Genetics and Reductionism*, Cambridge University Press, New York, 1998.
6. K. Schaffner, *Discovery and Explanation in Biology and Medicine*, University of Chicago Press, Chicago, 1993.
7. D. Wasserman, Personal identity and the moral appraisal of prenatal therapy, in L. Parker and R. Arkeny, eds., *Evolving Disciplines: Genetics, Medicine, and Society*, Reidel Publishing Company, Holland, in press.
8. R. Dreyfuss and D. Nelkin, The Jurisprudence of Genetics, *Vanderbilt Law Review* **45**, 313–348 (1992).

See other entries BEHAVIORAL GENETICS, HUMAN; HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY: OVERVIEW; MEDIA COVERAGE OF BIOTECHNOLOGY; PROFESSIONAL POWER AND THE CULTURAL MEANINGS OF BIOTECHNOLOGY; PUBLIC PERCEPTIONS: SURVEYS OF ATTITUDES TOWARD BIOTECHNOLOGY.

GENETIC INFORMATION, ETHICS, AND INFORMATION RELATING TO BIOLOGICAL PARENTHOOD

JEAN E. MCEWEN
National Human Genome Research Institute
Bethesda, Maryland

OUTLINE

INTRODUCTION

Genetic testing now makes it possible to determine parental and other familial relationships with a remarkable degree of accuracy. For this reason genetic testing in disputed paternity cases (at least in cases of the traditional type, where the paternity of a child born to unmarried parents is at issue) has gained increasing acceptance, and the ethical and legal concerns raised by such testing are lessening. However, new ethical and legal concerns are being raised in the growing number of cases in which either putative or presumed fathers seek genetic paternity testing to rebut the longstanding legal presumption that a child born to a mother during the course of a marriage is the biological offspring of the mother's husband. The practice of artificial insemination by donor is also creating new challenges in the area of forensic paternity determination.

As techniques such as surrogacy, egg donation, and embryo donation come to be used more widely, forensic challenges regarding *maternity* determinations can also be expected to arise with greater frequency. In addition the increased use of all types of alternative reproductive methods (coupled with advances in genetic testing more generally) is presenting courts with complicated questions regarding the posthumous determination of parentage or other familial relationships. Should human cloning eventually become feasible, the task of sorting through the myriad potentially recognizable familial relationships will become even more complex.

Unanticipated findings regarding parentage (or regarding other familial relationships) can also sometimes occur as a consequence of genetic testing undertaken for nonidentification purposes, raising challenging ethical and legal dilemmas. Incidental unexpected findings of misattributed paternity, in particular, occur quite frequently in the course of genetic testing. Several possible strategies are available for dealing with findings of this type, ranging from full disclosure, nondisclosure, partial disclosure, or disclosure only to the woman, to handling the issue through the informed consent process. However,

each approach carries with it a separate set of potential problems.

Previously undisclosed adoption, artificial insemination, or an incestuous mating within the family can also be inadvertently brought to light in the course of genetic testing. In rare cases, genetic testing may also reveal prior mix-ups, such as that two babies were switched in a hospital nursery, that sperm samples were switched in the course of artificial insemination, or that frozen embryos were switched during in vitro fertilization (IVF) procedures. These situations raise problems similar, although not identical, to those that arise when misattributed paternity is discovered, and they will require novel approaches.

GENETIC DETERMINATION OF PARENTAGE FOR FORENSIC PURPOSES

General Background

Objective scientific methods for the determination of biological parentage are a relatively recent development. Only a few decades ago the methods available for determining disputed paternity were limited and relatively primitive. Courts were typically required to calculate the date of the child's conception to determine whether or not the putative father could have conceived the child (1) and to make highly subjective assessments of physical resemblances (2).

Modern scientific tests provide far more reliable evidence of biological parentage by analyzing inherited characteristics: either the physical expression of the DNA in the child (phenotype) or the DNA itself (genotype) (3). The earliest phenotypic tests for paternity involved the ABO blood grouping system, but ABO testing, while capable of excluding an individual (i.e., ruling him out) as a child's father, could not make a definitive positive determination of paternity (4). However, ABO testing can be combined with other kinds of phenotypic testing, such as the analysis of red blood cell antigens, serum proteins, enzymes, and human leukocyte antigens (HLA) (5). When this is done, the cumulative impact of the test results including an individual as the child's father (i.e., placing him within a population of men who could be the father) becomes considerably stronger.

Direct genotypic testing of the DNA, which first became available in the late 1980s, provides an even more highly discriminating form of parentage testing. It can establish paternity (or maternity, if disputed) to degree that experts today agree is nearly conclusive (6). Every state in the United States now has a statute that provides for the admissibility of genetic paternity test results (7). Although the language of these laws varies, most states have patterned their laws on one of two model uniform paternity statutes: the Uniform Parentage Act (UPA) (8) or the Uniform Act on Paternity (UAP) (9).

State statutes vary in the statistical analysis methods (if any) that can be invoked in court to help describe the probative value of a test result including a man as the possible father. For example, some statutes require use of the "paternity index," which is the probability that a child born to the mother and the alleged father would have the observed phenotypes or genotypes divided by the chance of such types appearing in the child of the mother and a man randomly selected from the general population of men. Others require use of the "probability of paternity," which is obtained by multiplying the paternity index by the prior odds of paternity and converting the resulting "posterior odds" to probability form. Still others mandate the "probability of inclusion" (or exclusion) approach, which simply asks which proportion of the male population would be included (or excluded) as the possible father (10).

Regardless of the particular statistical analysis authorized, most state statutes mandate that when paternity test results reach a specified level (typically in the range of 95 to 99 percent), they create a rebuttable presumption of paternity; some statutes go even further and mandate a *conclusive* presumption of paternity when the results exceed a certain threshold (11). The growing confidence of legislators and courts in the reliability of genetic paternity testing, and federal child support enforcement legislation enacted in 1993 requiring all states to enact laws authorizing simple voluntary procedures for the establishment of paternity (12), have resulted in most disputed paternity cases now being settled by agreement of the parties without the need for trial (although statistical issues and questions regarding the adequacy of laboratory procedures are still often raised) (13). As the subjectivity and unreliability of the methods associated with earlier methods of determining paternity have been replaced by modern genetic testing, the ethical and legal issues associated with the determination of paternity (at least in the traditional situation involving unmarried parents and child support) have also gradually lessened.

Emerging Issues in Genetic Determination of Parentage

Rebutting the Presumption of Paternity. Although most paternity disputes continue to arise out of proceedings for the support of children born to unmarried parents, courts are increasingly being called upon to adjudicate paternity in other contexts. For example, notwithstanding the longstanding legal presumption that the parent of a child born during the course of a marriage is the mother's husband (unless the husband was sterile, impotent, or geographically distant at the time of conception), a growing number of putative fathers of children born to married mothers (i.e., men who are not married to the children's mothers but who believe they may have fathered the children) are filing lawsuits to rebut that presumption and assert various parental rights. Although the United States Supreme Court has held that such putative fathers have no right to challenge the presumption of paternity as a matter of constitutional law (14), the continued usefulness of the presumption might well be questioned in cases where undisputed evidence shows that the mother's husband could not possibly be the biological father.

In a related development, some former husbands of women who gave birth to children the former husbands believe were fathered by someone else have begun to challenge their own legally presumed paternity, typically in an effort to relieve themselves of postdivorce child support obligations. At the time of this writing, a petition for *certiorari* in a case raising this issue was pending before the United States Supreme Court (15). On the one hand, it

can be argued that no man should be compelled to support a child for whose birth he is not actually responsible, and that no court should be complicit in aiding such a fundamental deception. On the other hand, overfocus on biological parenthood in such cases, and underfocus on the role of social parenting, could harm the children involved, some of whom could be left altogether fatherless and with a great sense of betrayal should their presumed fathers be permitted so easily to relinquish responsibility. While a wrong may clearly have been perpetrated against the husbands in these cases, the question becomes whether the law should create the potential for an even greater wrong to be visited against the innocent children involved. Still, where a family situation has already deteriorated to the point that a man who has previously held himself out as a child's father announces publicly his belief that the child is not biologically his and states that he no longer wishes to be responsible, there may be little to be gained by perpetuating a legal fiction.

Paternity and Artificial Insemination by Donor. Continuing increases in the numbers of children conceived through alternative methods of reproduction, such as through artificial insemination by donor (AID), will further expand the range of cases in which courts will be asked to make forensic parentage determinations. For example, currently state statutes modeled either on the UPA or on the Uniform Status of Children of Assisted Conception Act (USCACA) provide that in a case of AID of a married woman, the woman's husband — not the sperm donor — is treated as the natural father of the child conceived (so long as the husband has consented to the procedure) (16). Likewise statutes patterned on the Uniform Putative and Unknown Fathers Act (UPUFA) accord no legal recognition to men who donate sperm under circumstances indicating that the donor did not anticipate having an interest in the resulting child (17). Moreover the current practice in AID is to keep the identity of sperm donors strictly anonymous, thus making it virtually impossible for anyone (including their resulting children) ever to trace their identity (18). For this reason genetic testing is unlikely to play a role in the establishment of biological paternity in the context of AID, at least for the immediate future. This could change, however, should the standard of practice for AID move toward the adoption of enhanced recordkeeping and liberalized disclosure policies. Significantly the anonymity that currently pervades AID is quite analogous to the secrecy and sealed records practices that for many years pervaded traditional adoption. The justification for such practices is increasingly being called into question as the public becomes more and more aware of the importance for every child of having access to a complete and accurate family history (19).

Genetic Determination of Maternity. Before the advent of modern reproductive technology, determining the *maternity* of a child was essentially never at issue. The woman who gave birth to the child was considered the child's mother as a matter of biological necessity. Increasingly, however, techniques such as surrogacy, egg donation, and embryo donation are raising questions regarding who should be recognized as the legal mother of a child so conceived when agreements involving the use of these technologies break down (20). The UPA expressly creates an action to declare the mother-child relationship but contains no specific provisions regarding the adjudication of such cases.

Surrogacy agreements are presenting especially novel challenges for the courts. There are two forms of surrogacy. In a traditional surrogacy arrangement, a woman agrees to be artificially inseminated with the sperm of the intended father (a man other than her husband), to carry to term the child thereby conceived, and to relinquish the child after birth to the intended father (and presumably also to his wife or partner). In a pure gestational surrogacy arrangement, by contrast, a woman agrees to carry an embryo created not through her own genetic material but through in vitro fertilization of the egg and sperm of the intended parents (or of a donated egg and/or sperm), and later to relinquish the child to the intended parents. The distinction between the two forms of surrogacy is thus that whereas the "traditional" surrogate mother both provides the egg for the pregnancy and gestates the baby to term, the pure gestational surrogate bears no actual genetic relationship to the child.

Judicial resolution of the question of who should be considered the child's legal mother in surrogacy cases has differed depending whether the underlying agreement was one for traditional surrogacy or gestational surrogacy. In the leading U.S. case, the New Jersey Supreme Court held that the woman who gestates (and provides the egg for) the child — not the partner of the man who contracted with her to bear the child — is to be treated as the child's mother (21). By contrast, in the leading case involving a purely gestational surrogacy agreement, the court held that while both genetic consanguinity and giving birth are recognized means of establishing a mother–child relationship, in cases where the two means do not coincide in one woman, the woman who intended to raise the child at the time the agreement was entered into — not the woman who gave birth — should be treated as the mother (22). The court distinguished this situation from a true "egg donation" situation, in which a woman gestates and gives birth to a child formed from the egg of another woman with the intent to raise the child as her own.

The USCACA is drafted so as to give states the option either to accord legal recognition to preapproved surrogacy contracts that meet specified statutory requirements, or to make surrogacy contracts unenforceable, and it specifically provides that except in those cases involving preapproved contracts, a woman who gives birth to a child is to be considered the child's mother. However, in states that have not adopted the USCACA or that have not otherwise clarified this issue legislatively, considerable uncertainty remains regarding the establishment of maternity in surrogacy cases, as well as in cases involving egg and embryo donation. Moreover, outside the United States, approaches to surrogacy have been quite different. For example, in most countries, in a contest between a genetic mother and a pure gestational mother, the gestational mother generally prevails (23).

Posthumous Determination of Parentage or Other Familial Relationships. Because DNA is present in almost all human cells and remains unchanged long after a person has died, DNA testing technology now makes it possible to make an accurate determination of paternity (or of grandpaternity, great-grand-paternity, or even more distant relationships) long after the putative father (or other more distant relative) has died (24). For this reason the advent of DNA testing has brought with it an increase in requests for exhumation of remains to conduct such testing. Persons may be interested in establishing paternity (or other family relationships) posthumously for a variety of reasons, ranging from establishing entitlement to inherit, immigrate, or receive government benefits, to satisfying concerns of purely genealogical interest.

In the 1990s DNA testing was used for the first time to support the claims of several persons long claiming to be descendants of deceased celebrities or historical figures, and there is reason to expect that the number of such cases will increase in the future (25). Issues may also arise posthumously concerning the parentage of children conceived through alternative reproductive methods. For example, the sperm of a man may be frozen and then used after his death to artificially inseminate a woman, resulting in the conception of a child months or even years later. In such a situation, resort to posthumous genetic testing may become necessary to establish the biological relationship between the resulting child and the deceased father. At least one such case, involving an application by Social Security survivor's benefits on behalf of a child, who was conceived by gamete intrafallopian transfer three months after the death of her biological father, has already been litigated (26). The Social Security Administration initially denied the claim, reasoning that because the child was born 13 months after her biological father's death, she could not have been his legal heir. However, the agency subsequently reversed its position, reasoning that conclusive proof of the child's biological paternity could provide a constitutional nexus for securing her entitlement (27).

The technology for the freezing of eggs has not yet developed sufficiently to allow for eggs to be fertilized and successfully implanted following the death of the woman from whom they were derived, but should this technology improve, analogous issues will arise as the resulting children seek posthumously to establish their genetic maternity. Related issues will also emerge as a growing number of children are created from embryos that have been cryopreserved for use in IVF procedures and that are not implanted until after (perhaps even many years after) the death of both biological parents. For example, in Australia, in the early 1980s controversy erupted when a married couple died simultaneously in an airplane crash after having frozen embryos for use in IVF. Because the couple was exceptionally wealthy, a question arose over whether the not-yet-implanted embryos should be provided to a third party for implantation — and if so, whether any resulting children would be legally those of the deceased couple (and thus eligible to inherit their estate) (28).

The USCACA would resolve some of the legal uncertainty regarding the posthumous determination of genetic parentage by providing that an individual who dies before the implantation of an embryo, or before a child is conceived other than through sexual intercourse, using the individual's egg or sperm, is not considered the parent of the resulting child. Once again, however, in states that have not patterned their laws on the USCACA, the resolution of these issues remains unclear. In fact, separate and apart from resolving issues of the *parentage* of children resulting from the use of such techniques, courts are still struggling with preliminary questions regarding the legal status of frozen sperm (29) and the frozen embryos (30) *themselves* — that is, whether they should be treated as property, as human life, or as something "in between."

Human Cloning and the Determination of Familial Relationships. Should human cloning (the creation of a human being through somatic cell nuclear transfer technology or a similar technology) someday become feasible, the task of sorting through the myriad potentially recognizable familial relationships will become even more complex. Although the National Bioethics Advisory Commission concluded in its 1997 report on human cloning that it would at this time be morally unacceptable for anyone to attempt to create a child using somatic cell nuclear transfer technology (31), it is nonetheless likely that successful attempts at human cloning will eventually occur, and when they do, the question will arise who is the "parent" of the clone. The process of cloning will result in a child having genetic material from as many as four individuals: the person from whom the cell nucleus was derived, that individual's two biological parents, and the woman contributing the enucleated egg cell (which contains a small fraction of mitochondrial DNA). In addition, if the egg with the transferred nucleic material is implanted in a surrogate gestational mother, the child will have two other potential parents: the gestational mother and (if she is married) her husband. There may also be intended parents unrelated to the person who is cloned, such as in cases where the cloned person is deceased or a celebrity (32). In such cases, not only will it be necessary to decide whether the child's genetic parent(s) should be given precedence over the biological (but nongenetic parent) or over the purely social parent(s), but it will also be necessary to determine who should be recognized as the child's genetic parent(s) in the first place.

GENETIC DETERMINATION OF PARENTAGE AS A CONSEQUENCE OF GENETIC TESTING FOR NONIDENTIFICATION PURPOSES

Findings of Misattributed Paternity

While genetic testing in the forensics context is the common way in which information regarding biological parentage or other familial relationships is brought to light, such information can also be uncovered inadvertently in the course of genetic testing undertaken for completely unrelated purposes, such as in clinical medicine. The common incidental finding that occurs in the context of nonforensic genetic testing is the finding of misattributed paternity (or sometimes grand-paternity). The true incidence of misattributed paternity is unknown,

and undoubtedly varies widely depending on geographical region, age group, and cultural or ethnic group, among other factors. However, 10 percent is the figure most commonly cited, and estimates as high as 30 percent have been proposed (33). While more recent studies suggest that both of these figures may be substantial exaggerations (34,35), the accumulated experience of large-scale genetic screening programs (e.g., newborn screening programs) shows that the aggregate number of children born each year whose paternity is misattributed is by no means insignificant.

The incidental finding that the presumed father of a child is not the biological father can arise in a number of situations, such as when several family members are being tested to locate a suitable donor for a bone marrow or organ transplant, to take part in genetic linkage testing, or to participate in other types of genetic risk assessment that require samples from multiple family members. In cases where testing for bone marrow or donor organ compatibility yields evidence of misattributed paternity, it is often possible to communicate the fact that the individual tested and the intended recipient (e.g., two half-siblings) would not be a good match without mentioning anything about misconceptions regarding the degree of their biological relatedness. This is because it is easy to explain the *fact* of a mismatch without going into the apparent *reason* for the mismatch; there may be many reasons other than misattributed paternity why two people might not be suitably matched for transplantation purposes. The nondisclosure approach in the transplantation situation can also probably be justified ethically and legally because at least in most cases the nondisclosure is unlikely to have any direct, potentially adverse, effect on the parties' future personal medical or reproductive decision making.

On the other hand, when a finding of misattributed paternity surfaces in the context of genetic risk assessment, the stakes are higher because genetic risk estimates are based on the assumption that the biological relationships assumed to exist within a family are correct. A person's misunderstanding about his or her biological relationship to other family members can confound the clinical determination of whether he or she is at increased risk for an inherited disorder or for passing on an inherited disorder — with crucial ramifications for health and reproductive planning.

A common situation regarding misattributed paternity occurs when genetic testing is sought to determine recurrence risk following the birth of a child affected with an autosomal recessive genetic disorder, for which both parents must be obligate carriers. In some cases the woman may already suspect that another man fathered her child, and may thus seek out counseling on her own without involving her husband or partner. However, where the woman does not realize (or is in denial of the possibility that) the child has a different father, the entire family may become involved. If carrier testing in such a case reveals the presumed father not to be a carrier, this means that he cannot be the biological father. The genetic counselor or other provider then faces a dilemma: how to convey to the couple the reason *why* they are not at increased genetic risk for bearing another affected child without simultaneously disclosing the fact that the child they already have must have been fathered by someone other than the husband (or other presumed father).

Reconciling the competing interests in cases like these can be very difficult, and no one strategy for resolving the issue is likely to be entirely satisfactory. In fact international surveys of genetic service providers performed as recently as the 1990s revealed a marked lack of consensus regarding the appropriate resolution of this dilemma, even though it is one that genetics professionals have been wrestling with for years (36,37).

Alternative Approaches to Handling Findings of Misattributed Paternity

Full Disclosure. One approach to the problem of misattributed paternity — the approach recommended in 1983 by the President's Commission for the Study of Ethical Problems in Medicine and Medical and Behavioral Research — is for the genetic counselor or other provider frankly to disclose the finding, including the conclusion that the recurrence risk in any future pregnancy of the couple is virtually zero because the existing affected child is not biologically the husband's (38). This approach accords maximal weight to the principles of autonomy and beneficence. It also reflects the practical consideration that deception regarding a child's paternity is likely eventually to be discovered in any event, and that in the long run, greater disruption to the family may result from this than from the frank revelation of misattributed paternity made when the information first surfaces in the clinical setting.

However, this approach can been criticized (at least in many cases) as placing form over substance, giving insufficient allegiance to the integrity of the family unit, and naive in its failure to recognize that many women — especially those who are in abusive relationships or who are economically disempowered — may suffer tangible detriment, in the form of physical, psychological, social, or economic harm, from the disclosure to their husbands or partners of misattributed paternity. In fact, in some cultures, the social environment may be such that an almost certain consequence of such a revelation would be clear harm to both the woman and the child. Indeed, it was this concern that led the Hereditary Diseases Programme of the World Health Organization to conclude that there is probably never a justification for a provider to reveal such a finding to a husband (39). Thus, at the very least, in cases where a decision is made to reveal information regarding misattributed paternity to both partners simultaneously, the provider should be prepared to offer appropriate psychological and other support.

Nondisclosure, Partial Disclosure, or Disclosure Only to the Woman. Another possible approach to dealing with findings of misattributed paternity is either to misrepresent the finding (or the basis for the finding) or skirt the issue in some other way, either through some form of partial disclosure or by telling the woman alone. The justification for this approach is that the genetic counselor owes greater loyalty to the integrity of the family unit than to any one family member, and that

for the reasons mentioned above, revealing the complete truth simultaneously to both the husband and wife could do more harm than good.

An approach that misrepresents the facts, however, is also problematic, both from an ethical and a legal standpoint. First, to the extent that overt deception is involved (e.g., explaining away the child's disorder, and thus the reason for the lack of recurrence risk, as merely a spontaneous mutation or some other anomaly, or explaining away the test results as having been confounded by a mix-up in the testing laboratory), it risks jeopardizing the provider's professional integrity and lowers the standards of practice (40). Moreover, if the explanations given are viewed by the couple as implausible, the approach is likely to engender suspicion and mistrust. In fact, should the deception eventually be discovered, the provider could conceivably be liable for medical malpractice.

A particularly risky practice from the standpoint of legal liability is for the provider to lie outright if asked by the husband whether he is in fact the biological father (41). Nevertheless, surveys indicate that many providers follow this approach, sometimes justifying it on the basis that because the genetic testing is not being done for the purpose of discovering paternity, they have no obligation to reveal the finding, even when asked. In fact two-thirds of all United States geneticists in one survey stated that they would not tell a man that he is not the father of a child, even if he asked (42). On the one hand, in cases where a genuine and serious risk of harm to the woman or to the family appears likely if the information were to be disclosed, this approach may have considerable justification. On the other hand, as earlier discussed, it can be argued that making a genetic counselor complicit in the woman's intentional deception is always unethical, that secrets of this type are in any event unlikely to remain buried forever, and that when the truth does come out, the well-being of the woman and the family may be even more seriously jeopardized than if the deception had never been perpetuated in the first place.

A form of partial disclosure that may be less risky legally but that still raises significant ethical concerns is simply to avoid any discussion of the specifics regarding actual recurrence risk (e.g., by characterizing the results as inconclusive), thus obviating the need to discuss the husband's noncarrier status. This approach, however, may lead the couple to make inappropriate future reproductive decisions based on the erroneous belief that they are both in fact carriers (and thus have a 25 percent chance of having another affected child), when the actual risk is close to zero. Based on this inaccurate assumption, the couple may later resort unnecessarily to artificial insemination by donor, forgo future pregnancies altogether, or even divorce and seek new noncarrier mates. If they do decide to conceive another child together based on the misapprehension that such a child is at increased genetic risk, they may suffer needless anxiety and incur needless risk and expense associated with amniocentesis or other prenatal testing that is not in fact medically indicated.

Another approach is to convey to the couple the actual risk (close to zero, in the example discussed) while withholding the information about genetic transmission that would explain the reason for the risk and raise suspicions regarding nonpaternity (the fact that in order for a child to be born with an autosomal recessive disorder, both parents must carry the gene). This approach, however, may leave the couple feeling anxious and confused and lead them to suspect that something important is being withheld.

Yet another approach is to relate the finding only to the woman (who is likely in many cases to suspect anyway that another man fathered her child) and leave with her the choice as to whether or how to tell her husband or partner. This approach — the approach recommended by the Institute of Medicine's Committee on Assessing Genetic Risks in its 1993 report (43) — avoids the above-described difficulties that may arise when such a finding is revealed by an outsider, and at the same time, requires no overt misrepresentation or skirting of the issues in the provider's conversations with the couple. However, the approach seems difficult to reconcile with the notion that the ethical and legal obligations of genetic counselors run equally to both partners (44). Thus, if this approach is followed, the potential psychological benefits of disclosure, including relief from the burden of keeping a secret and greater honesty in family relationships, should be stressed with the woman. However, the potential for adverse consequences should also be raised, and once again, the provider should stand ready to provide other necessary support.

Informed Consent Approach. An emerging approach to dealing with unexpected findings of misattributed paternity is to try to avoid many of the above-described problems by addressing the issue before the testing takes place, as part of the informed consent process. Under this approach the woman (or in some cases, both partners) are informed, prior to taking the test, of the possibility that misattributed paternity will be discovered. The woman (or the couple) can then (at least in theory) agree *in advance* on the way such information, if discovered, will be handled. This approach has the advantage of making the persons most likely to be affected active participants in any decision about disclosure. However, it too has limitations, due the practical realities of the context in which genetic testing typically occurs. The very inclusion of the subject of paternity among the subjects treated in an informed consent document may provoke anxiety, and where the woman is aware that misattributed paternity may be an issue, she may well "panic" in the situation, having never seriously thought before about the ramifications of that possibility. Even if the pre-test counseling is done separately, the woman may feel confused about how to now "get out of" a test she had previously seemed to agree to (before the possibility was called to her attention). In the end this approach could discourage some women who would like to obtain genetic information from participating in genetic testing — perhaps, in some cases, to the future detriment of themselves and their families.

Some genetic testing centers include in their standard informed consent form a reference to the possibility of an incidental finding of misattributed paternity, but

simply state what the center's policy is regarding the communication of such findings without giving the woman (or the couple) an opportunity to communicate her (or their) preference in this regard. This approach is problematic for many of the same reasons. In addition it is based on the (typically erroneous) assumption that a couple who is uncomfortable with a center's policy can simply "go elsewhere" for testing. Testing for some genetic disorders (particularly those that are relatively rare) may, as a practical matter, only be available at a single location. Insurance and other practical constraints may also limit a couple's ability to "shop around" for a center with a more favorable disclosure policy.

Other Possible Unanticipated Findings Regarding Familial Relationships

Misattributed paternity is not the only type of unanticipated finding regarding biological parentage that may surface when genetic testing is sought for nonidentification purposes. Genetic testing may also bring to light the fact that a child has been adopted, was conceived through donor insemination or another alternative reproductive method, or is the product of an incestuous mating (perhaps within the extended family). Situations like these raise ethical concerns similar to, but slightly different from, those present in cases where paternity has simply been misattributed by the mother to her husband or partner. Unlike in the misattributed paternity situation where disclosure of the finding is likely to come as a surprise to the presumed father, the basic facts surrounding an adoption, donor insemination, or incest will already (presumably) be known to both the husband and the wife. For this reason the risks associated with disclosure of the child's biological status are unlikely to interfere directly with the couple's relationship. The decision whether to disclose could, however, have crucial ramifications for the *parent–child* relationship, as well as for the child's own sense of psychological stability, because in such cases it is the child who has not yet been made aware of the family secret. Thus, once again, in these situations the genetic counselor is faced with a dilemma: Disclosing to the child his or her biological status may upset both the parents (who may arguably have sound reasons for opposing such disclosure) and the child (who may be upset by the revelation). On the other hand, *not* disclosing the information may make the counselor complicit in a deception that could ultimately be viewed by the child as a betrayal should the facts later come to light.

In general, where the child has not yet reached majority, it would generally seem appropriate for the counselor to respect the parents' wishes not reveal the information (although the counselor should advocate strongly that the child be told). Once the child has reached adulthood, however, the situation becomes more complex. In general, if the adult individual (perhaps suspecting that certain information has long been withheld) specifically *asks* whether the genetic test results reveal anything unusual regarding his or her parentage, it would seem ethically required for the counselor to reveal the information, even if doing so may upset the parents. If the adult individual does *not* ask, but if it appears that the information could be highly relevant to his or her own health or reproductive planning, the provider must balance the possible medical risks associated with nondisclosure against the possible psychological risks associated with a disclosure that turns out not to have been desired by anyone in the family (including the individual adult most directly affected) (45). The optimal approach is to try to anticipate such eventualities before the testing takes place, by raising the issue directly during the informed consent process.

Genetic testing performed in the context of clinical medicine may occasionally reveal that babies were switched (whether inadvertently or deliberately) in a hospital nursery. It may also bring to light the finding that sperm samples were mixed up in the course of artificial insemination (e.g., that the husband's sperm was confused with that of an anonymous donor, resulting in the birth of a child whom the parents had mistakenly believed was biologically related to both of them). Or, it may reveal that embryos used in the process of IVF were switched (e.g., where a woman gives birth to a baby whose phenotype suggests strongly that it had been fathered by a man of another race). In these cases any deception or other fault is unlikely to lie with either party to the marriage, but rather with the physician or other hospital personnel who were responsible for the mistake. For this reason the primary considerations underlying the decision whether to disclose will involve not so much potential risks to the mother's well-being as the potential for disruption to the family as a whole. The decision whether to disclose will also have important ramifications for the legal liability of the persons responsible for the original error. There are no easy resolutions in such extremely sensitive cases, and such cases are likely to arise even more frequently in the future as the number of children created through alternative reproductive methods continues to grow.

ACKNOWLEDGMENT

Note: The views expressed are solely those of the author and do not represent the views of NHGRI, NIH, or DHHS.

BIBLIOGRAPHY

1. *State v. Van Guilder*, 271 N.W. 473 (Minn. 1937).
2. *Andrade v. Newhouse*, 128 P.2d 927 (Cal. App. 1942).
3. J.W. Morris and D.W. Giertson, in D.L. Faigman, ed., *Modern Scientific Evidence: The Law and Science of Expert Testimony*, 1997.
4. 1 McCormick on Evidence, 4th ed., §205, at 895 (1992).
5. D.H. Kaye and R. Kanwischer, *Fam. Law. Q.* **22**, 109–116 (1988).
6. D.H. Kaye, *Fam. Law. Q.* **24**, 279–304 (1990).
7. A.Z. Litovsky and K. Schultz, *Jurimetrics* **39**, 79–94 (1998).
8. *Uniform Parentage Act, Uniform Laws Annotated*, vol. 9B, West Publishing, St. Paul, MN, 1973, pp. 287–345.
9. *Uniform Act on Paternity, Uniform Laws Annotated*, vol. 9B, West Publishing, St. Paul, MN, 1960, pp. 347–368.
10. R.H. Walker, ed., *Inclusion Probabilities in Parentage Testing*, 1983.

11. N.M. Vitek, *Disputed Paternity Proceedings*, 5th ed., Matthew Bender, New York, 1997.
12. Omnibus Reconciliation Act of 1993, Pub. L. 103-66, 42 U.S.C. §666(5).
13. L.W. Morgan, *Am. J. Fam. Law.* **12**, 204–211 (1998).
14. *Michael H. v. Gerald D.*, 491 U.S. 110 (1989).
15. *Miscovich v. Miscovich*, 688 A.2d 726 (Pa. 1997).
16. *Uniform Status of Children of Assisted Conception Act, Uniform Laws Annotated*, vol. 9B, West Publishing, St. Paul, MN, 1988, pp. 191–206.
17. *Uniform Putative and Unknown Fathers Act, Uniform Laws Annotated*, vol. 9B, Supp., West Publishing, St. Paul, MN, 1988, pp. 91–105.
18. A.T. Lamport, *Am. J. Law Med.* **14**(1), 109–124 (1988).
19. American Society of Human Genetics, Social Issues Committee, *Am. J. Hum. Genet.* **48**(5), 1009–1010 (1991).
20. M. Coleman, *Cardozo Law Rev.* **17**, 497–530 (1996).
21. In Re Baby M, 537 A.2d 1227 (N.J. 1988).
22. *Johnson v. Calvert*, 851 P.2d 776 (Cal. 1993).
23. Office of Technology Assessment, U.S. Congress, *Infertility: Medical and Social Choices*, OTA-BA-358, U.S. Government Printing Office, Washington, DC, 1988.
24. C.N. LeRay, *BC Law Rev.* **35**, 747–798 (1994).
25. H.B. Jenkins, *Alabama Law Rev.* **50**, 39–61 (1998).
26. *Hart v. Shalala*, Case No. 94-3944 (E.D. La. 1994).
27. E. Garside, *Loy. Law Rev.* **41**, 713–734 (1996).
28. *Syracuse Law Rev.* **36**, 1021–1053 (1985).
29. *Hecht v. Superior Court*, 20 Cal. Rptr. 2d 225 (Cal. App. 1993).
30. *Davis v. Davis*, 842 S.W.2d 588 (Tenn. 1992).
31. National Bioethics Advisory Commission, *Cloning Human Beings: Report and Recommendations of the National Bioethics Advisory Commission*, 1997.
32. N. Elster, *Hofstra Law Rev.* **27**, 533–555 (1999).
33. S. MacIntyre and A. Soosman, *Lancet* **338**, 869–871 (1991).
34. M.-G. LeRoux et al., *Lancet* **340**, 607 (1992).
35. P.J.H. Brock and A.E. Shrimpton, *Lancet* **338**, 1151 (1991).
36. D.C. Wertz, *Health Law J.* **3**, 59 (1995).
37. D.C. Wertz, *Clin. Genet.* **54**, 321–329 (1988).
38. United States President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Screening and Counseling for Genetic Conditions: A Report on the Ethical, Social, and Legal Implications of Genetic Screening, Counseling, and Education Programs*, United States Government Printing Office, Washington, DC, 1983.
39. D.C. Wertz, J.C. Fletcher, and K. Berg, *Guidelines on Ethical Issues in Medical Genetics and the Provision of Genetic Services*, World Health Organization, Hereditary Diseases Programme, Geneva, 1995.
40. B. Biesecker, in M.A. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 108–125.
41. P.R. Reilly, *Suffolk U. Law Rev.* **27**, 1327–1357 (1993).
42. D.C. Wertz, *J. Contemp. Health Law Policy* **13**, 299–346 (1997).
43. Institute of Medicine, Committee on Assessing Genetic Risks, *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994.
44. G.S. Omenn, J.H. Hall, and K.D. Hansen, *Am. J. Med. Genet.* **5**(2), 157–164 (1980).
45. S.M. Suter, *Mich. Law Rev.* **91**(7), 1854–1908 (1993).

See other entries GENETIC INFORMATION, ETHICS, FAMILY ISSUES; GENETIC INFORMATION, ETHICS, PRIVACY AND CONFIDENTIALITY: OVERVIEW; GENETIC INFORMATION, LEGAL, FDA REGULATION OF GENETIC TESTING; GENETIC INFORMATION, LEGAL, GENETIC PRIVACY LAWS; GENETIC INFORMATION, LEGAL, REGULATING GENETIC SERVICES.

GENETIC INFORMATION, ETHICS, ETHICAL ISSUES IN TISSUE BANKING AND HUMAN SUBJECT RESEARCH IN STORED TISSUES

CURTIS NASER
Fairfield University
Fairfield, Connecticut

SHERI ALPERT
University of Notre Dame
Notre Dame, Indiana

OUTLINE

INTRODUCTION

Research using human tissues and cells has contributed immensely to progress in the medical and basic biological sciences. These materials have been essential in developing and testing new drugs and vaccines, investigating infectious diseases, and exploring the mechanisms of virtually all disease processes. The use of human tissues and cells is also the foundation upon which much of the current biotechnological revolution has been based. The project to identify, map, and ultimately sequence the human genome would not be possible without the thousands of human tissue specimens from which DNA is routinely extracted and analyzed. Human tissues are also used for a variety of other medical purposes, including the transplantation of

whole organs (kidney, heart, liver, eye), the transfusion of blood and blood products, and diagnostic testing (blood chemistry, pathological identification of diseased tissues). In addition human tissues may be used in nonmedical contexts, such as the forensic identification of suspects based on trace tissues left at crime scenes, or the identification of the remains of soldiers.

As the collection, storage, and use of human tissues in biomedical research has increased, and the power of the scientific methods to analyze and unlock the secrets held within these tissues has grown, a number of ethical and public policy issues have been identified and debated in the research community. Traditionally excess tissues removed at the time of surgery or in other diagnostic procedures have been viewed as "waste" abandoned by patients and left to the disposal of the hospital or clinical lab. It has been presumed that patients would have no further interests in the disposition of their tissues and that most if not all would be glad to have medical researchers putting them to productive use. However, with recent advances in biotechnology, especially in the area of genetics, suddenly these "waste" tissues have become what one commentator has called a "coded future diary" of the individual and his/her family (1). Intimate knowledge about a person's medical condition (both their current status and potential future status) may be gleaned from the tiniest samples of human tissue.

Since many human tissues are used in research without the knowledge or consent of the persons from whom they are derived (tissue sources), the new-found powers of molecular and genetic analysis raise many difficult questions: What, if anything, should tissue sources be told about the results of research findings? Since information gleaned from tissue specimens, especially genetic information, can adversely affect an individual's employability and insurability, how can sources of research tissues best be protected from these social risks? Who *owns* these tissues and what rights, if any, do tissue sources have to financial gains derived from the use of their tissues in research? And, indeed, is "ownership" even an appropriate concept in this context? How do we protect individuals whose religious faiths or cultural practices impose special restrictions upon the disposition (or burial) of body parts? And, most important, should the informed consent of the tissue source be required for the research use of their tissues, and in what circumstances and how much information should be provided in the process?

HUMAN TISSUES: WHAT THEY ARE, AND HOW THEY ARE USED

Tissues (also referred to as "human biological material") can include everything from organs and parts of organs, cells and tissues (like bone, muscle, connective tissues and skin), to subcellular structures (e.g., DNA) and cell products, blood, gametes (sperm and ova), embryos and fetal tissue, and waste (urine, feces, sweat, hair, nail clippings, epithelial cells, placenta (2,3). Tissue specimens can be stored in many forms depending on both the reason for their collection, and their intended use: in paraffin blocks, slides, formalin fixed, frozen, tissue culture, extracted DNA, or dried blood spots (e.g., on Guthrie cards). Cryogenic storage is generally used for cord blood, gametes, and embryos.

Human tissues are collected and stored (or banked) through a variety of means and for a variety of purposes. They are most commonly collected in conjunction with diagnostic procedures or surgical treatment. For instance, at the time an individual has a surgical procedure to remove diseased tissue, that tissue may be examined to determine whether it contains malignant cells. The diseased tissue is generally preserved and maintained by pathology laboratories for several years (indeed, some tissue archives retain samples that were collected over 100 years ago). This storage of clinically derived specimens may be legally required in some states since they are regarded as part of the patient's clinical record. If questions arise at a later date about the adequacy of the laboratory testing of the specimen, the preserved specimen may be used to confirm (or refute) the original diagnosis.

Additionally excess diagnostic specimens may be used for follow-up clinical care, but most commonly they are used for educational and research purposes, as well as laboratory quality control. For instance, tissues can be used to ensure that equipment in diagnostic and pathological laboratories is functioning properly. Pathology slides may be used for the education of medical students and other specimens may be used to train technicians in testing procedures. Excess specimens are often made available to researchers for a variety of purposes. Fresh specimens may be cultured into immortalized cell lines which provide a perpetual source of DNA and may be used for a variety of other research purposes. Live tissues may also be cultured for use in pharmaceutical research. For instance, the first anti-viral drug for the treatment of AIDS — AZT — was initially demonstrated to be effective by testing it on HIV-infected cell cultures (4,5). Preserved pathological specimens may be used for a variety of research purposes, from studies of enzymes, proteins, and cell physiology to genetic analyses.

Many tissues enter into biomedical research through explicit research protocols. In gene mapping studies, for instance, DNA samples will be collected from entire families suspected of harboring a specific disease gene. Often, immortalized cell lines will be cultured from their blood samples to provide a constant source of DNA without having to return to the family members for further blood samples. Tissue specimens collected for one research purpose may also be used for secondary research investigations, either related to the initial research purpose or for completely unrelated research purposes.

Tissues may be stored in a variety of locations, depending upon the reason for their collection and their ultimate use. The storage facilities include military facilities, forensic DNA banks, government laboratories, diagnostic pathology and cytology laboratories, university- and hospital-based research laboratories, commercial enterprises, and nonprofit organizations (6). Tissue collections can range in size from fewer than 200 specimens to over 92 million, with a conservative estimated total of at least 282 million specimens (from over 176 million

cases) (6). The two largest collections of human tissue in the world (National Pathology Repository and the DNA Specimen Repository for Remains Identification) both reside within the U.S. Armed Forces Institute of Pathology (AFIP), which stores over 94 million specimens (3). An additional 13.5 million specimens are accounted for by newborn screenings, and an estimated 160 million or more specimens can be found throughout the various U.S. graduate medical education teaching institutions (3). These figures do not capture the additional information and tissues that may be found in cancer registries in many states, nor do they contemplate specimens that may be collected as part of the Human Genome Diversity Project.

In all these applications and uses, the tissues themselves can retain varying levels of identifiability — namely the ease with which the identity of the tissue source can be established. The protections afforded human research subjects in the United States are closely tied to how easily the identity of the tissue sources can be discerned. It is thus useful to lay out a general taxonomy of tissue identifiability:

- *Identified*. The tissue source is known and the individual's identity is tied to the sample. (This would be the case with specimens being analyzed for diagnostic and treatment purposes.)
- *Identifiable*. The tissue source is tied to the specimen through the use of a link (e.g., a code number), but the identity of the source is not directly known without tracing the link. Many pathology specimens, for instance, are archived according to a pathology record number.
- *Anonymized*. The tissue source's identity is irrevocably unlinked from the specimen, so that the individual's identity cannot be discerned (i.e., the tissue is unidentifiable).
- *Anonymous*. The tissue source's identity is never known, since the specimen is collected with no identifiers at all (i.e., the sample is unidentified).

It should be noted that as biomedical research progresses, becoming increasingly sophisticated in its ability to tease apart the molecular components of tissues and cells, the notion of having an "anonymous" or even "anonymized" sample will likely diminish, if not disappear altogether. In the future, determining the sequence of a tissue sample's (and tissue source's) DNA could become a routine procedure. If this were to happen in the clinical context, which is quite likely given the different analytical systems now under development (7), this information will invariably end up in a patient's medical record. Since these records are increasingly being stored and processed electronically, it could require only a small endeavor to match sequences from a medical record against the sequences from stored tissues. When that happens, the protections to safeguard tissue source identity currently in place, discussed in the next section, could become wholly inadequate to the task.

CURRENT FEDERAL REGULATIONS FOR PROTECTING HUMAN SUBJECTS IN BIOMEDICAL RESEARCH

The U.S. federal regulations (8,9) governing the use of human subjects in biomedical research establish three primary requirements intended to protect human subjects. First and foremost is the requirement for informed consent, which includes among other things a statement of the purpose of the research and its probable risks and benefits to the subject. Second is the requirement that all research involving human subjects be reviewed by an Institutional Review Board, (IRB) that is composed of other scientists and physicians, at least one nonscientist, and at least one member not affiliated with the institution. This committee is charged with evaluating the adequacy of informed consent, establishing that the research poses a favorable risk benefit ratio to the subject, and ensuring that the research design will yield useful results.

The third requirement is that institutions conducting federally funded research file an "assurance" with the Department of Health and Human Services (DHHS) Office for the Protection from Research Risks (OPRR), which administers and enforces the regulations. This assurance contains a statement of the ethical principles to be followed in conducting the research as well as the constitution of the IRB and its operating procedures. Institutions which conduct multiple research protocols may apply for a "multiple project assurance" which permits a single IRB to review each protocol. Institutions holding an MPA are required to ensure that *all* research involving human subjects conducted by their employees and affiliates complies with the regulations, including projects not funded by the federal government. Thus a wide array of research is subject to these regulations. There are approximately 450 institutions in the United States currently holding an MPA (10). Institutions which do not regularly conduct federally funded research may not be covered by these regulations. Privately funded research institutions, such as pharmaceutical firms, may engage in research involving human tissues outside of the regulatory context, including review by an IRB and requirements for informed consent. The Food and Drug Administration maintained until 1991 separate regulations governing research involving human subjects that was undertaken by drug and medical device manufacturers. In 1991, regulations governing research on human subjects conducted by the many agencies of the Federal government were consolidated into what is known as the Common Rule. However, the FDA maintains a separate system of enforcement (largely by audit) and its regulations do not cover the use of human tissues or medical records, but apply only to research conducted under the auspices of an Investigational New Drug application (IND) or an Investigational Device Exemption (IDE) as part of the FDA approval process prior to marketing. Thus, basic research conducted by private corporations and research labs prior to an application to the FDA for marketing approval is not regulated, unless these labs are otherwise regulated through their receipt of Federal funds.

The current regulations and the IRB system of peer review are recent phenomena that grew out of revelations

in the 1960s and 1970s of unethical research practices, including the infamous Tuskegee syphilis study. These abuses led to the first formal regulations issued in 1974 by the U.S. Department of Health, Education, and Welfare (HEW) (11). That same year, Congress passed the National Research Act, which established a National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, which was charged with reviewing all aspects of the involvement of human subjects in research and making recommendations to improve the system of protections. The National Commission issued a series of reports, including the *Belmont Report* (12) which set out the broad ethical guidelines that research involving human subjects should follow. The Commission also issued specific recommendations regarding the IRB review process. HEW was required by law to codify these recommendations into its regulations. Proposed regulations were published in 1979 by HEW (13) and after public comment, were finally adopted in January 1981 by what had then become DHHS (14).

The research use of tissues (and medical records) was explicitly considered by the National Commission and as part of its recommendations to HEW, the Commission suggested that these research projects need not obtain informed consent provided that "the importance of the research justifies the invasion of privacy" (15). The National Commission recognized that the use of human tissues and medical records constituted an "invasion of privacy," but argued that if adequate safeguards to protect individual's confidentiality were provided, this research was of minimal or no risk to subjects and could proceed in the absence of informed consent. Nevertheless, the Commission's recommendation would have required IRB review of all such research protocols. The National Commission also recommended that institutions that anticipate using tissues and records in research should in lieu of informed consent, notify all patients and provide a mechanism (i.e., a blanket consent) by which they could opt out, if they so desired.

However, in codifying these recommendations, HEW opted not to follow completely the National Commission's recommendations. Concerned about IRB workloads and convinced that much of this research was of minimal or no risk, HEW instead proposed two exemptions for such research from the IRB regulations. In addition, in these cases, HEW provided that IRB's could waive or modify informed consent.

The first exemption from the requirements of IRB review and informed consent is implied by the definition of "human subject" and the statement [at 45 CFR 46.102(f)] that the regulations apply only to research involving *human subjects*. It is possible that some research protocols can claim that their use of human tissues, under narrowly defined circumstances, do not involve human subjects as defined, and therefore the regulations do not apply. The second exemption is an explicit exemption governing the use of existing specimens and records, provided only that investigators do not record any patient identifiers or links to identifiers. Protocols that meet the requirements of either exemption may proceed in the absence of IRB review and in the absence of any informed consent, unless state laws or institutional policies provide otherwise.

For protocols that do not meet the requirements for either exemption, it is possible that the IRB nonetheless may permit a waiver or modification of informed consent. There are four conditions to be met before an IRB may waive or modify consent. Because of their ambiguity and because they point to important ethical issues, the discussion of these conditions below will extend to some of the important ethical and policy questions surrounding the use of human tissues in research, including the recontact of tissue sources with research findings, the ownership and commercial use of tissues, as well as the potential risks to which new biotechnologies may give rise.

Research protocols involving the use of human tissue that do not meet any of the above conditions for exemption or waiver of consent are required to obtain the informed consent of the subjects. Many tissue protocols may follow standard informed consent requirements used at the investigator's institution, but some protocols, especially those involved in genetics research, may require special considerations.

Exemption Implied by the Definition of "Human Subject"

This first exemption from IRB review arises out of the regulation's definition of "human subject." If a research protocol can demonstrate that it does not involve human subjects, then the regulations do not apply. The regulations define "human subject" as a "living individual about whom an investigator (whether professional or student) conducting research obtains" either of the following: (*1*) data through intervention or interaction with the individual, or (*2*) identifiable private information (§46.102(f)).

Interventions and Interactions. Criterion 1 is an important consideration because obtaining human tissues may often require some intervention or interaction with a person, usually an invasive procedure such as a needle stick to draw blood, or a tissue biopsy. Since these procedures may involve pain or risk harm to subjects, the regulations require that the interventions be subject to the oversight of an IRB, though it is possible that the IRB may waive the requirement for informed consent (see below).

One circumstance that easily gives rise to confusion in the application of this criterion is when tissues are removed from a patient for clinical reasons. If the attending physician is also a researcher who hopes to use some of the tissues for research purposes, and if the diagnostic procedure is genuinely clinically indicated, then the physician might assume that there is no interaction or intervention which would not have taken place in the absence of the research use of the tissue. However, the regulations do not distinguish between such dual roles of clinician and investigator. If the clinician is also an investigator, then, by definition, there is interaction, if not intervention, by the clinician qua investigator — whether or not the procedure is ordered solely for clinical purposes. The duality of roles in such cases presents the possibility that individuals may be subjected to research procedures under the guise of clinical care, or at the very least that a clinician's judgment may be influenced by his or her research interests, and thus would modify the patient's

care accordingly. Such modifications of patient care for research purposes, however, present the possibility of further risks to these subjects, and as such, the regulations require at least the oversight of an IRB to evaluate those risks, and to advise the investigator on appropriate procedures and safeguards.

Identifiable Private Information. Many protocols that seek to collect, store, and experiment on human tissues use excess tissues from clinical procedures with which the investigators are not associated. These tissues may be gathered from blood banks, blood chemistry labs, pathology labs, cytogenetics labs, or any other clinical diagnostic laboratory. Since more diagnostic specimens are routinely collected than are strictly necessary to perform the clinical tests and analyses, there is often excess tissue which may be made available to investigators. In these cases, research protocols can clearly meet the first criterion in the definition of human subject, namely that the protocol does not involve any intervention or interaction with an individual.

In order to meet the second criterion of the exemption, however, the investigators must not obtain any "identifiable private information." OPRR has interpreted this phrase to mean that if investigators have access, *however briefly*, to individual identifiers (such as names or social security numbers) or to links to identifiers (such as a medical record number or pathology record number or other such code) then the research involves human subjects as defined and the protocol would fail this exemption (16).

While some research protocols involving human tissue may be able to use anonymized tissue specimens, more powerful research can be conducted on identified or linked specimens. By maintaining a link back to the originating laboratory and/or the medical record, it is possible to update a tissue bank's database with subsequent morbidity and mortality data from the sources of the tissue specimens. Several tissue banks and tissue procurement services exist (17) or have been proposed in the literature (18) that are designed to maintain a one-way flow of information from medical record to tissue bank and on to individual investigators, while ensuring through the use of codes that no research information finds its way back to individual patients or subjects. Since some research, especially in genetics, can generate sensitive information about an individual, it is important to maintain a clear distinction between research results and clinical information in order to protect the source from the harmful consequences of research as well as from the uncertainty of results that have not been further validated. While such one-way tissue banks make very powerful research tools, they would fail to meet the exemption implied by the definition of "human subject," since they must maintain linking codes back to the originating laboratories and ultimately back to patient identifiers.

Exemption for Existing Tissues and Medical Records

The second exemption for research involving human tissues is in Section 46.101(b) of the regulations and it lists a number of specific exempt categories of research. Most of these categories pertain to educational, social or psychological research which involve the use of surveys and questionnaires. However, §46.101(b)(4) explicitly addresses the use of tissues and medical records in research. It exempts:

> [r]esearch involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

The key determination in applying this exemption is that the research materials must be *existing*. OPRR has clarified the interpretation of "existing" to mean that the materials (specimens, medical records, etc.) must be already existing *at the time the research is proposed* (19). Thus this exemption only applies to tissues and records that have already been produced and are being stored when the protocol solicits exemption status from the IRB. Such protocols are typically referred to as *retrospective*. Although the historical record of the promulgation of these regulations does not discuss the reasoning behind this limitation, it can be easily seen that it serves two purposes. First, because the specimens must already be in existence and therefore will already have been removed from individual patients, it would be impossible for investigators to influence clinicians — diagnostic technicians, surgeons, and so forth — to alter their procedures to obtain more or different kinds of tissue than they would otherwise remove in the course of clinical care and diagnosis.

Second, it would be much more difficult and in many cases impossible to recontact former patients to inform them of the intentions of the investigators to use their tissue specimens in research and to ask their consent. Many patients will have moved or died and recontact could only be made at great expense; expense that would be prohibitive for many research protocols. It is also possible that sampling bias might result if investigators were limited to only those tissues for which consent could be obtained.

It should be noted that this exemption does permit investigator access to patient identifiers, though it requires that investigators *record* no identifiers or links to identifiers. The purpose of this limitation is the same as that expressed in the limitation under the definition of "human subject" that investigators obtain no "identifiable private information." Because the exemption for existing specimens is more liberal in permitting investigators access to identifiers, the first exemption based upon the definition of "human subject" applies primarily to *prospective* protocols, that is, protocols that seek to collect specimens produced or procured after the protocol is proposed to the IRB.

While tissue banking and collection protocols that meet either of these exemptions are therefore exempt from the federal requirement of informed consent, the regulations do not preempt state laws. While some states laws defer to the federal regulations in matters concerning the protection of research subjects, other states may have independent laws and regulations governing human

subjects research. Before investigators proceed on the basis of exemption from the regulations they, or their IRBs, would be prudent to confirm their compliance with local law.

Who Decides What is Exempt? One problem in applying the exemptions that is not addressed in the regulations is deciding *who* determines what is exempt in the first place. Applying the exemption categories can be complex and confusing. Investigators may easily misinterpret them and fail to submit to the IRB research that in fact would not be exempt and may in fact even require informed consent. The MPAs at some institutions may require that the IRB or IRB chair make this determination. In other cases, it may be institutional policy that only the IRB may make exemption determinations. This is the safest policy, provides for a uniform application of the regulations, and limits the possibility that institutional funding could be jeopardized by the failure of an investigator to properly interpret the regulations.

Conditions Necessary for the Waiver of Informed Consent

For protocols that do not meet the criteria for exemption from the regulations, as described above, the possibility remains that the IRB could waive or modify the requirements for informed consent, as provided for in §46.116(d). In order to do so, the IRB must find and document that a protocol meets all four of the following conditions:

1. The research involves no more than minimal risk to the subjects.
2. The waiver or alteration will not adversely affect the rights and welfare of the subjects.
3. The research could not practicably be carried out without the waiver or alteration.
4. Whenever appropriate, the subjects will be provided with additional pertinent information after participation.

The discussion that follows will look at the ethical issues that arise in the application of each of these four conditions, since many of the issues that arise in tissue research will do so within the context of one or more of them. The more general ethical question of whether informed consent should or should not be required for these protocols (and this includes the exemptions) will be taken up later in this article.

Minimal Risk. Minimal risk is defined in the regulations at §46.102(i):

> Minimal risk means that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.

This definition presents two pairs of issues that need to be clarified. First, the scope of the term "harm" needs to be established, and second, the distinction between the *probability* and the *magnitude* of harm that may result from participation in a research protocol needs to be clarified.

Physical Harms. Many research protocols involving the collection, use, and storage of human tissues involve no contact with the tissue sources, the tissues being procured from excess diagnostic or pathological specimens which have been generated for clinical (or other) purposes. In such protocols, there is virtually no risk of physical harm to these sources that may result from the research use of their tissues. Thus the use of genuinely excess tissue specimens, procured by whatever means, is generally of minimal *physical* risk.

One exception to this may arise when the investigators explicitly solicit tissues from the clinicians who order the diagnostic or surgical procedures, or when these clinicians are the investigators themselves. In such cases, it is possible that clinical procedures may be altered to accommodate the research need for more tissue. Whether or not the extraction of additional tissue constitutes a greater than minimal risk to the subject depends on the kind of procedure the patient will undergo. If it is simply a matter of drawing an additional vial of blood, then the additional physical risk is minimal if not non-existent for most subjects. The same may be true of procuring additional bone marrow during a biopsy procedure. But risks may be heightened, for instance, if additional spinal fluid is procured during a spinal tap or additional liver tissue is procured during a liver biopsy. The risks of taking additional tissue will depend on the specific circumstances of the procedures being used. See Holder and Levine (20), who argue that in many cases, small amounts of additional tissues taken during surgery for research are of minimal or no risk and do not require informed consent. There are no hard and fast guidelines in this area and ultimately IRBs must make their own judgment of whether the procurement of additional tissue for research purposes during a clinical procedure is of minimal risk or not.

It is important to note, however, that the risks that result from the procurement of additional tissue for research purposes do not include the risks of the procedure itself, which is being conducted for clinical purposes. Thus the risks of infection from a spinal tap procedure are not a result of the research use of the additional spinal fluid taken, but result from the procedure itself which is ordered for clinical purposes. Likewise the risks of surgery are not a *research risk* if the surgery is undertaken for clinical purposes, while a small amount of additional tissue is taken for research. On the other hand, if an invasive diagnostic procedure is undertaken strictly for research purposes, then the full risks of the procedure must be considered in determining whether it is of minimal risk. A blood draw, for instance, is a routine diagnostic procedure and may be judged to be of minimal risk. A bone marrow biopsy, however, entails greater risks of infection and greater possibility of pain and discomfort and therefore may not be of minimal physical risk.

Psychosocial Harms. In addition to the physical risks of harm and discomfort, tissue sources may be at risk of social, psychological, economic, and legal harms if research results and any medical record information collected from the tissue source are not kept confidential. Such risks

will depend on the specific experiments to be conducted upon the tissue specimens. Analyses of protein structure may have little impact on one's self-image, insurability, or employability, but some genetics research results can have a devastating impact on the source if the results were to find their way back to the individual's medical record or into the hands of insurers and employers. Although the definition of minimal risk adopted in the current regulations does not clearly identify these risks as relevant considerations, the National Commission in the *Belmont Report* clearly indicated that IRBs should be concerned with these possible consequences:

> Many kinds of possible harms and benefits need be taken into account. There are, for example, risks of psychological harm, physical harm, legal harm, social harm and economic harm and the corresponding benefits. While the most likely types of harms to research subjects are those of psychological or physical pain or injury, other possible kinds should not be overlooked (12, p. 15).

A number of commentators have argued that, depending upon the specific circumstances, genetic information developed in research protocols may entail more than minimal risks to subjects (21–24). Others have argued that the excellent record of researchers in the United States in maintaining the confidentiality of research data is a sign that the risks of a breach of confidentiality are so low that we may judge all research involving tissues to be of minimal risk (25–28). Part of this judgment will depend on what, if any, information is communicated back to the tissue sources. If the sources are identified and research results either are or can be communicated back to them, then clearly these persons may be at risk of learning things about themselves which they never consented to finding out. If sources are not identified, or there are no intentions to provide sources with research results, then the risks that these results may have an adverse impact on them are greatly minimized.

However, despite the intentions of investigators not to inform sources of research findings, there may be occasions when research records are used in public health or even criminal investigations. If codes are maintained that would link to individual identifiers, it is possible that individuals will become involved in such activities as a result of the use of their tissue specimens or medical records. For instance, tissue specimens from the Navaho Health and Nutrition Survey maintained by the Centers for Disease Control were used in the investigation of the hanta virus outbreak in the four corners region of the southwest, even though the sources of these tissues never consented to such use (29). Public health investigations are not governed by these regulations and because they often involve infectious diseases and seek to minimize imminent risks to the public, they may override the usual considerations for the protection of human subjects. The same justifications can be made in criminal investigations, though for certain classes of research, especially research involving illicit drug use, psychiatric problems and even genetics, investigators may obtain a "certificate of confidentiality" from DHHS (30,31) which protects research data and materials from most court ordered subpoenas. Certificates of Confidentiality are issued under the Public Health Service Act §301(d), 42 U.S.C. §241(d). Categories of research information for which such certificates are issued include what "would normally be recorded in a patient's medical record, if the disclosure ... could reasonably lead to social stigmitization or discrimination," "information ... damaging to an individuals financial standing, employability, or reputation," and "genetic information."

Distinguishing between Probability and Magnitude of Harm. Whether the risks are physical, psychological, or social, it is essential to clearly distinguish the difference between the *probability* of harm and the *magnitude* of harm that may result from a subject's participation in a research protocol (12, p. 15). Some harms may be of very low probability. For instance, the record of researchers maintaining the confidentiality of private information is excellent. Thus, in general, the probability of a breach of confidentiality may be quite low in many research projects, especially when investigators take steps to protect confidentiality by the use of codes and locked file cabinets. However, the magnitude of the harm such a breach would cause depends on the sensitivity of the information itself and the context in which the breach occurs.

For instance, disclosure of the histological tissue type of an identified specimen may be of little consequence to the tissue source, since this information has little or no clinical or social relevance, except in organ transplantation. However, disclosure that a given identified tissue contains the genetic mutation for Huntington's disease may have a profound impact upon the source, both psychologically as well as socially and economically. Such an individual may find it difficult or impossible to obtain health insurance, or even find a job. While the risk of a breach of confidentiality may be low, the magnitude of social, psychological, or economic harms that may result if confidentiality is breached may be quite high. Again, the definition of minimal risk does not provide a great deal of guidance on how IRBs are to weigh the risks and magnitudes of harms and the current literature is divided on this question as well. Under the current regulatory framework it remains to the individual IRBs to make this judgment as best they can.

Rights and Welfare. The second criterion to be met is that the waiver or modification of informed consent does "not adversely affect the rights and welfare of the subjects." This has been a difficult phrase to interpret. The phrase "rights and welfare" has served as a catch-all idea at least from the earliest days of federal involvement in the protection of human subjects of biomedical research (32,33). It is hard for investigators and IRB members alike to know precisely what is meant by this wording (34). On the one hand, if informed consent is a "right" which the regulations bestow (or recognize), then clearly the waiver of informed consent violates such a right (35). On this interpretation, requiring that the waiver not violate a subject's rights would be self-contradictory. Furthermore, since "minimal risk" is a

requirement for the waiver of informed consent, it is difficult to understand how this waiver could be construed as adversely affecting the welfare of a subject, since a protocol would have to already minimize any such adverse effects.

Nevertheless, it is necessary for IRBs to interpret this clause and document that the waiver of consent does "not adversely affect the rights and welfare of the subjects." Since the promulgation of these regulations in 1981, biological science and technology have advanced at an ever-increasing pace, and issues which at that time were perhaps just beyond the horizon have come to the fore with increasing frequency. In particular, two issues that may fall within the scope of "rights and welfare" have been debated and are relevant to the waiver of informed consent for the research involving human tissues: (*1*) the ownership and/or commercial exploitation of human tissues and cells and the patenting of gene sequences; (*2*) the adverse impact some genetic research may have upon ethnic, racial, or other groups.

Ownership and Commercial Use of Human Biological Specimens. A potentially contentious issue is that of "ownership" of tissues, for it may juxtapose the rights and welfare of tissue sources against the interests of researchers in freely pursuing scientific knowledge. From this pursuit of scientific knowledge comes most of the breakthroughs that allow new drugs and therapies to be developed, and any concomitant financial rewards. Part of this tension arises because the issues of "ownership" are enmeshed in the language of "property rights," and the attendant legal lexicon. Additionally "ownership" implies that all interests associated with the tissues can be couched in economic terms.

Probably the most famous court case involving these issues is that of John Moore. In its 1990 decision in *Moore v. Regents of Univ. of Cal* (36), the California Supreme Court ruled that Mr. Moore did not have property rights in his removed tissues/cells (which had been transformed into a profitable product by his physician, a colleague of the physician, and a pharmaceutical company, without his knowledge or consent). However, the court also ruled that his physician had breached his fiduciary duty to Mr. Moore by not disclosing his financial interest in treating and extracting tissue specimens from Mr. Moore. This case has become the touchstone for how many institutions deal with issues of tissue "ownership," even though, as legal precedent, it applies only in California.

The Moore case does not exhaust the possible ways to address "ownership" of tissues. Indeed, there are at least four different ways in which to view these issues:

1. Tissue sources have no "ownership rights" in their tissues; researchers do (e.g., the Moore paradigm).
2. Tissue sources share "ownership rights;" the question becomes how best to compensate them and when.
3. Tissue sources do not have "ownership rights" but researchers owe them recompense when their tissues become profitable.
4. "Ownership" is not an appropriate construct in this context, either for researchers or sources.

1. *Tissue sources have no "ownership rights," researchers do*. On this view, tissue sources are presumed to either abandon tissues (if provided in a clinical encounter) or donate their tissues for research use. Consent forms for clinical encounters (e.g., where surgery is involved) may have merely stated that removed tissues would be disposed of by the institution. Most patients would probably interpret this to mean that their excess tissues would be thrown out, even though this has generally not been the case. "In many — perhaps most — cases, individuals were not aware that their specimens were being stored or had no knowledge that they might be used for various research purposes by a number of investigators" (6, p. 41). Furthermore, according to a 1995 study of informed consent forms for genetics research, of the 23 documents reviewed, 4 explicitly mentioned that the investigator or institution were the sole owners of any tissue samples or transformed cell lines. The other 19 were silent on the issues of ownership or recompense to tissue sources in the case of profits realized (37).

In the language of John Locke, the sort of property right in tissues that would be claimed by researchers would be a natural right claim — namely that they have a right to any amount of wealth the fruits of their labor have produced. On this moral theory, whenever researchers come into possession of tissues they presume are abandoned, which they then modify in a way that renders the materials commercially valuable, the right to the financial rewards inheres to the researcher — the tissue source does not share in the profits, for s/he did not contribute to the labor that produced the value from the tissue.

2. *Tissue sources share "ownership rights," and the question is one of compensation*. According to this view, tissue sources and researchers share the profits realized from the transformation of tissues. There would be several challenges in trying to implement this sort of model, many of them logistical in nature. First, there is the problem of identifying at the outset which tissues might, through transformation, produce a marketable product. For instance, the case of John Moore is highly exceptional in that his doctor saw the commercial potential in the unique characteristics of his cells at the outset of his clinical treatment. This has rarely been the case, not only in terms of the early recognition of potential value, but also in the exclusive existence of the cells within one person. In most cases, any individual's "raw" tissues are of low economic value (leaving aside any questions of organ donation).

Even if one were able to predict at the outset which extracted tissues could produce value, the next challenge would come in instituting a scheme that could compensate the tissue sources. The issues here include keeping track of individuals whose tissues eventually do lead to a valuable commodity, often years after the tissues are obtained, which could eventually cost more than the actual compensation realized from any commercial product made from those tissues. In addition it is infrequent that any *one* person's tissues will lead to a viable product. That means that an additional logistical hurdle would have to be overcome, that is, calculating proportionality. In other words, researchers would have to determine what proportion of each person's tissues led to the product

(assuming there was a desire to achieve compensation based on the value of each contribution of the unique tissues), and distribute compensation accordingly. This presumes that researchers would have a methodology to keep track of whose tissues were transformed in what ways, and that this "inventory" carried forward to other researchers who might also get the tissues for related or unrelated research. Finally, most profitable discoveries are made on the basis of a variety of tissues, and the proportional compensation of each tissue source would create a logistical ordeal to administer.

In a different context, a similar issue has led some scholars to argue that individuals should have property rights in the personal information about them that is bought and sold. The sale of personal information forms the basis of and supports a multi-billion dollar industry. In this context, these scholars propose forming a clearinghouse (similar to ASCAP and BMI in the music industry) that would receive royalties on behalf of individuals whose personal information is bartered (38,39). While their schemes have not been adopted by the information industries, the idea is still one worth considering for compensating tissue sources for the contribution their tissues make to science.

3. *Tissue sources do not have "ownership rights" but they are owed recompense when their tissues become profitable*. Because the body and its parts are generally not held by most people to be commodities, some have suggested that as an alternative to "ownership" in their tissues, when tissues do form the basis of a profitable product, some recompense should be provided to those individuals. As Thomas Murray has stated,

> There is something very special about human organs and tissues, even when removed from the body. We do retain moral interests in them, so that at the least they are not misused or treated in an undignified manner. And we have certainly recognized that body parts, whatever their dignity, can also have a price. But, on balance, we have rejected the idea that they should be bartered on the market (40).

Researchers finding a way to compensate tissue sources for the contribution their tissues have made to a profitable product would demonstrate respect for the people from whom the tissues came, and an appreciation for their contribution. While it may still not be easy to determine precisely whose tissues led to the profitable product, the issue does not have to be determined definitively, since "ownership" by tissue sources is not involved. For instance, for groups of related individuals (see the discussion that immediately follows), products that are derived from their tissues can be fairly easily compensated, since the group's collective contribution makes the product possible. Consider some of the new genetic tests that are now available. Research done on tissues collected from Ashkenazi Jews looking for a genetic basis for some types of breast cancers led not only to the discovery of BRCA1 and BRCA2 but also led to the marketing of genetic tests for the mutation. One possible way for researchers and pharmaceutical companies marketing these genetic tests to show their respect for the research subjects would be to donate some of their profits to Jewish organizations or synagogues that serve the communities from which the research subjects came. Alternatively, the groups that would most likely be tested (if they formed the pool from which researched tissues came) could receive a discounted price on the testing. In the former case, the compensation is spread over more of the community, while in the latter, it is concentrated on those potentially most likely to directly "benefit" from the testing.

4. *"Ownership" is not an appropriate construct in this context, either for researchers or sources*. Finally, there is a case to be made for the outright rejection of the concept of ownership in this context. It may be that it is more productive to think of the holding of tissues for research purposes as a "custodial" relationship between the tissue source and the researcher. Certainly the legal discourse surrounding property "rights" could be abandoned. Moreover this approach can also accommodate more readily the noneconomic interests people have in their tissues. Some have suggested that an alternative is an independent trust model, where a disinterested nonprofit organization could hold cell lines (or other tissues) in a custodial arrangement, and grant licenses for use of the cell lines or other tissues to researchers and others (41). The licensing agreements could conceivably contain conditions under which tissue sources do not want their tissues used (e.g., genetic research). Any such alternative model will require great care and thought to ensure that the shortcomings of the current "property paradigm" are actually accommodated.

There is little regulatory guidance for researchers and IRBs alike in this controversial area, and it thus remains to IRBs to individually determine how they will handle questions of ownership and the commercial exploitation of human tissues in biomedical research. Some tissue banking protocols, such as the Cooperative Human Tissue Network (17), have stipulated that their tissues may not be used for commercial purposes nor patented by the researchers who use them. But it must also be recognized that for better or worse, private enterprise is increasingly becoming the economic engine that drives scientific progress. How these economic interests fit into the other ethical considerations — particularly the presumed altruism which has historically justified researcher access to tissues and records — which must be considered when weighing the waiver of informed consent remains an open question.

Risks to Ethnic and Racial Groups. The rights and welfare of groups of related individuals are becoming more of a concern in the context of biomedical (and, for now, genetic) research. The concept of "groups" used here is identical to that which was used in a paper written for the National Bioethics Advisory Commission on privacy issues in analyzing tissues (42):

> . . .a collective of individuals who are culturally or ethnically related, where shared genetic characteristics are either likely or possible (or perhaps simply inferred). "Groups" can usually be characterized by a demographic label; e.g., African-American; Pacific Islander; American Indian; Scandinavian American; Ashkenazi Jew; etc. This notion of "group" does not necessarily extend to nuclear families as the unit of analysis, although certainly nuclear families may be members of larger

> cultural/ethnic groups. Culturally or ethnically related people are, for the *present time*, the most easily recognizable as members of particular groups, within social contexts, and therefore, potentially the most readily stigmatized by genetic characteristics predominantly associated with that group. The distinction amongst different types of groups may become less relevant in the future, as we accumulate more knowledge about the genetic makeup of the entire population and all its constituent groups.

With the increase in molecular genetic research, groups of genetically related individuals are increasingly becoming a desirable "unit of analysis," particularly where there is thought to be a "group" component to the genetic trait — namely that the trait is more prevalent in certain genetically-related individuals. That members of these groups might have concerns about how their group is understood and characterized should not be too surprising, since much of one's self-identity comes from their interactions with others like them (i.e., other members of their group). This is particularly the case where some groups historically have been the targets of discrimination and stigmatization. Many individuals find that their associations with groups make important contributions to their self-development, self-discovery, and even their self-image. The often mutually supportive nature of groups and collectivities plays a key role in making these contributions. This may be even more the case in groups in which the members have an ethnic, racial, or cultural commonality. In other words, group identification, particularly in these latter cases, can be as important to the development of an individual's self-definition and self-respect as it is to the group's self-definition and continuity. When an encroachment on an individual (as a member of a group) or on the group itself occurs, the violation may be felt as being an affront to both the individual and the group (42). As Larry Gostin writes:

> Derogatory information associated with a group can result in real harms such as discrimination against members of the group in employment, housing, or insurance. Derogatory information can also cause intangible hurt to groups such as lowering their self-esteem or racial or cultural pride. Derogatory information about a sub-population can stigmatize and wound its people as much as breaches of confidentiality can affect an individual. The information collected from groups, just as information about individuals, need not be blatant or intentional to cause harm or hurt. Even the best intentioned and careful research can trigger concerns about privacy (43).

As discussed above, the federal regulations exempt both the retrospective and prospective collection of some tissues and records from IRB review, provided investigators meet the specific requirements for access to identifiers. From the standpoint of public policy, these exemptions from the regulations may need to be rethought, where the tissues are known to originate *within a particular group*. The current policies make the presumption that *individual* identity is the only form of identity that is relevant to the research being conducted — and to ensure the protection of human subjects. Of particular concern is the possible impact upon various ethnic and racial groups of genes associated with personality, behavior and intelligence, though other less socially charged genetic traits may give rise to economic discrimination based on ethnicity or race. Such concerns have led several commentators to argue that the current exemptions from IRB review should be abandoned (23,42,44,45) in order that there be some "independent, social mechanism to ensure that research is ethically acceptable and that the rights and welfare of subjects will be protected" (13, p. 47692).

Where a group or identifiable community is the "unit of analysis" for the genetic research, researchers should involve members of the affected community in the research process, from recruitment through to the publication of results. The Human Genome Diversity Project (HGDP) provides much guidance in this regard. Henry Greely states that:

> Research inevitably provides information about a group, as well as the individuals who constitute it. The group . . . is really the research subject. It is the group's collective autonomy that is challenged if researchers, with the informed consent of only a few individuals in the group, can probe for information about the whole group (46).

Indeed, it is partly for this reason that the model protocol for HGDP requires that researchers obtain the informed consent of the population, "through its culturally appropriate authorities where such authorities exist" (47) prior to sampling. Furthermore, if the population's authorities choose not to participate, HGDP would not accept any samples from any member of that population. "We believe . . . that the population-based nature of this research requires population-based consent, and we will insist on it" (47, p. 1444).

Within the United States, finding the "culturally appropriate authority" can be a difficult, if not impossible task for many groups. While it may be possible to find people who can facilitate discussions within the community (e.g., religious leaders), many groups are simply too populous and dispersed (e.g., Scandinavian-Americans or African-Americans) to have an authority with the power to make decisions for the entire group. In these cases, it is still important to hold frank discussions within the community to facilitate trust in the process (42).

Moreover, even though the regulations to protect human subjects address only the protection of individuals and not the protection of identifiable groups, when researchers propose additional retrospective research on tissues belonging to such groups, they should utilize the full IRB process (e.g., not expedited review) to justify that additional research. Part of the justification should be an indication of how the researcher can mitigate the harm that can be caused by the information obtained as a result of the research. In other words, even though the regulations treat these tissues as anonymous, researchers and IRBs would be prudent to do more than the regulations require in this case, and treat protocols using these tissues as using identified samples.

"Practicality" of Research without the Waiver. Of the four conditions this requirement is the easiest to interpret, though it may be the condition upon which most protocols

founder. The waiver or modification of informed consent cannot be granted unless it can be demonstrated that "the research could not practicably be carried out without the waiver or alteration." The key term to interpret is "practicably." While the regulations provide no further guidance, this criterion was explicitly commented upon by the National Commission in the *Belmont Report*:

> In all cases of research involving incomplete disclosure, such research is justified only if it is clear that (1) incomplete disclosure is truly necessary to accomplish the goals of the research ... Care should be taken to distinguish cases in which disclosure would destroy or invalidate the research from cases in which disclosure would simply inconvenience the investigator (12, p. 12).

The primary concern in considering the "practicality" of the research without the waiver is the *scientific validity* of the study. The National Commission had in mind primarily social and psychological research which required some element of deception or incomplete disclosure, and the Commission expressed its concern that such alterations and deceptions not be taken lightly, but be justified by scientific necessity. This concern was underscored in the Commission's recommendations to HEW in 1978:

> *Nondisclosure must be essential to the methodological soundness of the research* and must be justified by the importance or scientific merit of the research (48).

The Commission was also clear to distinguish such methodological issues from the question of whether obtaining informed consent was *inconvenient* to the investigator (see also Ref. 35, p. 62). Accordingly, then, if we are to base our interpretation of this clause on the National Commission's own reflections, "practicality" refers to scientific necessity and not to the extra work an informed consent requirement might entail.

Although the National Commission had in mind primarily deception and incomplete disclosure in social and psychological research in its comments on this criterion, their application to tissue collection, use, and storage protocols is relatively straightforward and is best analyzed in light of those categories of protocols that fail to meet the exemption criteria discussed above: (*1*) retrospective and identified and (*2*) prospective and identified. Following this discussion, we will (*3*) analyze the role of prior "blanket consent" in the waiver of consent for specific protocols.

Retrospective and Identified. Many studies involving tissue specimens and/or medical records benefit by the inclusion of identifiers, or indeed may methodologically require such identifiers. Such studies are not therefore candidates for exemption under §46.101(b)(4). Especially in epidemiological studies based on medical record reviews, a number of commentators have argued that waivers of consent are methodologically necessary to the statistical validity of the studies (27,49). The same considerations arise in some epidemiological research involving the collection of tissue samples that already exist (50). Such protocols can involve thousands of tissues and medical records, making the attempt to obtain consent not only exceedingly expensive but impossible in many cases where the patients have moved and no forwarding address is available (51). The impossibility of obtaining consent in these cases undermines the validity of the study by introducing selection bias in the data.

On the other hand, some studies may involve smaller numbers of tissues, and in other cases the issue of selection bias will not be relevant. In those cases recontact of individual patients to obtain their consent may be practical, though an inconvenience to the investigators. It will remain for individual IRBs to judge what number of subjects is too many to make obtaining informed consent impractical, as well as when the scientific validity of the study depends on the waiver or modification of informed consent.

Prospective and Identified. There are two issues to be analyzed here. First, there is again the problem that obtaining consent will lead to selection bias, since some subjects will inevitably refuse. As in the case with retrospective protocols, investigators would have to demonstrate that selection bias is genuinely a problem for the study they are conducting, that there are no alternative methods that would provide equally valid results without the waiver, and that the value of the research to society justifies the violation of subjects' privacy and autonomy. Since the validity of some studies will not suffer as a result of selection bias, it is disingenuous to propose that *all* tissue protocols should be exempt from requirements for informed consent.

The second problem concerns the time and expense of obtaining informed consent. Although these protocols are *prospective* and hence do not face the obstacle of contacting subjects that *retrospective* studies face, still investigators may argue that because they have no professional relationship with the subjects, who may be surgical patients or patients undergoing routine or invasive diagnostic procedures, contacting these patients is *impractical*. Some have argued, for instance, that even if contact were practical, the consent process itself is more burdensome to the subject than the minimal risks involved in the research (20).

Again, however, the validity of these arguments will depend on the specific nature of the protocols and where the patients interact with the health care system. When the investigators are directly involved in ordering or performing the clinical procedures from which the tissues are collected, obtaining informed consent for the research use of tissues and review of medical records is convenient and feasible. Other studies, however, may involve the collection of tissues from a large number of satellite hospitals and clinics and enormously increase the burdens on the investigators if consent is required. One solution to this problem is to name local personnel at these institutions as co-investigators, who can then make the necessary arrangements to obtain informed consent.

We should be careful though, to distinguish between *impractical* in the sense that a study would be methodologically impossible to perform and *impractical* in the sense that the consent process is simply burdensome to the investigators. Would it be just an inconvenience to investigators

in a pathology lab to contact surgical patients to inform them about the nature of a tissue bank to which they hope to send samples of the patient's tissues, or would such a requirement make the collection impossible or impractical? How burdensome must the consent requirement be in order to find it impractical? Is the financial cost of obtaining informed consent a relevant consideration in these cases? Several commentators have argued that the increased costs to research of obtaining informed consent wastes limited financial resources because the informed consent process provides virtually no further protection from harm to subjects since the protocols under consideration for the waiver are by definition "minimal risk" (27,49,50,52). Melton has argued that since well over 95 percent of patients surveyed in Minnesota would gladly consent to the research use of their medical records (27), the requirement for informed consent is an added burden to patients at a time of stress, and the high costs of requiring consent are not justified by consideration for the abstract right of patient autonomy. As Phillip Reilly puts it:

> The cost of such research would be indirectly increased by the invocation of rules that, to me, only abstractly protect individual autonomy. So few people are likely to forbid their samples to be used for anonymous research that the expense attached to asking the question and tracking the few samples that are not available for study seems a poor use of resources (53).

However, as Veatch has argued, since disclosure of the purposes of research is an essential element of informed consent (35, p. 46), the federal regulations take seriously the right of subjects to decide for themselves what research purposes they wish to contribute to, a point not discussed by those who argue that consent should not be required for the use of tissues in research. Individuals may have a variety of reasons for wishing not to have their tissues used for different research purposes, from religious convictions regarding the disposition of body parts, to concerns about specific types of research and the commercial exploitation of their tissues or the patenting of genes (see Ref. 54 for a full discussion of these issues). The question of increased costs to research of requiring informed consent, and at what point those costs make the research *impractical*, leads to a direct confrontation with the respect for individual autonomy. We will discuss this dilemma later, but note here that judgments regarding the practicality of research without a waiver of informed consent are ethically complex, and there is little guidance in the regulations for IRBs and investigators alike. At the very least, it is prudent for IRBs to judge the impracticality of the consent requirement for each protocol on a case-by-case basis.

Blanket Consent. One important modification of informed consent should not be overlooked in applying this criterion to tissue collection protocols. Some hospitals contain either in their admissions literature or in their surgical and diagnostic consent forms provisions notifying the patient that tissues and medical record information may be used for research purposes. A few institutions provide the patient with the option to dissent from this use of these materials. These "blanket" consents or notifications simply tell patients in advance that their records and specimens may be used for research purposes. The National Commission had recommended that this type of notification or blanket consent be used in institutions that anticipated using tissues and records for research, and indeed, the National Commission tied this recommendation to their recommendation that explicit informed consent may then generally be waived for specific research protocols. Unfortunately, HEW/DHHS opted not to include this recommendation in the regulations. Nevertheless, institutions which employ blanket consents give researchers and IRBs the opportunity to use notification as a *modification* of informed consent that would not violate individuals' autonomy. Explicit informed consent may be waived since the subject has already consented to the use of their tissues and records for any research purpose.

The adequacy of these blanket consents and notifications, typically included in the small print of surgical consent forms, has been questioned however. In the context of genetics research, some commentators have argued that broad statements that tissues and medical records may be used for research purposes are inadequate to the complexities of genetics research (22, p. 1791; 35, p. 173; 55–58). Others, including the National Commission and the Privacy Protection Study Commission, concluded that the benefits of relatively unrestricted access to tissues and records, coupled with the minimal risks to subjects involved in their use, justify the use of blanket consent or notification measures (15,25, pp. 111–112; 59–61). Lacking further regulatory guidance, IRBs must rely upon the collective judgments of their members in determining the adequacies of blanket consent or notification for the specific protocols that they review.

Providing Subjects with Additional Pertinent Information after their Participation. The history of the promulgation of these regulations makes it clear that this requirement was intended to be applied primarily to psychological and social research which for methodological reasons involved deception or incomplete disclosure. In these studies, debriefing subjects afterward can help allay anxieties and stress that may have arisen through the deception, and in general speaks to the respect of the subjects as persons. In research involving solely the use of human tissues and associated medical records such issues do not arise, and in the absence of informed consent prior to the use of the tissues, recontacting subjects would place researchers in the difficult position of explaining to them that — through their tissues and medical records — they had been involved unawares in a research protocol.

Nevertheless, in some research involving identified or identifiable human tissues, researchers may discover clinically relevant information either as a direct result of the research or by happenstance. If informed consent is required at the outset, then it is possible to state up front under what conditions, if any, subjects will receive research results, and when necessary, be provided with adequate pre- and post-test counseling. A difficult problem arises when *unanticipated* research findings are discovered, as happened for instance, when a strong

correlation was discovered between the apo-E4 gene and Alzheimer's disease upon analysis of data gathered to study the relationship between apo-E family of genes and hypercholesterolemia and heart disease (62). These cases present the same dilemma that arises in studies for which the requirement for informed consent has been waived by the IRB. Such studies will typically involve identified or identifiable tissues (studies involving anonymous and anonymized tissues for the most part being exempt from the regulations), and therefore researchers may find that their results could be of clinical relevance to the tissue sources, who in these cases, will never have been informed of, much less consented to, the research use of their tissues.

This requirement for the waiver of informed consent stipulates that "*where appropriate*, the subjects will be provided with additional pertinent information after participation." The central question then is whether it is *appropriate* to provide research findings to subjects who have not given prior consent to the use of their tissues. The literature on this question, which has focused primarily on genetics research, is split. Those who argue for a duty to contact tissue sources do so based on two arguments: (*1*) a legal argument drawn on analogy to the clinical duty to recontact past patients with new information regarding their treatment, and (*2*) the obligation of researchers to *benefit* subjects whenever possible. Arguments against the duty to contact cite (*1*) the increased financial and administrative burdens that contact would place on research protocols, (*2*) the uncertainty of research generated results which may not be fully understood or validated by subsequent investigations, and (*3*) the psychological, social, and economic harms to subjects that may result if research findings are disclosed without adequate prior consent and counseling.

Arguments for a Duty to Contact

1. *Do Researchers Using Tissues Have Clinical Obligations?* To date there have been no cases litigated that would establish an investigator duty to contact tissue sources. Legal arguments suggesting such a duty have been based on analogy to cases in which physicians and health care institutions have been held liable for not re-contacting former patients with important findings regarding treatments previously provided. One of the most famous cases is that of the drug DES given to women in the 1950s. The University of Chicago was held liable for delaying notification to women who received this drug at their hospital four to five years after its toxic side effects became known in 1971. Pelias comments that the court found that the doctor–patient relationship is "on-going, especially when future injury to a client can be attributed to the relationship" (63). Pelias suggests that courts may also view the relationship of investigator to subject along analogous lines.

This point is also made by Hannig et al. who argue that the obligations of physician to patient are recognized, both legally and morally, to spread over the entire health care team, including consultants whom some state courts have found to have a legal duty of care to patients they may not have even seen in person (64). Research is then understood as just another part of a complex health care system, and the obligations placed on clinical health care providers should apply to researchers as well:

> Yet the law imposes a host of requirements on the practice of medicine regardless of the individual physician's actual motives. One cannot avoid these obligations simply by asserting that research is somehow different. Where, as here, research requires the assistance of certain individuals because they or their relatives have a problem that is the object of study and where the research is directed toward the diagnosis or treatment of this condition, research assumes the mantle of health care. In that setting, the law should not hesitate to impose on the researchers some duties of care toward those subjects as well, at least in the absence of explicit agreements to the contrary (64, p. 259; 65).

There are three responses to these arguments. First is the claim that their position risks confusing the separate roles of researcher and clinician. A number of commentators (25, p. 2; 66–67), including the National Commission, have argued that a clear separation between clinical and investigator roles must be maintained, even (and especially) when they are borne by the same person. Though this argument has historically been addressed to research involving the evaluation of therapeutic interventions, it applies equally well to the use of *information* derived from tissue specimens that may be of clinical relevance.

Ultimately the ground for clearly distinguishing research from clinical practice is based on the second argument, which concerns whether research findings are of sufficient validity on which to base clinical decisions. We will take this argument up below, but in short, the claim is that research findings are preliminary and may not be fully validated. The use of such research findings in the clinical context risks basing clinical decisions on incomplete or possibly false information that may lead to substantial harms to the patient.

The third argument seeks to dissolve the entire dilemma by ensuring that subjects are *informed* and counseled up front about what, if any, research findings they may expect to receive. Adequate planning at the start coupled with an explicit informed consent process will head off most dilemmas that may arise out of the research findings, anticipated or not (56, p. 87; 57). But this requires informed consent, and so the argument here is really against the waiver of informed consent entirely, if it is possible that research findings may be of clinical value to the subject. This question will be analyzed later in the general discussion of the ethics of the waiver of informed consent.

2. *The Obligation of Researchers to Benefit Subjects*. Several commentators have argued that researchers always have an obligation to serve the best interests of their subjects. Robert Veatch in particular has argued that human subjects should be viewed as *partners* in the research enterprise (35), a view also expressed by Jonas in a famous essay of 1969 (68). Veatch argues subjects of research are entitled to any benefits that the research may produce, including being informed of research results. If the results are merely preliminary, then subjects may be counseled regarding their uncertain status. Veatch does not, however, address the problems of contacting subjects who have never consented to the use of their tissues in research, since he argues that the waiver of consent

(and the exemptions) are inconsistent with the respect for individual autonomy that underpins the regulation and ethics of human subjects research.

Arguments against a Duty to Contact

1. *Increased Administrative and Financial Burdens*. The burdens to a research protocol of contacting subjects whose tissues have been used in research without their consent will depend on the individual protocol itself and the origin of the tissues used. Some protocols may involve relatively small numbers of tissues. In these cases the added record keeping, time, and resources spent in contacting the tissue sources would be small. Other protocols may involve larger numbers of tissue specimens, anywhere from hundreds to the tens of thousands. For instance, the NHANES III study conducted by the CDC (69) contains over 17,000 blood and DNA specimens (55,70). Many genetic epidemiological protocols similarly will involve very large numbers of tissue specimens. Although these research protocols may generate genetic or other information that would be of value to the tissue sources (and may even be of high quality and certainty), the costs of keeping track of and contacting large numbers of tissue sources could often exceed the entire budget of the research protocol in the first place.

Second, and in rejoinder to the argument that researchers owe subjects the benefit of research results, some commentators (55,71) have argued that the goal of biomedical research is to provide benefits to society at large and that insisting on providing benefits to research subjects whose participation is of minimal risk diverts resources that otherwise would contribute to medical and scientific progress and ultimately to the public good. If and when research results find their way into clinical practice, individuals whose tissues are used for research will have the same opportunity as the rest of the population to benefit from this new knowledge.

2. *The Uncertainty of Research Generated Results*. Data and conclusions derived from individual research protocols are often tentative and uncertain. The path from an intriguing research result to a validated clinical test, treatment, or procedure is long and arduous, requiring a series of clinical trials with increasing numbers of subjects. Many results that flow directly from research protocols are unsuitable for use in the clinical context without further testing and evaluation. If the conclusions derived from a research protocol are flawed and yet the results are passed along to the tissue source and their physician, it is possible that they will decide on an inappropriate course of action that may end up producing more harm than good for the patient. Here again, commentators on this problem have pointed to the necessity of maintaining clear distinctions between the often dual roles of researcher and clinician. Merz et al. in particular have recommended that tissue procurement and banking protocols be constructed so as to permit only a one way flow of information from patient to tissue bank to investigators so as not to confuse research and clinical information (18; see also Ref. 28).

In any event, the clinical validity of research results must be evaluated by investigators and IRBs on a case-by-case basis. The possibility that some research results are relevant to clinical care should not be ruled out, especially when there are interventions available that may reduce or minimize harm to the patient. When contact is made with an unconsenting subject, however, a number of problems may arise which are discussed in the next section.

3. *Disclosing Research Results without Adequate Prior Counseling*. When informed consent has not been required for researchers to obtain and use human tissues, contacting subjects presents a dilemma. In standard clinical practice, diagnostic tests typically are preceded by either the tacit consent of the patient whose presence in the physician's office indicates their willingness to investigate a particular health problem, or by the explicit and sometimes written authorization of the patient if the test is particularly invasive or will produce sensitive information. For instance, most genetic testing clinics require extensive pre-testing and post-testing counseling of their patients (72–77). Such counseling includes information about the nature of the test itself, its relation to symptoms and disease, including the penetrance of the gene and limits of the predictive value of the test, and the social consequences of the genetic information, including the possibility of various forms of economic and social discrimination. When the test results come back, patients are more psychologically and emotionally prepared to cope with the consequences of bad news and often undergo further counseling.

It is a well documented fact that depending on the disease gene in question (and especially whether any treatment for it exists), many patients who contemplate a genetic test opt not to perform the test after initial counseling (78,79). Some patients decide that the burdens and risks of knowing are not justified by the benefits. However, if individuals have never consented to the research testing of their tissues, there is no way for investigators to know whether they would want to know the results or not. It is largely a value judgment that individual patients make for themselves when they decide for or against having a given diagnostic test. If unconsenting subjects are recontacted, some subjects may be forced to learn things about themselves and their potential medical future to which they never would have consented in the first place had they been given the opportunity.

One solution to this dilemma advocated by a number of commentators is the use of a general newsletter from researchers describing aggregate results to subjects (56). This presupposes prior informed consent in order to justify this form of contact with the subjects, but publication of study results to the general population from which the tissue sources were derived may fulfill the same role. It should also not be overlooked, as pointed out above, that as research knowledge finds its way into clinical practice in the form of validated procedures, the unwitting tissue sources will have the same opportunity to benefit from these procedures as the general population and in a manner in which their autonomy and values are respected.

The duty to contact is a complex issue, but despite the ongoing debates, the research community has generally tended to limit contacting unconsented tissue sources. There may be instances where this has not been the case, and IRBs and investigators should look at each

case carefully, but in general it is more prudent not to permit contact of unconsented tissue sources with research findings. Investigators who find themselves with information that they feel is compelling enough to warrant the risks of contact should certainly consult their IRB before proceeding.

General Requirements for Informed Consent and Special Provisions for Genetics Research

Protocols that fail to meet the criteria for either exemption or the waiver of informed consent are required by the regulations to be reviewed by the IRB and must conform to the requirements for informed consent. Since these protocols may be eligible for expedited review we will in this section examine (*1*) the conditions for such review, and then turn to the general discussion of (*2*) informed consent and (*3*) the special provisions for genetics research.

Expedited Review. The review of tissue protocols, if they present no more than minimal risk to subjects, may be done on an expedited basis according to the conditions outlined in §46.110 of the regulations. These regulations refer to the Expedited Review List published by DHHS in 1981 (80). A large number of the items contained in this list refer to the collection of human biological specimens, including the collection of hair and nail clippings, external secretions, amniotic fluid and placenta (at the time of birth), excreta, blood draws of less that 450 ml, dental plaque and saliva, and the study of existing pathological and diagnostic specimens. This latter category would involve only studies in which identifiers or links are recorded by investigators, since otherwise these protocols would be exempt.

The expedited review may be carried out by the IRB chair person or an experienced member of the IRB. This reviewer may exercise all the powers of the IRB, including the waiver of informed consent, but protocols under expedited review must meet all the conditions for IRB approval stipulated in the regulations. The expedited review process simply permits a more timely review of protocols and was designed to minimize the work of the full IRB.

DHHS recently expanded the list of expedited review categories (81). The list increases the amount and frequency of blood drawn and most notably includes research which collects identifiable records and pathological and diagnostic specimens, whether they are collected retrospectively or prospectively, provided that they are excess or have been produced for nonresearch purposes. However, a report by the DHHS Inspector General's Office on IRB performance noted plans to give IRBs "added responsibility in the areas of genetics and confidentiality" (10). Permitting more of these and related protocols to be reviewed on an expedited basis would seem to work against the need for greater IRB vigilance and expertise in these areas.

General Requirements for Informed Consent. Section 46.116 lists eight elements required for informed consent:

1. A statement that the study involves research, an explanation of the purposes of the research and the expected duration of the subject's participation, a description of the procedures to be followed, and identification of any procedures that are experimental.
2. A description of any reasonably foreseeable risks or discomforts to the subject.
3. A description of any benefits to the subject or to others that may reasonably be expected from the research.
4. A disclosure of appropriate alternative procedures or courses of treatment, if any, that might be advantageous to the subject.
5. A statement describing the extent, if any, to which confidentiality of records identifying the subject will be maintained.
6. For research involving more than minimal risk, an explanation as to whether any compensation and an explanation as to whether any medical treatments are available if injury occurs and, if so, what they consist of, or where further information may be obtained.
7. An explanation of whom to contact for answers to pertinent questions about the research and research subjects' rights, and whom to contact in the event of a research-related injury to the subject.
8. A statement that participation is voluntary, refusal to participate will involve no penalty or loss of benefits to which the subject is otherwise entitled, and the subject may discontinue participation at any time without penalty or loss of benefits to which the subject is otherwise entitled.

Not all of these requirements are appropriate for tissue protocols: Specific consent requirements depend on whether the intervention is research related or for clinical purposes, or whether only excess tissue specimens are used or the collection of additional tissue is required. Many research institutions maintain standard informed consent templates that investigators may then customize to fit the individual needs of their protocol. It is important to remember that the risks of procedures ordered for clinical purposes need not be detailed in the research consent process. Only those risks that arise from the specific research activities need to be detailed, since the clinical consent process should cover the risks attendant to the diagnostic or surgical intervention.

Special Provisions for Genetics Research. Some research protocols using human tissues requiring informed consent will be of little risk to the subject. For instance, research that uses normal blood specimens to study the clotting process is relatively straightforward, and the informed consent process need not be lengthy or complex. However, the use of tissues, especially for genetics studies, can present an array of problems that must be dealt with in the informed consent process. The purposes of the research, the communication of research results to the sources, the psychosocial risks of such disclosure, questions of ownership of tissues and rights to commercial profits, the use of specimens for other research purposes or the

disposal of specimens after the research is complete, may all be relevant considerations.

While the clinical genetics community has reached a consensus on the elements of clinical informed consent, good counseling practices, and the protection of patients from psychosocial harms arising from the dissemination of genetic testing results, the research community continues to debate both the need for, as well as the extent of, informed consent for genetic studies. On the one hand are those like George Annas, who advocate for a robust informed consent requirement. Annas has proposed that:

> No collection of DNA samples destined for storage is permissible without prior written authorization of the individual that a) sets forth the purpose of the storage; b) sets forth all uses, including any and all commercial uses, that will be permitted of the DNA sample; c) guarantees the individual (i) continued access to the sample and all records about the sample and (ii) the absolute right to order the identifiable sample destroyed at any time; and d) guarantees the destruction of the sample or its return to the individual should the DNA bank significantly change its identity or cease operation (56).

Annas would further require that DNA samples be used only for the purposes for which they were originally collected and would prohibit the use of open-ended consents that permit the use of the sample for any other research purpose. He would also prohibit third-party access to research results but would require notification and adequate counseling of sources concerning research results that may have a significant health impact.

Annas's proposals as embodied in the Genetic Privacy Act (GPA) (82), a proposed model law, have been influential in a variety of legislative proposals around the country, but they have met with opposition from the research community. Some commentators have argued that singling out genetic information from other types of medical information ("genetic exceptionalism") is impossible to apply in practice, since a variety of medical information can be considered "genetic" (28,83). It is also suggested that elaborate informed consent, more consistent with interventional research, is inappropriately applied to the collection and use of tissues and medical records, since little or no risk is involved in the research use of tissues. Furthermore an elaborate consent may frighten subjects and raise the spectre of significant harms that are of extremely low probability, leading to what Korn has called "uninformed denial" (28, p. 25). Most of the objections to Annas's position concern the question of whether informed consent should be required at all for such protocols, and we will discuss this question in detail below.

An alternative idea — a tissue "advance directive" — has been proposed by Robert Weir (84). This document would be filled out by patients as they enter a health care institution and could limit the research use of their tissues to specific types of diseases, allow no research use at all, or permit any research use. This detailed blanket consent could then be entered into the patient's computerized medical record and easily tracked by pathologists and researchers. This proposal would have the virtue of avoiding a detailed consent process each time a researcher wanted to use an individual's tissue. Individuals could update their "tissue directive" as well as specify whether they desired to be informed of any research results. Whether such a system is feasible warrants further study, for it would obviate the need for specific informed consent for many research uses of tissues and records, while giving individuals more control over the purposes for which these materials are employed.

AUTONOMY, PRIVACY, AND THE SOCIAL GOOD

Although some controversy still surrounds the extent and detail of informed consent for the use of human tissue in research, the primary debate has centered on the question of whether informed consent should be required at all for these protocols. As seen above, the regulations support several exemptions from both informed consent and IRB review in addition to the waiver of informed consent. Those who support these measures and would seek to expand them make six general arguments: (*1*) requiring informed consent is administratively and financially burdensome; (*2*) requiring informed consent may introduce selection bias into research; (*3*) the informed consent process is itself a burden to patients and subjects and as such outweighs any benefits the consideration for autonomy and privacy would produce through informed consent; (*4*) there is a long tradition in medicine of free and unfettered access to tissues and medical records for research purposes, and requirements for informed consent and/or IRB review for currently exempt categories of research would infringe upon academic freedom; (*5*) the social benefits of minimal regulation of research involving human tissues and medical records, including the omission of informed consent, greatly outweighs any violation of individual autonomy; and (*6*) there is no violation of individual privacy in the use of tissues and records, since confidentiality is scrupulously protected by researchers.

In response, those who advocate for greater individual control over their tissues and records as well as regulatory and IRB oversight of this research argue that (*1*) the ethical foundation of informed consent is a respect for individual autonomy, and permitting research on tissues and records without informed consent violates a growing tradition in research ethics that has served patients and subjects very well; (*2*) social benefits are not a sufficient ground to undermine the respect for autonomy by not requiring informed consent, and indeed, social benefit is incommensurate with respect for autonomy and cannot simply be weighed against it; (*3*) as a result of the successes of the Human Genome Project, human tissues contain a vast amount of medical information that easily can increasingly be tapped, and subjects have a right to determine whether and how this highly personal information is used; (*4*) public trust in the health care industry and in biomedical research is undermined when researchers use individuals' tissues without their consent or knowledge for purposes they may not share, and especially when researchers might profit financially from them; and (*5*) our current health care system does not distribute its benefits equally across the population, and

many minorities and the poor are excluded or underserved. Justifing the violation of individual autonomy by the claim that research on tissues serves the public good fails to recognize that this public good does not return equally to all segments of the population.

With the exception of this last argument which is predicated upon the principle of *justice*, it can readily be seen that the core of the ethical debate over whether to require informed consent pits those for whom social benefit — a concept derived from the general principle of *beneficence* — is held as the highest moral principle, against those for whom the principles of individual *autonomy* and *privacy* take precedence over competing principles and values. Beneficence, according to the National Commission, implies two separate obligations: "do not harm and maximize possible benefits and minimize possible harms" (12). These benefits and harms may serve to advance or hinder the interests of either individuals or society in general. Those advocating less regulation and no informed consent typically will view autonomy as a species of this more general category of beneficence: autonomy and privacy simply represent *interests* that are important to individuals but that must be weighed against competing *interests* such as general social benefits. On the other hand, those who advocate for the requirement of informed consent typically will understand autonomy as a principle independent of beneficence and either conceive of it as a moral *right* or as entailing a moral *duty*.

How the Regulations Encode and Interpret Autonomy and Beneficence

The involvement of the U.S. government in the regulation of human subjects research evolved out of revelations in the 1960s and 1970s of various abuses that risked or caused grave harm to research subjects. It was recognized that such abuses would be less likely to occur if subjects were fully informed of the purposes of, and risks (and benefits) of the research, since subjects would be reluctant to submit to risks without the corresponding possibility of benefits. But informed consent also came to serve another purpose: respect for individual autonomy, both moral and legal. Legally, those who practiced medical interventions upon unconsenting individuals were liable for *battery*, defined as unwanted or unconsented touching of one's body. In the legal domain, autonomy is expressed as a right to control one's body and what is done to it. This legal right entails an obligation or duty on the part of others not to violate this right. Morally, the requirement of informed consent is an expression of the respect for the independent will or freedom of the individual to associate and participate in activities of their own choosing. In the moral domain, autonomy is expressed first as a duty, not as a right, though some have inferred a corresponding moral right of autonomy. This duty is one that demands that in all our actions we respect the free will and self-determination of others. As a moral right, autonomy has come to mean, in our society, the freedom to choose.

In most cases the principles of autonomy and beneficence work hand in hand for the protection of human subjects in biomedical research. However, research involving human tissues and medical records, insofar as it is genuinely of minimal or no risk requires that we evaluate the justification for informed consent solely on the basis of autonomy, since the beneficence role informed consent plays in minimizing harm is, by definition, out of play. The role that the principle of beneficence then plays is to ground the research project itself as contributing to the social good. Thus the stage is set for the conflict between autonomy (*informed consent*) and beneficence (*social good*).

In its seminal *Belmont Report*, the National Commission identified three basic ethical principles that apply to research involving human subjects: justice, beneficence, and autonomy. The National Commission recognized that informed consent served two ethical masters (beneficence and autonomy), but constructed its recommendations to HEW in such a manner that neither came into conflict with the other. The commission did indeed recognize that research involving human tissues and medical records involved an "invasion of privacy," but this invasion, in the absence of informed consent, was justified, on the one hand, by the social value of the research and, on the other hand, by the recommended requirement that institutions conducting such research institute a blanket consent at the time of admission. The commission also recognized that IRB review was necessary to evaluate the social value of the research, which, in the absence of informed consent, would entail a judgment that the research is of a socially acceptable nature — that is, that the research would not likely be objected to by those whose tissues and records are conscripted without their informed consent.

In its reasoning, the National Commission was guided by the conclusions of the Privacy Protection Study Commission (PPSC), which had published its report in 1974. The PPSC had not considered the use of tissues but did examine carefully the research use of medical records and recommended that this use is legitimate provided that individuals are *notified* in advance (59). The National Commission took a stronger stance with regard to tissues and records, recommending blanket *consent* rather than mere notification. The debate over the adequacy of blanket consent aside, this provision represents the recognition by the National Commission of the independence of autonomy as a ground for informed consent. IRB review would stand as proxy for the *informed* consent of individual subjects, what Veatch has called "constructed consent" (35, p. 63; see also Ref. 25, p. 150, and Ref. 85), while blanket *consent* for the research use of tissues and records would permit those who might object, for whatever reasons, to opt out of any research use of their tissues and records.

As we have seen in the discussion above, HEW did not follow the National Commission's recommendations for a standard waiver of consent for tissue and medical record research coupled with a blanket consent requirement at institutions conducting such research. Instead, HEW proposed, and DHHS adopted, the two exemption categories as well as a general provision for the waiver of informed consent. But neither the proposed nor the final regulations contained any provision for blanket consent. There is no comment on this omission in the public record, but HEW/DHHS may have decided that it would have been an unwarranted intrusion of government regulation into *clinical* affairs to require blanket consent at all

hospitals and health care facilities that might supply tissues and records for research. But blanket consent was an essential element in the National Commission's recommendations that preserved respect for autonomy while greatly simplifying research access to tissues and records. And IRB review was essential to ensure that the purposes of the research were generally unobjectionable in the public eye. There are two ways, then, to interpret this omission by HEW/DHHS. Either the compromise reached in the final regulations does not recognize the independent validity of autonomy as a ground for informed consent, or the compromise reached simply distorted the ethical ground so carefully crafted by the National Commission.

There is good reason to believe that the latter is the correct interpretation. In his *Patient as Partner*, Veatch has argued that the informed consent requirement (§46.116(a)(1)), that subjects be informed of the *purposes* of the research protocol, can only be justified by appeal to subject autonomy and cannot be based on any beneficence considerations to prevent or minimize harm (35, p. 99). By requiring that the purposes of the research be stated clearly in the consent form, subjects are given the opportunity to accept and adopt such purposes as something to which they are willing to contribute. Informing a research subject that a protocol will contribute to the treatment of a particular disease or, perhaps, make abortion less risky to women, plays no role in the subject's evaluation of the physical risks of participating in the protocol but speaks instead to the values they wish to promote or not promote.

Veatch argues that the exemption categories as well as the provision for the waiver of informed consent are fundamentally at odds with the independent role of autonomy in the justification and constitution of the informed consent requirement. By exempting a large class of research protocols as well as permitting the waiver of informed consent, the regulations usurp the ability of individual subjects to make value judgments regarding the goals of the research protocols to which their genomes, tissues, and medical records contribute. It was perhaps the view of the authors of the regulations that biomedical research was of such a high and uncontested value that no one would reasonably object to contributing their tissues and records to the common good. But this notion would apply equally well to invasive and physically risky research protocols as it does to protocols involving only the use of tissues and records. There is no sufficient ground for explaining to subjects the purposes of research in the one case but not in the other. It therefore follows that the regulations must recognize the independent validity of autonomy as an ethical justification for informed consent (86).

Defining Privacy

We noted above that the National Commission viewed researcher access to tissues and records in the absence of informed consent to be an "invasion of privacy." It is useful at this point to clarify the concept of privacy, since it is intimately related to the concept of autonomy and suffers from the same ambiguity of interpretation.

One of the most influential articulations of what is meant by the legal "right to privacy" (as distinct from "privacy" as a moral interest or value) appeared in a now-famous *Harvard Law Review* article in 1890 by Samuel Warren and Louis Brandeis. The right to privacy, they said, is "the right to be let alone" (87). It was from this simple articulation that common law began to recognize a right to privacy in certain circumstances. This recognition, however, has not made defining the concept of "privacy" any easier. Indeed, privacy is a "notoriously vague, ambiguous, and controversial term that embraces a confusing knot of problems, tensions, rights, and duties" (88).

Privacy is usually described as being related to notions of solitude, autonomy, anonymity, self-determination, and individuality: It is experienced on a personal level. Within socially and culturally defined limits, privacy allows us the freedom to be who and what we are as individuals. By embracing privacy, we exercise discretion in deciding how much of our personhood and personality to share with others. Moreover we generally feel less vulnerable when we can decide for ourselves how much of our personal sphere we will allow others to observe or scrutinize (89). Complicating the process of defining what privacy "is" is the fact that it often means something different to nearly everyone, and the experience with and perception of what invades privacy will likely differ significantly from person to person (42).

In trying to "break apart" the notion of privacy, much of the literature focuses on the following elements:

1. *Autonomy*. Respecting the dignity of each individual to make decisions for themselves, free from coercive influences. It also encompasses our need for solitude and intimacy. A National Research Council report stated that the protection of individual autonomy is a fundamental attribute of a democracy (90). Autonomy is also addressed in other analyses as "decisional privacy" (91).
2. *Informational privacy*. Defined by how much personal information is available from sources other than the individual to whom it pertains. Informational privacy encompasses the ability to limit access to one's personal information (which also supports the autonomy aspects of privacy), from both a quantitative and qualitative perspective — namely the amount and type of information one surrenders, either voluntarily or by coercion. It also involves when such information should be communicated or obtained, and what uses of it will be made by others. It includes the collection, storage, use, maintenance, dissemination/disclosure, and disposition of personal information.
3. *Freedom from intrusion / surveillance*. Encompasses, in part, an individual's desire to preserve his or her anonymity and solitude (both physical and emotional solitude). This notion includes not only the individual's desire to limit access to information about him/herself but also to be free from physical intrusion and observational surveillance by others. Surveillance can have a chilling effect on individuals,

> as noted by many sociologists and studies of electronic monitoring. Individuals often change their behavior to conform to what they believe those monitoring their movements/actions will find "acceptable" or "normal" (92,93). Freedom from intrusion is addressed in other analyses as "physical privacy" (91).

The concept of privacy is often confused with or treated as synonymous with two other distinct concepts: confidentiality and security. Confidentiality

> refers broadly to a quality or condition accorded to information as an obligation not to transmit that information to an unauthorized party. . . .Confidentiality has meaning only when the promises made to a data provider can be delivered, that is, the data gatherer must have the will, technical ability, and moral and legal authority to protect the data (90).

The following is a simple way to differentiate between these three concepts: *security* measures provide the technical (and sometimes physical) means to safeguard the *confidentiality* of personal information, which in turn protects the *privacy* of individuals. Within the doctor–patient relationship, confidentiality is used to describe the relationship of trust that must exist for appropriate clinical care to be rendered. In its essence, confidentiality advances the protection of personal information that is exchanged or generated between doctor and patient (whether through verbal exchanges of information or information generated through physical examinations). This is the most fundamental way in which the patient's privacy is preserved (42).

One of the arguments made against the unconsented use of human tissues and medical records is that it violates this fundamental trust between patient and physician. Even if health care is delivered by a team of nurses, physicians, and other personnel, there is a fundamental expectation on the part of patients that their medical information will be kept confidential. What this means to most patients is that this information will not be divulged to other persons or institutions for purposes other than what the patient initially provided it for, namely their own medical benefit and, correlatively, for purposes of payment. The Privacy Protection Study Commission (PPSC) concurred with this understanding in its 1974 report when it recommended that medical records could be used for:

> . . .conducting a biomedical or epidemiological research project, provided that the medical-care provider maintaining the medical record: (i) determines that such use or disclosure does not violate any limitations under which the record or information was collected (59, p. 306).

Of course, virtually all standard medical record information is collected with the expectation of confidentiality. It was thus necessary for the PPSC to insist in its recommendation 12:

> . . .that each medical-care provider be required to notify an individual on whom it maintains a medical record of the disclosure that may be made of information in the record without the individual's express authorization (59, p. 313).

The National Commission, as we have seen, took a stronger stance and recommended blanket consent rather than mere notification. But regardless, only through such notification or blanket consent procedure could private medical information be disclosed for research purposes in a manner that did not violate the *expectation* upon which the information was first gathered.

We may therefore understand and interpret "privacy" according to the intentions and expectations of patients when they provide specimens or medical information to their providers. Confidentiality is the *respect* providers pay to these intentions of their patients by not disclosing medical information. Or, confidentiality is a respect for the autonomy of the individual to determine the purposes to which information (and their diagnostic specimens) are put. In the medical context, privacy becomes a species of autonomy and confidentiality a species of the respect for this autonomy, or more generally, a respect for persons. Thus privacy is not violated if the patient knowingly *consents* to the disclosure of information (or the use of tissue). It should therefore be clear that privacy and autonomy with respect to one's tissues and records is of equal concern whether the records are identified or used anonymously. Although the use of identifiers or links to identifiers may raise the risks of harm to the source and correlatively the use of anonymous tissues protects *confidentiality*, it is the expectations placed upon the information and tissues by patients which determines them as private in the first place. Only through *consent* can their use respect the privacy and autonomy of sources (for a contrary view, see Ref. 85, p. 176).

As thus interpreted, either privacy may then be evaluated as an *interest* patients have in controlling their medical information and therefore may be evaluated alongside other such interests and goods within a consequentialist framework, or privacy may be understood deontologically as a duty on the part of others to respect what is private, that is, not to violate the intentions according to which one may receive private information. The ambiguity that attends the interpretation of autonomy — whether it may be subsumed by beneficence or is independent of beneficence concerns — also attends the interpretation of privacy. Indeed, autonomy and privacy are so intertwined that it may well be the same ambiguity in either case.

Consequentialist Interpretation of Autonomy

As we have seen the exemptions and waiver provisions of the regulations appear to simply disregard the independent role of autonomy as an ethical justification for the requirement of informed consent. Because the regulations permit research access to tissues and records in the absence of specific informed consent and yet do not require any form of blanket consent, the regulations appear to authorize an "invasion of privacy" that two federal commissions recognized as violating patient autonomy.

There remains one route out of this dilemma that would save the consistency and coherence of the regulations. It

is the argument offered by some proponents of greater research access to tissues and records without informed consent. The only way in which autonomy can be traded off against beneficence concerns (informed consent balanced against, or sacrificed for, the social good) is if autonomy itself is understood as the expression of certain *interests* that individuals and society regard as especially important. Once autonomy is understood as the expression of interests, the task is either to minimize the harm to these interests that research may cause or to demonstrate that such harms to them as do occur are minimal and do not *outweigh* the benefits to society of the research. In a paper commissioned by the National Bioethics Advisory Commission (NBAC), Allen Buchanan has made precisely this argument. Since this argument has formed the basis of many of the recommendations the NBAC report (6) on the use of human tissues in research, it is worth while analyzing this argument in detail.

This argument begins with the assertion commonly made by those who would require informed consent for tissue research that individual autonomy is a *right* that permits patients to control private information about them, including their tissues. The language of rights is an appealing framework for moral argument, since rights often stand as moral trump cards overriding other interests.

If we begin with the assertion of individuals' autonomy rights, or privacy rights, against the interest researchers have in freer access to tissues and medical records, the outcome of the argument will be determined by how these putative moral rights are justified or grounded. Buchanan suggests, in line with some traditions of moral and political thought, that rights are "protectors of morally important interests" (94). Specifically:

> . . .rights-statements are assertions that certain interests are of such importance from a moral point of view that they deserve especially strong protections (94).

The claim that such interests are protected by "rights" simply asserts that the interest is of such high priority as to overrule other competing interests and rights. As Buchanan points out, such assertions are the conclusions to moral arguments that must be provided in order for the assertion to be justified.

Having established that a right is simply a proxy for an important interest, the task is to demonstrate that the interest that justifies the right is of greater or lesser priority relative to competing interests. Interests are defined as "an ingredient in someone's well-being" (94, p. B-5). Well-being in turn relates to the goods and benefits that individuals pursue, either as a matter of necessity or to satisfy their desires. Interests thus represent goods or benefits. Insofar as a right is violated, the interest in that benefit is set back; that is, the individual is harmed. This argument has succeeded in converting the principle of autonomy into a species of the more general principle of beneficence, and the deontological force of autonomy, that is, the *duty* to respect another's autonomy, is measured according to the consequences of respecting or violating the interest expressed by this putative right. One simply reckons up the benefits and harms according to a utilitarian calculus.

In the case of the research use of human tissues and medical records, one need only reckon up the harms (and benefits) to individuals alongside the harms and benefits to the social good that the use entails with and without consent. The argument from this point is straightforward. Since only in special circumstances does society permit the unconsented risk of harm to individuals for the sake of social benefit (e.g., military conscription), the task of research interests is to demonstrate that harms to individuals are minimal or *inconsequential*. By protecting confidentiality, adverse psychosocial consequences to subjects are minimized. We are then left with what Reilly has called the abstract protection of individual autonomy (53) weighed against the obvious social benefits of biomedical research. Since little of *consequence* hinges upon the exercise of autonomy in this instance, and the benefits of biomedical research to health care and the economy in general are so high, overriding the "right" of autonomy is easily justified. Thus Buchanan concludes that even blanket consent may not be necessary:

> But it would be hyperbole to say that a system that does not include the requirement of blanket consent violates anyone's "right to autonomy." For one thing, . . . not all choices warrant the stringent protections that talk about a right to autonomy implies; some choices are relatively insignificant because they are largely irrelevant to a person's well-being and value (94 p. B-18).

If we interpret autonomy along these lines, the exemptions and waiver of informed consent provisions in the regulations in the absence of a blanket consent requirement no longer appear inconsistent with the otherwise independent validity of the principle of autonomy. On this interpretation, stating the purpose of research when informed consent *is* required is justified by the respect for the autonomy *interests* of the subjects. It costs research little to include this. The requirement for informed consent itself is justified by the beneficence concern to avoid harm by affording individual subjects the opportunity to evaluate the risks and benefits, as well as the overriding legal concern not to put individuals at risk without their prior informed consent. In the case of the research use of tissues and medical records, however, since there are few or no risks that cannot be managed through the maintenance of confidentiality, the beneficence ground of informed consent does not apply. Lacking any further ground for informed consent, it is morally justified, according to this argument, not to require it.

This argument forms the foundation for the six arguments cited at the beginning of this section supporting researcher access to tissues and records without the informed consent of sources.

1. The requirement for informed consent, in addition to the requirement for IRB review, adds administrative and financial burdens to research protocols. It requires the filing of paperwork and the tracking of subject consent documents in addition to the time and personnel resources to contact patients

and go through the informed consent process with them. When protocols involve multiple centers for the collection of tissues and records, these burdens are multiplied many times over. In addition some epidemiological studies include such large numbers of subjects that the costs of obtaining consent would far exceed the protocol budgets, thus making the studies impossible to perform (27,49). If the risks of confidentiality are minimized through appropriate protections, the social benefits of research can be maximized by eliminating the costly burdens of informed consent.

2. Selection bias may be introduced into research by requiring informed consent. Many protocols examine trends and small statistical differences. Permitting some subjects to opt out of the use of their tissues and records can introduce uncontrollable statistical bias. Furthermore, in some areas of research, because of the sensitivity of the topic (e.g., psychiatric or sexual dysfunction), informed consent may make it difficult to obtain sufficient specimens and records, making such research impossible. As a result some areas of research could be abandoned that could produce important benefits for these same populations (27,49,50,53).
3. Requiring informed consent for the use of excess surgical and diagnostic tissues would place an added burden upon patients, often at times of emotional stress. Approaching patients prior to surgery may confuse them and raise unnecessary fears in their mind because they may fail to understand the complexities and subtleties of the research use of their tissues. Furthermore, since the risks to subjects are minimal if not nonexistent, there is little benefit to be gained by informing subjects and asking their consent. The informed consent process itself is a burden to the subjects with little or no corresponding benefit (20, p. 76).
4. Researchers have long enjoyed unfettered access to human tissues and medical records. The benefits of this access are innumerable, and the record of researchers protecting the few risks that might arise is impeccable. This argument is partially predicated upon the claim that sources simply do not have any rights to control the use of their tissues, rights which they gave up when the tissues were removed (53). Researchers and hospitals consider tissues not used for direct patient benefit in diagnostic testing or treatment to be "waste," which the institution may dispose of as it sees fit, provided it does not directly harm the patient in the process (51, p. 6).
5. The social benefits of research outweigh any putative violation of individual autonomy (53, p. 380). Indeed, since autonomy has no other meaning in this argument than to promote certain benefits and prevent certain harms, and since researchers control for these harms by maintaining confidentiality, autonomy has no role to play in such research. This is the essence of the consequentialist argument presented above.
6. There is no violation of any putative "right" to privacy as researchers scrupulously protect the confidentiality of the information and tissues they receive. Patients in the contemporary health care system must consent to a wide array of individuals and institutions gaining access to their medical records. Research is an integral part of the health care system and provided protections are in place to maintain confidentiality, subject privacy is maintained (95).

It is worth noting here that there is also one argument that employs the idea of the social good as a justification for the requirement of informed consent. Clayton et al. argue that respect for autonomy is an important social good that the informed consent requirement promotes. We should not sacrifice this good for the lessor goods of scientific progress, which could undermine public trust in the research establishment (22, p. 1787). In a similar vein Hans Jonas has argued that scientific progress is an *optional* good, one which brings benefits but not benefits that are so *essential* to the public welfare that they justify overriding individual autonomy (68, p. 230). Although a majority of individuals would gladly consent to such uses, they also find benefit in the asking. For instance, in a 1993 Harris Poll survey, 64 percent of patients surveyed indicated that permission should be obtained before medical records are used for research, even if individuals are not identified in any publication (96). Whether the *benefits* of respecting the autonomy of patients and subjects by requiring informed consent outweigh the social benefits of research conducted in the absence of this requirement is fundamentally a value judgment, which in the regulatory context must be decided through political process.

Deontological Approach to Autonomy

The consequentialist interpretation of autonomy is predicated upon two premises: (*1*) moral duties derive from moral rights and (*2*) moral rights are nothing more than the expression of strong moral interests. Accepting these two premises leads inevitably to the conclusions analyzed above. The fateful move, however, is to accept the first premise. Although the use of the language of rights in moral discourse may be quite common today, it is a relatively recent phenomenon. In the history of ethics, ethical *duties* have rarely been interpreted according to rights language, and even more rarely actually derived from rights.

The derivation of duties in a deontological system may be seen by way of contrast to the consequentialist approach. We may evaluate actions according to their *consequences*, that is, the value of the results of particular actions. This is the core of utilitarian ethics in its many different forms. It depends, then, on how we define the *good* we seek to promote by our actions before we can evaluate these consequences. Alternatively, we may evaluate the actions according to the motivations or reasons for the action. This is the way of a *deontological* ethics. In the former case, it is the ends that justify the means, whereas in a deontological system, the means must be justified independent of the ends, that is, the action itself must

have moral worth independent of any consequences which flow from it. The method according to which we define the *good* in action will depend upon the ethical system we employ. In a deontological moral system, agents have a *duty* to perform those actions which are good, regardless of the consequences.

There have been several articles in the literature that employ deontological reasoning to evaluate the question of requiring informed consent in general and in particular for the research use of tissues and records. In a general analysis of consent, both Lebacqz and Levine (97) and Marshall (98) cite autonomy as defined by Immanuel Kant as the foundation of informed consent alongside the beneficence concern to minimize harm. Clayton et al. cite a deontological interpretation of autonomy as the ground for requiring informed consent for research involving tissues and records (22, p. 1788). And Faden et al. follow the historical rise of autonomy in the regulation of research (and in the practice of clinical medicine) as an increasingly important ethical foundation for informed consent (33).

Capron provides an intuitively clear argument that highlights the difference between a deontological and consequentialist ethics by distinguishing between harms and wrongs (99). Suppose someone enters your house without permission and looks through your belongings without taking or damaging any of them. Capron argues one may not be *harmed* by this invasion of privacy, but one is certainly *wronged* by it, since one's private life has been exposed without one's permission. This analogy is then extended to the medical record that contains a great deal of intimate information about the individual, which is also the case with human tissues. The use of medical records and tissues *without consent* may not *harm* the individual (though harm is possible), but it does violate the individual's privacy, or *wrongs* them. In the consequentialist domain, wrongs are simply measured by the severity of harms. In a deontological framework, the wrong is constituted not by any harm to the individual but rather by the violation of their own self-determination. The intruder into the home and the medical researcher who obtains private medical information without consent, both fail to respect the self-determination, or autonomy, of the individual, regardless of whether harm is suffered or not.

The concepts of *respect for persons* and *autonomy* have long been used interchangeably. The National Commission interpreted respect for persons in the *Belmont Report* as the requirement that individuals "should be treated as autonomous agents" (12, p. 4), and defined such an agent thus:

> An autonomous person is an individual capable of deliberation about personal goals and of acting under the direction of such deliberation. To respect autonomy is to give weight to autonomous persons' considered opinions and choices while refraining from obstructing their actions unless they are clearly detrimental to others (12, p. 5).

The key to this interpretation of autonomy is the idea of self-determination, that individuals may set their own goals and act accordingly, within the limits of not acting in a detrimental fashion toward others, that is, either harming them or violating their autonomy. Autonomy is thus related to the concept of *freedom*, and is typically understood as a matter of *choice*.

Our contemporary concept of autonomy derives for the most part from the moral philosophy of Immanuel Kant, and is intimately related to the concept of freedom (100). For Kant, freedom has two aspects: negative and positive. Negative freedom is the absence of constraint. One is free to the extent that one is not constrained in one's actions by external constraints (chains, iron bars, poverty), or by internal constraints, such as the influence of desire upon one's decision making. But negative freedom is incapable of steering one toward any particular action. For this, we need *positive* freedom — a ground for the determination of our actions. This positive freedom Kant argued is *autonomy*, which he understands in its etymological meaning of *self* (auto) *law* (nomos). Far from being the freedom to do anything that one wishes or wants, autonomy for Kant is the freedom to legislate the moral law *for oneself* and in this sense is genuinely self-determining. Positive freedom is thus the ability to self-legislate the moral law. For Kant, autonomy is a much richer concept than the mere "freedom to choose" by which it is typically understood.

In Kant's famous categorical imperative, the moral worth of an action (and the moral law) may be determined by whether its maxim — a simple statement of the principle of the action — can be willed as a universal law. If the maxim can stand as a universal law for all to follow and does not lead to logical contradiction in the process of universalization, then the maxim is lawful and it becomes a duty of the person to act in accordance with it. For Kant, moral worth, or the *good*, is determined simply and solely by the *lawfulness* of the maxim of the action, that is, its ability to be applied universally and without contradiction, and duty is the necessity to act in conformity to this law which one has legislated for oneself. Thus, autonomy derives from each person's ability to formulate for themselves the moral law that applies to everyone, since a law that applies only to some is not universal. It is by virtue of the light of reason in each person that they have moral freedom and moral responsibility, since it is through our reason that we are able to conceive and indeed legislate the moral law.

This capacity — autonomy — is the *highest good* according to Kant, for it is the condition of all other goods, which are of relative worth. Only *persons* may be accorded this *respect*, all other goods and values being merely *things*. "Things," according to Kant, have a relative price and can be exchanged with one another, but persons are accorded a *dignity*, that is, a worth that is beyond all measure and comparison. Therefore persons cannot be bartered or traded for something of equivalent value, for each person has the capacity to legislate the moral law. Consequently the *value* of autonomy cannot be exchanged, balanced, weighed against, or superceded by other values, for instance, social benefit.

Insofar as individuals determine the moral good through their self-legislation of the moral law, and thereby determine themselves to action, it follows that each individual sets his or her own moral *ends* or purposes. Kant

thus offered an alternative formulation of the categorical imperative which is most helpful in this context: "Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end." "Rational nature," writes Kant, "exists as an end in itself" (100, p. 96). Respect for persons is the necessity or duty to always treat others as ends in themselves, that is, as persons who set their own goals and purposes and who in their actions determine the moral law.

Kant recognized that in everyday interactions we are constantly involved in relations with others through which we are reciprocally means to each others' ends. The force of this imperative is to stipulate that in these relationships we always *also* respect each other as self-determining persons — autonomous — that is, as persons who set our own goals. Thus in our everyday activities it is wrong to use other persons for purposes (ends) that they do not share. To do so is to treat other persons merely as *means*, that is, as mere *things*, and thus to violate their dignity as persons (101).

There are three conclusions to be drawn from Kant's argument. First, there is no claim here to a moral *right* of autonomy. Rather, duty derives from the rational capacity of each agent. Respect for autonomy is a *duty* of the researcher who seeks to use human subjects as a means to the production of scientific knowledge, just as it is a duty of every individual to always treat others as ends, even as we use each other as means. Only through such respect can actions have moral worth. This duty is the necessity of the researcher to submit their purposes to the dignity and sovereignty of the person who is the research subject, which is to say, it is their submission to the moral law itself. Moral duty thus does not derive from a "right," the right of the other to determine their own ends, but rather duty derives from the moral law itself. By starting with the premise that autonomy is a right that simply represents important moral *interests*, the consequentialist interpretation of autonomy fundamentally loses sight of the deontological dimension of the concept, that duty depends on the moral law. Autonomy may indeed represent certain "interest," but these are not the basis of the deontological concept of autonomy and the duty to respect others as ends in themselves.

Second, although other goods may be achieved through failure to respect the dignity of persons and the moral law, it is the goodwill of each rational agent that is the highest good and prerequisite of all other goods. To sacrifice this highest good, from which all other relative goods derive their value, for the sake of a dependent and relative good is to undermine the very good that one would hope to achieve in the action. It is self-defeating. From this deontological point of view, the respect for autonomy has nothing to do with the consequences of action, good or bad, but is rather the condition for achieving any good at all. Social good is not an end in itself but rather is a relative good, or means, to the realization of the highest good, the achievement of morally good action. The fundamental claim Kant is making is that there can be no good at all, including social good, without moral freedom, or autonomy.

In usurping this freedom by failing to inform and ask the consent of patients for the use of their tissues and records, we limit the moral freedom of individuals and thereby diminish the possibility of realizing the good to be achieved through the social benefits of the research. It is the problem of putting the cart (social good) before the horse (individual moral freedom). Individuals whose privacy has been violated are indeed *wronged* as Capron argues. The wrong is a violation of their moral freedom and consequently a violation of their *dignity*. The pejorative *abstract autonomy* to which Reilly refers, and which the consequentialist interpretation produces in this context, fails to recognize that the very possibility of moral freedom and responsibility is at stake when we fail to respect the autonomy of others.

Third, it is necessary that research subjects, in order that they not be treated *merely* as means, be informed of the purposes of the research and freely accept these purposes as their own in their participation in the research protocol. This is the insight for which Veatch argues so strenuously. It is also the basis of the National Commission's definition of autonomy. It follows, as Veatch argued, that the exemption from, and waiver of, informed consent is immoral, since only through informed consent is it possible to respect the autonomy of the persons who are subjects of research. Failure to inform subjects and ask their consent for participation is to treat them merely as *means* to ends which they may or may not share. It is to treat persons as *things* which lack dignity.

Autonomy, understood in this light, conflicts with the consequentialist interpretation which would subsume autonomy as a species of beneficence. On the contrary, the deontological interpretation of autonomy places consequentialist reasoning as subordinate to and derivative from the concept of good which autonomy determines. This good is incommensurate with the relative goods to be achieved through the consequences of particular actions. For the deontologist, arguments concerning the increased social benefits to be derived from research simply miss the point or end up subverting the concept of autonomy. Beneficence concerns are legitimate but not overriding.

It can readily be seen that the absoluteness of the duty to respect autonomy in the deontological interpretation is behind some of the arguments for requiring the informed consent which we cited at the beginning of this section.

1. The ethical foundation of informed consent is a respect for individual autonomy, and permitting research on tissues and records without informed consent violates a growing tradition in research ethics that has served patients and subjects very well. We have demonstrated the independent validity of autonomy as an ethical underpinning of the requirement for informed consent, and the National Commission gave a deontological interpretation of autonomy. Although the independence of autonomy has emerged through the historical rise of government regulation in human subjects research, it has served as an important and powerful check against researchers who would privilege the social benefits of scientific knowledge over the free will of research

subjects, which in the past has led to disasterous consequences.

2. Social benefits are not a sufficient ground to undermine the respect for autonomy by not requiring informed consent. This is the heart of the deontological argument which finds that benefits and harms in general are incommensurate with deontological principles.
3. As a result of the successes of the Human Genome Project, human tissues contain a vast amount of medical information that can be tapped with increasing ease. There was a time when tissues represented very little of consequence, either directly to patients or indirectly by virtue of the purposes to which researchers might put them. But the production of scientific knowledge is less of an unequivocal good now. Not all research subjects may share the research purposes of some or even many protocols. Thus the moral freedom of individual subjects is at stake when the research community argues for less stringent requirements for informed consent.
4. Public trust in the health care industry and in biomedical research is undermined when researchers use individuals' tissues without their consent or knowledge for purposes they may not share, and especially when researchers profit financially from them. This argument, as we saw above, may be grounded on a consequentialist analysis, but it also may be grounded on the deontological requirement not to involve other persons in purposes that they do not share or have not accepted. Especially when research is increasingly motivated by profit, the presumed altruism that has played a role in justifying researcher access to tissues and records without informed consent is undermined. This presumed altruism is really the presumption that individuals *share* the goals of biomedical research. But the ulterior motives that also animate research to an increasing degree cannot so easily be presumed to be shared by the general public.

The idea of patient altruism has a correlate in the idea that patients who benefit from biomedical advancement owe a *debt* to past research subjects and thus have an *obligation* to contribute to the future advancement of biomedical science. One relatively easy way for patients to discharge this obligation is to provide easy access to tissues and medical records for research purposes. However, both Caplan (102) and Jonas (68) have argued that while we may owe some debt of gratitude to past subjects of biomedical research, those subjects cannot possibly have considered that this debt would be discharged through violating the autonomy of subjects that follow, though they may well hope that others would make similar contributions to the common good. Furthermore, were this debt one that the research community could *exact* from the population of patients, there would be no reason why the requirement to participate in research would not extend to all biomedical research, including the testing of invasive procedures and drugs. Contributing to biomedical progress is like the general moral obligation to be charitable. We do indeed have an obligation to be charitable, but charity must be given freely and according to the individual's own view of the good to be achieved thereby. To require that patients contribute their tissues and records to biomedical research amounts to a *tax* on patients, in much the same way that welfare programs for the poor are funded by a tax on the public. It would be a serious mistake to refer to one's tax contributions to welfare programs as charitable, and it is no less a mistake to presume that the health care industry can collect on patient debts to medical progress by collecting their tissues without consent.

The last argument for requiring informed consent derives from the fact that our current health care system does not distribute its benefits equally across the population. Many minorities and the poor are excluded or underserved. The justification that the violation of individual autonomy by the claim that research on tissues serves the public good fails to recognize that these benefits do not return equally to the individual sources of tissues and records. This argument is founded on neither beneficence nor autonomy but the independent principle of justice. Justice in this context concerns the *fair* distribution of benefits and burdens and relates to the previous argument concerning the obligations of patients to participate in research. Clearly, those who have not shared in the benefits of biomedical research as fully as others cannot be said to have an equal obligation to that system, especially when the inequities of our health care system are largely the result of a conscious political choice not to remedy them. Autonomy does play a role here insofar as those who have received less than equal share in the benefits of biomedical science and treatment may no longer share the goals of this industry, at least, until they and others are included as full partners.

CONCLUSION

Having formulated the contrast between the consequentialist and deontological interpretations of autonomy so starkly, or what amounts to the same thing, the contrast between beneficence and autonomy, we must be careful how we proceed. It would be tempting, perhaps, to simply *weigh* these two principles against each other. But this would beg the question, for we must have some criteria by which to evaluate these competing principles, a common scale as it were, and by definition there can be no such common scale. Marshall (98) has argued that the National Commission fell prey to this error by viewing autonomy, beneficence, and justice as principles to be *weighed* or *balanced* against each other when they lead to contradictory conclusions. The metaphor of *balancing* implies that the contradictory terms are commensurate. But as Marshall concludes "...the point of the Kantian principles is precisely to say that certain things cannot be 'balanced out'" (98, p. 6). The metaphor of "weighing" and "balancing" is primarily a utilitarian strategy of reckoning up the goods and ills of consequences. To suggest that a

deontological principle may be balanced against a consequentialist principle is to eviscerate the deontological core of the concept, and thus begs the question.

We are left therefore with incommensurate principles which, when applied to research involving tissues and records, lead to contradictory conclusions. Each principle appears to assert its priority over the other. Either we must sacrifice autonomy as an independent deontological principle if we wish to privilege beneficence in justifying not requiring informed consent, or we must accept the deontological principle of autonomy as taking precedence over beneficence considerations, and thus sacrifice some measure of the common good to be achieved through the use of tissues and records without informed consent. Notice that in this dilemma, privileging beneficence over autonomy does violence to its deontological interpretation, while recognizing the priority of autonomy does not in any way corrupt the principle of beneficence. Consequentialist reasoning is inherently blind to its deontological counterpart, while from the deontological point of view, there is no inherent contradiction in reckoning up consequences, provided only that in doing so we conform to our moral duty. If we are to accept the deontological interpretation, the inevitable conclusion is that autonomy is an independent moral principle that cannot be subsumed under beneficence.

If public policy is then to recognize the independent validity of autonomy, we must reconsider the exemptions and waiver of informed consent that the current regulations permit. This does not necessarily mean that we must default to always requiring the full informed consent of subjects for the use of their tissues and records in research. As we have seen, the National Commission was reluctant to give up the social benefits of scientific research that would be lost if specific informed consent were required for all uses of tissues and records. But recognizing that individual autonomy could not simply be sacrificed at the altar of social expediency, the commission recommended two provisions that would preserve the respect for autonomy. First, it required IRB review of all research involving human tissues and records, and second, it required that institutions notify all patients that these materials may be used for research and provide patients with a blanket consent to opt out if they so chose. The requirement of IRB review ensures the social acceptability and value of the research and thus stands in as proxy, or "constructed" consent, for subjects who do not specifically consent to these protocols. The requirement for blanket consent recognizes, however, the sovereignty of the autonomous subject to choose whether to participate in such research activities. Clearly, this compromise crafted by the National Commission limits the exercise of individual autonomy, but at the time it was perhaps much less likely that subjects would object to the research use of their tissues.

Research involving tissues, however, has changed dramatically since the late 1970s when the National Commission issued its recommendations. The revolution in genetics and biotechnology in general raises a host of new risks to subjects, and the erosion of privacy in our evolving information society and within health care itself has made the public more aware and more cautious in guarding privacy. It may be that blanket consent is no longer an adequate expression of respect for the autonomy of patients. Robert Weir's concept of an advance directive for tissues would go a long way toward returning to patients and subjects the control that is an expression of their self-determination. It is at least worthwhile exploring the feasibility of instituting such a system.

If we are to pay more than lip service to autonomy, it will be necessary for IRBs to pay close attention to the scientific necessity of the waiver of informed consent. The current categories of exemption should be abandoned in favor of IRB review in which IRBs may decide that the waiver of the informed consent requirement is necessary to accomplish the goals of the research protocol, and not just a matter of inconvenience to the investigators. The current exemption categories permit investigators to use human tissues and records for virtually any purpose without the knowledge and consent of the sources. While most such purposes are generally laudable and serve the public good, this may not always be the case. If the waiver of informed consent is genuinely necessary to conduct the research, then the research community owes it to the individual sources of tissues and records to submit the research to an independent evaluation to ensure that the research is not foreseeably objectionable. This does not provide for those who may have particular objections, but those who are concerned not to have their tissues and records used for purposes to which they might object would have the opportunity to opt out of participation ahead of time through a blanket refusal of consent or an advance directive prohibiting the research use of tissues and records.

In order for IRBs adequately to take on this increased responsibility, the conditions for waiver of informed consent need to be tailored specifically to research involving tissues and records. The ambiguities in the waiver criteria analyzed in the second part of this article should be clarified, and specific criteria relating to tissues and records should be made explicit. The questions concerning risks, "rights and welfare," ownership and commercial use, recontact of subjects and especially the "practicality" of research, that is, the scientific necessity of the waiver, need to be specifically tailored to research involving tissues and records.

Second, IRB membership, as noted in the Inspector General's report on the reform of IRBs (10), needs to be broadened with additional noninstitutional and nonscientific members representative of community interests so that evaluations of the social value or objectionability of research conducted without specific informed consent can be more representative of the local populations. IRBs also need greater expertise in the areas of genetics and biotechnology in order to be in a position to recognize and analyze the many issues these protocols may raise.

These modifications to the current system of protections would ensure that both autonomy is respected and that

the benefits of scientific research may accrue through as minimal interference with the research enterprise as is consistent with this respect for autonomy.

BIBLIOGRAPHY

1. G.J. Annas, *J. Am. Med. Assoc.* **270**(19), 2346–2350 (1993).
2. Nuffield Council on Bioethics, *Human Tissue: Ethical and Legal Issues*, Nuffield Council on Bioethics, London, 1995, p. 19.
3. E. Eiseman, *Stored Tissue Samples: An Inventory of Sources in the United States*, in Research Involving Human Biological Materials: Ethical Issues and Policy Guidance; Volume II, Commissioned Papers; National Bioethics Advisory Commission, Rockville, MD, 2000, p. D-5.
4. U.S. Pat. 5,409,810 (December 1, 1992), B.A. Lardar and S.D. Symons (to the Burroughs Wellcome Co.).
5. B.A. Lardar, K.E. Coates, and S.D. Kemp, *J. Virol.* **65**(10), 5232–5236 (1991).
6. National Bioethics Advisory Commission, *Research Involving Human Biological Materials: Ethical Issues and Policy Guidance; Volume I, Report and Recommendations of the National Bioethics Advisory Commission*, National Bioethics Advisory Comm., Rockville, MD, August 1999, p. 1.
7. C.R. Naser, in A. Thompson and R. Chadwick, eds., *Genetic Information: Acquisition, Access and Control*, Kluwer Academic, Plenum, New York, 1999, pp. 105–120.
8. U.S. Department of Health and Human Services, Office of Protection from Research Risks, *Federal Policy for the Protection of Human Subjects*, Title 45, Code of Federal Regulations, Part 46, Subpart A. All citations to this policy will appear in the text by section number, for example: "§46.101(b)(4)."
9. U.S. Department of Health and Human Services, Office of Protection from Research Risks, *Fed. Regist.* **56**(117), 28003–28032 (1991).
10. U.S. Department of Health and Human Services, Office of Inspector General, *Institutional Review Boards: A Time for Reform*, OEI-01-97-00190; USDHHS, Washington, DC, 1998, p. B-1.
11. U.S. Department of Health, Education, and Welfare, *Fed. Regist.* **39**(105), 18914–18920 (1974).
12. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, U.S. Government Printing Office, U.S. Department of Health, Education and Welfare, Washington, DC, 1978.
13. U.S. Department of Health, Education and Welfare, *Fed. Regist.* **44**(158), 47688–47734 (1979).
14. U.S. Department of Health and Human Services, *Fed. Regist.* **46**(16), 8366–8392 (1981).
15. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *Institutional Review Boards: Report and Recommendations*, U.S. Government Printing Office, Washington, DC, 1978, p. 21.
16. Personal communication from J.T. Puglisi, Ph.D., Director, Division of Human Subjects Protections, Office for Protection from Research Risks, March 28, 1997.
17. V.A. LiVolsi et al., *Cancer (Philadelphia)* **71**(4), 1391–1394 (1993).
18. J.F. Merz et al., *J. Invest. Med.* **45**(5), 252–257 (1997).
19. G.B. Ellis, Director, Office of Protection from Research Risks, *Memorandum: Response to Frequently Asked Questions*, April 17, 1996, Chart 2.
20. A.R. Holder and R.J. Levine, *Clin. Res.* **24**(2), 68–77 (1976).
21. J. Merz, *IRB* **18**(6), 7–8 (1996).
22. E.W. Clayton et al., *J. Am. Med. Assoc.* **274**(22), 1786–1792 (1995).
23. E.W. Clayton, *J. Law, Med. Ethics* **23**(4), 375–377 (1995).
24. K.T. Kelsey et al., *J. Am. Med. Assoc.* **275**(14), 1085–1086 (1996).
25. R.J. Levine, *Ethics and Regulation of Clinical Research*, Urban & Schwarzenberg, Baltimore, MD, 1981.
26. P.R. Reilly, M.F. Boshar, and S.H. Holtzman, *Nat. Genet.* **15**(1), 16–20 (1997).
27. L.J. Melton, *N. Engl. J. Med.* **337**(20), 1466–1470 (1997).
28. D. Korn, *Genetic Privacy and the Use of Human Tissues in Research, Presented at Risk, Regulation, and Responsibility: Genetic Testing and the Use of Information*, American Enterprise Institute, Washington, DC, 1997. Available at: *http://www.aei.org*
29. Personal communication from J. Cheek, Ph.D., Principle Epidemiologist for the Indian Health Service, June 1998.
30. Privacy and Protection of Research Subjects: *Certificates of Confidentiality*, Information sheet available from PHS through the NIH Office of Resource Management. (Investigators may obtain a Certificate of Confidentiality by applying to the NIH. To receive a complete information packet, call 301-443-3877.)
31. C.L. Earley and L.C. Strong, *Am. J. Hum. Gene.* **57**(3), 727–731 (1995).
32. W.J. Curran, in P.A. Freund, ed., *Experimentation with Human Subjects*, George Braziller, New York, 1970, pp. 402–454.
33. R.R. Faden, T.L. Beauchamp, and N.P. King, *A History and Theory of Informed Consent*, Oxford University Press, New York, 1986, p. 208.
34. R.S. Dresser, *IRB* **3**(4), 3–4 (1981).
35. R.M. Veatch, *The Patient as Partner: A Theory of Human-Experimentation Ethics*, Indiana University Press, Bloomington and Indianapolis, 1987, p. 43.
36. Moore v. The Regents of the University of California, 215 Cal. App. 3d 709, 249 Cal. Reptr. 494 (1988), *aff'd in part, rev'd in part*, 51 Cal. 3d 120, 793, 271 Cal. Reptr. 146 (1990).
37. R.F. Weir and J.R. Horton, *IRB* **17**(5,6), 1–8 (1995).
38. K.C. Laudon, *Dossier Society*, Columbia University Press, New York, 1986.
39. J.B. Rule, *Newsday*, July 30, p. A40 (1997).
40. T.H. Murray, *IRB* **8**(1), 1–5 (1986).
41. E.R. Gold, *Body Parts: Property Rights and the Ownership of Human Biological Materials*, Georgetown University Press, Washington, DC, 1996, pp. 161–162.
42. S. Alpert, *Privacy and the Analysis of Stored Tissues*, in Research Involving Human Biological Materials; Ethical Issues and Policy Guidance; Volume II, Commissioned Papers. National Bioethics Advisory Commission, Rockville, MD, 2000.
43. L. Gostin, *Law, Med. Health Care* **19**(3–4), 191–204 (1991).
44. L.M. Kopelman, *J. Med. Philos.* **19**(6), 525–552 (1994).
45. L.H. Glantz, *IRB* **1**(6), 5–6 (1979).
46. H. Greely, *Houston Law Rev.* **33**, 1397–1430 (1997).
47. North American Regional Committee of the Human Genome Diversity Project, *Houston Law Rev.* **33**, 1431–1473 (1997).

48. U.S. Department of Health Education and Welfare, *Fed. Regist.* **43**(231), 56174–56198 (1978).
49. K.J. Rothman, *N. Engl. J. Med.* **304**(10), 600–602 (1981).
50. W.W. Grody, *Diagn. Mol. Pathol.* **4**(3), 155–157 (1995).
51. O. O'Neill, *J. Med. Ethics* **22**, 5–7 (1996).
52. E.G. Knox, *Br. Med. J.* **304**, 727–728 (1992).
53. P.R. Reilly, *J. Law, Med. Ethics* **23**, 378–381 (1995).
54. L. Andrews and D. Nelkin, *Lancet* **351**, 53–57 (1998).
55. E. Marshall, *Science* **271**, 440 (1996).
56. G.J. Annas, in T.F. Murphy and M.A. Lappé, eds., *Justice and the Human Genome Project*, University of California Press, Berkeley, 1994, pp. 75–90.
57. P.S. Harper, *Br. Med. J.* **306**, 1391–1394 (1993).
58. J. Benbassat and M. Levy, in P. Allebeck and B. Jansson, eds., *Ethics in Medicine: Individual Integrity Versus Demands of Society*, Raven Press, New York, 1990, pp. 159–165.
59. Privacy Protection Study Commission, *Personal Privacy in an Information Society*, U.S. Government Printing Office, Washington, DC, 1977, pp. 596–598.
60. LeR. Walters et al., *IRB* **3**(3), 1–3 (1981).
61. A.R. Holder, *IRB* **3**(3), 4–5 (1981).
62. G. Kolata, *N.Y. Times*, October 24, pp. A1, C6 (1995).
63. M.Z. Pelias, *Am. J. Med. Genet.* **39**, 347–354 (1991).
64. V.L. Hannig, E.W. Clayton, and K.M. Edwards, *Am. J. Med. Genet.* **47**, 257–260 (1993).
65. M. Feinleib, *J. Clin. Epidemiol.* **44**(Suppl. I), 73s–79s (1991).
66. E.D. Pellegrino, in H.T.J. Engelhardt, S.F. Spicker, and B. Towers, eds., *Clinical Judgment: A Critical Reappraisal*, Reidel Publ., Boston, MA, and Dordrecht, The Netherlands, 1979, pp. 169–194.
67. K. Lebacqz, *Villanova Law Rev.* **22**, 357–366 (1977).
68. H. Jonas, *Daedalus* **98**(2), 219–247 (1969).
69. National Health and Nutrition Examination and Survey (NHANES), conducted by the Centers for Disease Control, Atlanta, GA, July 1994.
70. K. Steinberg, in R. Weir, ed., *Stored Tissue Samples: Ethical, Legal, and Public Policy Implications*, University of Iowa Press, Iowa City, 1998, pp. 82–88.
71. C.R. Naser, in R.F. Weir, ed., *Stored Tissue Samples: Ethical, Legal, and Public Policy Implications*, University of Iowa Press, Iowa City, 1998, pp. 160–181.
72. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Screening and Couseling for Genetic Conditions*, U.S. Government Printing Office, Washington, DC, 1983.
73. D.L. Breo, *J. Am. Med. Assoc.* **269**(15), 2017–2022 (1993).
74. K.F. Hoskins et al., *J. Am. Med. Assoc.* **273**(7), 577–585 (1995).
75. A. Clarke, *Lancet* **338**, 998–1001 (1991).
76. F.M. Giardiello et al., *N. Engl. J. Med.* **336**(12), 823–827 (1997).
77. G.G.S. Singer, in B. Gert et al., eds., *Morality and the New Genetics: A Guide for Students and Health Care Providers*, Jones & Bartlett International, London, 1996, pp. 125–146.
78. M. Decruyenaere, G. Evers-Kiebooms, and H. Van den Berghe, *J. Med. Genet.* **30**, 557–561 (1993).
79. M. Huggins et al., *Am. J. Hum. Genet.* **47**, 4–12 (1990).
80. U.S. Department of Health and Human Services, *Fed. Regist.* **46**(16), 8392 (1981).
81. U.S. Department of Health and Human Services, *Fed. Regist.* **63**(216), 60364–60367 (1998).
82. G.J. Annas, L.H. Glantz, and P.A. Roche, *The Genetic Privacy Act and Commentary*, Boston University School of Public Health, Boston, MA, 1995.
83. T. Murray, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, and London, 1997, pp. 60–73.
84. R.F. Weir, in R.F. Weir, ed., *Stored Tissue Samples: Ethical, Legal, and Public Policy Implications*, University of Iowa Press, Iowa City, 1998, pp. 236–266.
85. B.M. Dickens, *Law, Med. Health Care* **19**(3–4), 175–183 (1991).
86. R.J. Levine, *Hastings Cent. Rep.* **9**(3), 21–26 (1979).
87. S.D. Warren and L.D. Brandeis, *Harv. Law Rev.* **4**, 193 (1890); reprinted in F.D. Schoeman, ed., *Philosophical Dimensions of Privacy: An Anthology*, Cambridge University Press, Cambridge, MA, 1984, pp. 75–103.
88. C.J. Bennett, *Regulating Privacy: Data Protection and Public Policy in Europe and the United States*, Cornell University Press, Ithaca, NY, 1992, p. 13.
89. S.A. Alpert, *Santa Clara Comput. High Technol. Law J.* **11**(1), 97–118 (1995).
90. G. Duncan, T. Jabine, and V. de Wolf, *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics*, National Academy Press, Washington, DC, 1993, p. 27.
91. A. Allen, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 31–59.
92. A.F. Westin, *Privacy and Freedom*, Atheneum, New York, 1967.
93. V.M. Brannigan and B. Beier, *Datenschutz und Datensicherung*, April 1985.
94. A. Buchanan, *An Ethical Framework for Biological Samples Policy*, in Research Involving Human Biological Materials: Ethical Issues and Policy Guidance; Volume II, Commissioned Papers, National Bioethics Advisory Commission, Rockville, MD, 2000, p. B-4.
95. M.B. Visscher, *Mod. Med.*, pp. 62–64 (Feb. 15, 1975).
96. Harris-Equifax, *Health Information Privacy Survey 1993*, Study No. 934000009, Harris-Equifax, New York, 1993, p. 45.
97. K. Lebacqz and R.J. Levine, *Clin. Res.* **25**(3), 101–107 (1977).
98. E. Marshall, *IRB* **8**(6), 5–6 (1986).
99. A.M. Capron, *J. Clin. Epidemiol.* **44**(Suppl. I), 81s–89s (1991).
100. I. Kant, *Groundwork of the Metaphysics of Morals* (H.J. Paton, trans.), Harper & Row, New York, 1956.
101. G. Dworkin, in T.L. Beauchamp et al., eds., *Ethical Issues in Social Science Research*, Johns Hopkins University Press, Baltimore, MD, 1982, pp. 246–254.
102. A.L. Caplan, in S.F. Spicker et al., eds., *The Use of Human Beings in Research*, Kluwer Academic Publishers, Boston, MA, 1988, pp. 229–248.

See other entries GENETIC INFORMATION, ETHICS, PRIVACY AND CONFIDENTIALITY: OVERVIEW; GENETIC INFORMATION, LAW, LEGAL ISSUES IN LAW ENFORCEMENT DNA DATABANKS; GENETIC INFORMATION, LEGAL, GENETIC PRIVACY LAWS; PATENTS AND LICENSING, ETHICS, MORAL STATUS OF HUMAN TISSUES: SALE, ABANDONMENT OR GIFT.

GENETIC INFORMATION, ETHICS, FAMILY ISSUES

ERIC T. JUENGST
Case Western Reserve University
Cleveland, Ohio

OUTLINE

INTRODUCTION

Of all the many "publics" that are affected by advances in genetics, families that use genetic services are the closest to our society's grass roots: They cut across all other sociological categories and lie behind all the usual interest groups that contend over our society's health policies and practices. Moreover these families and their members have always been portrayed by the biomedical community as the focus of the genetic services it generates. In the absence of effective therapies, the promise of accessible genetic information lies almost entirely in its ability to allow families to identify, understand, and sometimes control their inherited health risks. Against the excesses and abuses of the eugenicists' population-oriented concerns, contemporary geneticists are firm in their conviction that "the fundamental value of genetic screening and counseling is their ability to enhance the opportunities for individuals to obtain information about their personal health and child-bearing risks, and to make autonomous and non-coerced choices based on that information" (1). This puts families at the moral fulcrum of the enterprise: If genetic services are to be judged a success, it must be from the recipients' point of view, in terms of their ability to use the results to support their flourishing as individuals and families. That, in turn, gives a special urgency to getting clear about the impact of genetic services on family life, and the ethical issues they can raise for family members. This article summarizes what is known about these "family matters" in genetics and suggests an agenda for futher research in this crucial area.

FAMILY VIRTUES AND GENETIC TESTING

Over the last decade a large literature has evolved anticipating and addressing the ethical, legal, and social implications of advances in human genetics. For overviews, see Juengst (2) and Thomson (3). One of the weakest spots in that literature, however, is work addressing the ethical issues faced by the individuals and families who might avail themselves of the fruit of those advances. This relative neglect is not entirely surprising, despite the centrality of the issues to the success of the genetic enterprise. First of all, most of our efforts have been directed toward the development of policy capable of optimizing our uses of genetic advances: rules, guidelines, agreements, and positions that can be generalized across all the situations that raise the issues they address. That is much easier to accomplish for entities like states, institutions, and professions than it is for "the family," since our many different families subscribe to no unified process for making universally binding "family policy" on ethical issues. Moreover there is no well-received generic account of the moral dynamics of family life to draw upon in even attempting to develop such policies. With some recent exceptions (4,5), the ethics of family interactions has been a black box for contemporary applied ethics, protected from intellectual scrutiny as well as state intervention by our liberal traditions. In part, this is because there is such a rich pluralism of strongly held specific theories on the subject, which reflect the convictions and experiences of different family histories and traditions (6). Since these specific theories are usually intertwined tightly with our most important beliefs and values, contemporary moralists tend to give them the deference that we give to other transrational matters of conscience, like personal religious commitments. Against that variegated and sensitive background, the thinking goes, one would be foolish to try to generalize about what ethical issues any new genetic service might raise for families outside of some parochial cultural perspective (7).

Be that as it may, the existing literature does suggest three ways in which advances in genetics will challenge the ethics of at least many American families. Despite our differences, there are three familial virtues that most socialized Americans would not be surprised to find listed among the qualities that have traditionally been ascribed to the "good family" in our culture. I will call these qualities the familial virtues of loyalty, intimacy, and security. While these three virtues are neither necessary nor sufficient to an adequate ethical theory of family life, they are keys to the familial ethics of genetic testing, because they are value commitments that do seem particularly challenged by our emerging genetic testing practices. To the extent that these virtues are at least widely intelligible as ingredients in the moral dynamics of family life, their analysis provides a starting point for further discussion. New advances in genetics will challenge the ability of the families who endorse them to live up to the ideals that they represent.

Loyalty

The members of good families accept special obligations to serve their kin. Whether this means grown children joining their parents' businesses, siblings helping each other out of debt, or cousins hosting visiting cousins, members of a good family are expected to aid and assist one another without having made any of the explicit promises,

offers, or agreements to do so that would govern such service between strangers. Moreover this familial loyalty can supersede the interests of individual family members in significant ways: when mothers delay careers to raise unexpected children, for example, or when children's educational interests are sacrificed in order to care for their infirm grandparents, or when siblings come home for the holidays at the cost of their holidays. Of course, like any virtue, familial loyalty can become a vice *in extremis*, as when excessive concern for "family honor" generates vendettas between families, or self-sacrifice becomes unnecessarily self-destructive. Identifying the proper demands of family loyalty and balancing them against our other interests is one of the perennial challenges of moral life within a family. However, the way we approach that balancing problem is itself oblique evidence for the value we give to the virtue. As Nelson and Nelson (5, p. 76) point out, "Moral relationships among family members can certainly be strained by betrayal or violence, but it takes a catastrophe to dissolve them." ... In our culture we understand the vices of extreme loyalty precisely as a problem of overdoing a good thing: we can sympathize with their perpetrators in ways that we cannot with those who abandon or betray their kin. On the whole, we applaud the families that stick together because we see a reciprocal concern for each other's best interests as a constitutive ingredient of what it means to be a good family. For example, Post (4, pp. 81–108) argues that the balance between giving and receiving love in the context of the family is a common moral expectation that is a truly universal intuition.

New advances in human genetics challenge the virtue of familial loyalty in two ways. They simultaneously illuminate both the connections among and the differences between family members. First of all, genetics is, by definition, a science of family connections. Clinically useful genetic information about individuals often requires knowing the background against which the individual's genome presents itself: the pattern of inheritance of the traits and markers in question within the larger family. This means that in order for a genetic test to be useful to an individual family member, other members of the family have to be willing to provide that background and, in the process, discover their own status within the pattern.

Moreover, like most medical interventions, genetic testing usually motivated by a crisis — someone diagnosed with breast cancer or genetic disease — which creates a sense of urgency to "get the family in for testing." For those other family members, the decision to participate in a testing program raises a basic moral question: What are the demands of my loyalty-based obligation to help my kin learn their genetic risks? In particular, must I sacrifice my own "right not to know" in order to help my relative enjoy the "right to know," and join him or her in braving the psychosocial risks of having that personal information known about me? When family members decide to protect their own interests and decline to participate, the same question is passed "downstream" to their children and grandchildren: If those downstream kin should decide to be tested, the status of the declining member could be revealed as a simple matter of deduction. What interests must they sacrifice, then, in order to give the decliner the filial respect that they deserve (8)?

Finally, if a decliner's kin do become interested in learning their own genetic risks, but cannot do so without involving the reluctant relative, to what lengths may the family go to persuade the unwilling to do their familial duty? Split decisions about genetic testing have already been observed by genetic counselors to lead to familial discord in some cases (9), while unanimous decisions in other cases have raised suspicions of undue familial pressure to participate (10).

The moral bite of these questions within families can be seen by the ways in which they seem to be provoking health professionals involved genetic testing to clarify their own allegiances. Despite the long-standing reluctance of clinical geneticists to interfere in the personal choices that their clients make about genetic testing and the use of test results, some now argue that clinicians should help their clients persuade reluctant family members to participate, by reminding them of their familial obligations (11). Others stand by the conviction that each family member has equal standing as a client, and thus, "it is important that individuals within families are supported to make their own decisions about testing and are not coerced by eager relatives" (9, p. 25). Still others now argue that, if the clinical geneticist's "patient" is best understood as the family as a whole, perhaps clinicians have no more business attempting to regulate the influences on the family's collective decision than they would in second-guessing an individual's informed consent. From that point of view, whatever internal dynamics produce the decision are protected from the clinician's interference by the same sphere of familial privacy that grounds their commitment to nondirective counseling at the individual level (12). Just like the family members involved, the professionals are forced by genetic testing to think about the moral significance of the genetic connections that link up the individuals with whom they interact.

Of course, genetic information is as much about our differences as it is about our shared traits, and illuminating those differences is another way in which genetics can challenge the familial virtue of mutual loyalty. As we are able to sort out which lineages within families, and which individuals within a lineage, carry a family's risk-conferring mutations, tension will be created between the divergent interests of the two groups. Whatever their commitment to family solidarity, the family will have to face the fact that it will be in its "normal" "members" interests to reveal their noncarrier status in some circumstances and in the interests of the carriers to conceal theirs. Family members free of the mutations in question, for example, will find it in their interests to use that information to counter their family history of a disease in applying for insurance (13). In doing so, however, they will inevitable raise questions about their kin who do not volunteer their test results in turn. Should families be expected to stick together "in sickness and in health" as we ask of married couples, or do the "limited sympathies" of human nature give us leave to concentrate on the welfare of our own threads within the familial patterns of inheritance?

For families in our society, market institutions like risk-rated health and life insurance that create serious competitive advantages out of genetic differences make it increasingly difficult for even conscientious families to present a united front on the matter of disclosing test results. As a spokesman for the insurance industry recently wrote, for high risk families faced with gaining access to adequate health and life insurance, "Harsh as it may sound to ears of a society that subscribes to egalitarian principles, solidarity ends with a negative genetic test" (14). Since, short of cloning, human families will always weave together a combination of different genetic threads, new abilities to identify those differences will continue to expose families to this kind of external pressure, as long as we live in a society that uses such differences to allocate its opportunities.

Intimacy

A second traditional virtue of family life is intimacy. As our concept of "family secrets" suggests, we expect family members to communicate with each other about private matters more readily than they do with their neighbors, coworkers, or even friends. Family members are allowed to know more about each other, and expected to share important personal information with each other (births, deaths, marriages) before they tell the rest of the world. Again, it is possible to see this virtue reflected in the way we treat its vicious distortions. Thus excessive communication (i.e., gossip) within families is a recognizable moral problem for family life, but we are usually more concerned about the keeping of secrets between close family members, particularly when the secret bears on their own relationship. In theory, disclosures of behavior that would be scandalous in public can be safely made within families not only because the secret is safe there but also because the expected reaction is not a moral judgment; it is rather an attempt at understanding. Intimacy in family life is not only about exposing our vulnerabilities to one another; it also requires taking no offense at the disclosures.

Unfortunately, despite our commitments to intimacy, information about shared genetic health risks seems particularly hard to share with our kin. For families that are already aware that they have a hereditary history with a particular disease, revealing that one is a mutation carrier can invite stigmatization as a bearer of the "family curse," and premature assignment to the sick role. On the other hand, "survivor guilt" seems to be relatively common among those testing negative in these families, and it also inhibits disclosure (15)? In these "at risk" families, genetic test results place individual members on one side or the other of the "watchful waiting" that characterizes their family's corporate identity and, in the process, segregates them from one of the central dramas of their family's life. For example, Alice Wexler explains her own decision to remain uncertain about whether she carries the gene for Huntington disease this way:

> One man who tested negative felt as if he were missing a limb, a part of his identity. "I had lost my creative terror" he said. . . . Choosing not to take the test is a decision one can easily revoke, unlike the situation after testing. A Geri Harville said, "You and your family will be affected by this information forever. Once you have the information, you cannot give it back." I have made my peace, more or less, with uncertainty. . . . Perhaps I even enjoy the ambiguity, resisting sharp categories and binary definitions, the border guards insisting that we place ourselves in one camp or the other. (16, p. 238)

Clearly, genetic test results can fall in with mis-identified paternity as among the hardest kinds of personal secrets to share with kin: secrets that seem to deny the very kinship that licenses their intimate disclosure.

On the other hand, genetic test results are also almost always "about" more than the individual who undertakes them. For some, this overlap serves to strengthen the obligation to share genetic test information, because they understand themselves to be privy to a secret about their at risk kin (17). For members of families that do not have a recognized history of the genetic risk, their test reveals this poses another kind of challenge to familial intimacy. In these situations the disclosure does not segregate the source from kin but rather exposes a new and troubling connection between them. Here, feelings of shame and the urge to shield the family from harm seem to combine to encourage keeping genetic risk information secret from kin (18).

These pressures raise for family members the question: What are the limits of my intimacy-based obligations to share my genetic test results? Can my personal sphere of privacy extend to information that is also about my relatives without becoming secrecy? Of course, the problem with keeping specific secrets from other family members is that families are particularly hard organizations to keep individual secrets in without leaving the family altogether. Yet, when secrets have been kept and do emerge, it immediately calls into question the true intimacy of the relationships involved, with potentially devastating effects. Thus Alice Wexler writes that, in the wake of her mother's diagnosis with Huntington disease, and its exposure of her long-held secret about her family history,

> [W]hat my sister and I thought we knew about our family suddenly shifted and everything had to be rethought, reinterpreted. Who we were had suddenly been called into question, and everything had to be reconfigured taking into account the presence of the disease. It was as if we had been experiencing fallout from some unseen bomb for all these years, and suddenly the great mushroom cloud had come into view and we could see the source of all that radiation (16, p. 75).

Looking to law and social policy for indications of social consensus on these questions is only partially helpful. For example, the fact that there are laws that ensure adoptees access to genetic information about their birth parents suggest that we do feel strongly about the obligations of parents, like the Wexlers, to share their genetic health secrets with their children (19). On the other hand, most policy statements on the genetic testing of children stress the need to protect their genetic privacy even in the face of parental requests, suggesting that these intimacy obligations are not always reciprocal (20). Moreover, as Lori Andrews points out in her review of the issue, the law is particularly silent on familial duties outside the

parent-child relationship: "With respect to an individual's duties towards other relatives (such as siblings or cousins), there might be a moral duty to disclose a person's genetic information, but there is probably would not be a legal one" (19, p. 273).

Again, the importance of these ethical questions for families can be seen by the extent to which they bleed over into professional ethical questions for clinicians. Thus clinical experience with patients' reluctant to warn their families of their risks has spawned a renewed discussion of the limits of the professional's commitment to confidentiality in the genetic testing context, with some arguing that the patient's familial obligation to disclose shared risks should translate into a professional duty to warn third parties (21), others standing by the individual proband's authority to make disclosure decisions, and still others arguing that if clinicians take the family seriously as a collective client, they should never make promises of privacy to any particular piece of it but interact with the family as a whole throughout the testing process (22).

Security

Good families are safe places for their members, as the virtues of loyalty and intimacy both suggest. In part, that is because good families accept special obligations to protect their vulnerable members from harm, even when those members cannot ask for help. Thus we expect adult family members to practice a benevolent paternalism with their immature and infirm kin, and we expect our kin to defend us against calumny and slander even when we are unaware of it. Of the three familial virtues, this is the one whose applications and dynamics have been most discussed in biomedical ethics, because it is the one that is called into question in health care when parents are faced with making medical decisions about the welfare of their children (23), and adult children are called to make medical decisions for their aged parents (24). At both ends of life, our bioethical discussions are often struggling to define the limits of people's duties to stand guard over the bodies of their loved ones in particular circumstances, but they almost always start from the premise that they do have those duties.

Like the others, this familial virtue is also thrown into relief by our responses to the vices on its borders. We can sympathize with overly protective parents even as we criticize them as excessive, but we find cultural practices that seem to deny special protections for the vulnerable (e.g., prenatal sex selection) morally callous and hard to understand when they occur in our society. Moreover the strength of our allegiance to this virtue is reflected again in professional ethics and law, with the recognition that "In general, the appropriate presumption is that the family of the incompetent individual is to be the principal decision-maker." Thus it is received wisdom in medical ethics that:

> The chief reasons in support of the presumption that the family is to serve as surrogate decision-maker are both obvious and compelling. The family is generally both knowledgeable about the incompetent individual's good [e.g., intimacy] and his or her previous values and preferences [e.g., loyalty], and most concerned about the patient's welfare. [e.g., security]" (25).

With genetic testing, the limits of this virtue are pressed most vividly by the questions raised by the prospect of testing our offspring. How far should parents go in seeking to protect their vulnerable children from genetic harms? Does prenatal testing and selective abortion count as a "preventive" or "protective" intervention for the children tested, or simply a means of choosing which children to bear (26)? Studies already show that declining to have prenatal testing is perceived as causing the bad outcome when a disabled child is subsequently born, which suggests that many families do think that prenatal screening falls within the protections that parents should afford their offspring (27). This perception, in turn, can put the pressure on couples to be "responsible reproducers" in their own eyes and the eyes of other family members in ways that can seriously compromise their own reproductive autonomy (28). Clinicians, however, have traditionally viewed prenatal testing as enhancing familial rather than fetal security: It is understood as a service to the prospective parents that allows them to reduce the risk of burdening themselves and their family with more than they can handle. This perception of prenatal testing simplifies the professional ethical challenges that the practice poses, but challenges parents to clarify the limits of their ability to nurture children with disabilities, and to define for themselves what will count as an acceptable quality of life for their children (29).

Similar challenges accompany decision making about pediatric genetic testing. Should parents include screening for diseases of late adult onset amongst their obligations to protect the welfare of their children? There has been a lively debate in the literature on this topic, aimed at establishing clinical policies for practitioners (30). Interestingly, however, much of the argument is parental ethics, not professional ethics. Critics of such practices see in them the vice of over-protectiveness and "pre-emptive paternalism," while defenders argue that:

> Children are not discrete monads who develop in isolation until they reach adulthood when they can seize autonomy and begin to make significant decisions about their lives. They are shaped within the context of their family and community as they make their way to adulthood. It is the responsibility of parents to provide for their children's nurture and education. . . . While such socialization can affect the choices children will make as adults, it is considered so much a part of parental responsibility that parents who fail to teach their children a value system can be said to have failed in their duty to them. . . . Yet responsible parenthood would have it no other way. Moreover, refraining from predictive testing is, in effect, teaching children that ignorance is a good way of life. That is at best a controversial message (31).

FAMILY STYLES

Of course, it would be naive to believe that the families in our society that still subscribe to the three familial virtues of loyalty, intimacy, and security all resemble one another. The institution of the family has undergone tremendous change in our society, and it continues to evolve in several directions at once (32). Unsurprisingly, the challenge that

genetic testing will pose to a family's ability to uphold the three virtues will vary with the kind of family involved.

For convenience, consider three of the many kinds of American families that sociologists describe: The sessile organic family, the blended social family, and the diasporadic virtual family. Each of these will feel the challenge of genetic testing differently. Interestingly, however, it is the first, most traditional form of the family, presumably the source of these virtues, in which the challenge may be most severe.

Organic Families

The organic family is the kind of family that best exemplifies what some authors call the "biological" concept of family (33): a multigenerational clan that lives together in geographic proximity, like the Walton's on Walton's Mountain, for example, or the Venezuelan "Huntington disease families" of Lake Maricaibo. In some ways this kind of family might be best situated to benefit from a genetic testing program. Experience with presymptomatic testing for Huntington disease, for example, suggests that in families of "blood relatives" of any particular proband, intrafamilial communication is efficient, and the familial nature of any recurrent health problem will be readily understood. Clinical case studies show that familiarity with the disease in relatives may also improve family members' motivation for testing and diminish the stigmatization of those affected (34).

On the other hand, this kind of family is also more likely to be hierarchically organized, with clear lines of internal authority over decisions that affect the family as a whole. Under this structure the same features that make it possible to support strong family loyalties, intimacies, and protections can also exacerbate ethical problems for individual family subunits. For example, consider the following case report, described in a recent review of the experience of Indiana University's testing program for Huntington disease:

> Mr. Crawford's son married a woman at risk for HD. The couple, now divorced, were married for eight years and had a son and a daughter. These children are at 25 percent risk for HD. The former daughter-in-law has repeatedly said that she is not interested in testing for herself. Mr. Crawford is contemplating setting up a trust fund for his son's children to ensure adequate care for them if they should develop HD. In order to know whether he should make these arrangements, Mr. Crawford wants to have his grandchildren tested (35).

Cases of this sort pose ethical problems for genetic counselors to manage because of the way they challenge the individualistic, "nondirective" ethos that characterizes the profession's ideal. From the counselor's point of view, the principal issue in this case is the appropriateness of testing the children for a late onset disorder like Huntington disease before they are able to give their own autonomous consent to the procedure, and that is how the case is discussed in the review that reports it. But for the family, the scenario raises more fundamental moral questions than the children's ability to consent. Rather, they are more likely to be preoccupied with the relative weight of the various familial loyalties, protections, and intimacies involved. In seeking to improve the security of his grandchildren, has Mr. Crawford overstepped the proper limits of his role? Is it fair to ask the ex-daughter-in-law to risk discovering her own status in order to help Mr. Crawford in his attempt to improve her children's welfare? What role should the history of her divorce from his son play in the deliberations? As this case suggests, families that are organized in ways that facilitate power differentials between members may be most at risk for becoming coercive in pressing members to be tested on behalf of their relatives, and for holding individuals to familial standards of "responsible parenting" that curtail their parental autonomy and procreative liberty. This is also the family structure in which it is most difficult to keep a secret, raising concerns about how to protect the interests of those who choose not to learn their own genetic risks and what information is fair to use in deciding to override them. Moreover all these challenges can emerge insidiously within these families, because they all begin in positive affirmations of the familial virtues. Unlike abandonment, neglect, and silence, the risks of coercion, paternalism, and gossip will be hard to detect and difficult to demonstrate, especially by the less powerful family members who, by definition, will bear the burden of their harms.

Social Families

At the other extreme are families that are the products of serial monogamy, adoption, and new reproductive technologies: the "Brady Bunch," for example. These families exemplify the "household" concept of family: "an aggregate or group of actual (living) members, who are closely associated by living arrangement or by commitment, for better or worse (36). For these families the medical value of genetic testing is lowered by the fact that fewer family members are expected to share the same genetic risks. For these families genetic testing may be less threatening, if only because they are likely to be less attractive. Families seeking to build their identities on ties of love and commitment rather than blood will tend to downplay the importance of shared health problems, which will lower pressure to test "loyalty" through genetic testing. On the other hand, by underlining the differences between the lineages that make up the family, genetic testing does pose another kind of risk for social families: the potential divisiveness that testing could create within the blended family structure between at-risk and low-risk members. For example, consider this case report:

> Angela Smith, a 30-year-old Caucasian woman, came to an outreach clinic for genetic counseling because her husband, David, had two half siblings who died from cystic fibrosis. She was interested in finding out her risk to have an affected child. . . . Angela was tested first; the results showed that she carried the F508 deletion. David had blood drawn for testing. No mutations were detected. Based on this information and family history, the lab calculated his carrier risk to be 1 in 8. His carrier status could potentially be clarified by testing his father, an obligate carrier. Mr. Smith Sr. was tested but no mutation was identified. For further clarification, studies were proposed for David's stepmother, Sue, and his surviving half siblings, Robert, Dale and Karen. If Sue had an identifiable

> mutation which was also present in one of the half siblings, linkage analysis could potentially establish whether David had inherited a chromosome with a high risk of carrying as mutation. Sue, Karen and Robert lived in another state, but agreed to be tested. Sue and Karen had the F508 mutation. Robert was negative for any mutation. Because Robert fell into the same category as David, his sample could not be used to clarify risk. An additional complication arose when Ruth learned that linkage studies indicated that Karen was not Mr. Smith Sr.'s daughter. The family was told that linkage studies were uninformative. If they wished to proceed, a sample would be needed from Dale. The family has made no further efforts to pursue testing (37).

In discussing this case report, its authors focus on issues raised for the genetic counselor by the inadvertent finding of Karen's mis-identified paternity. But behind the counselor's dilemma, the outlines of this blended family's own moral struggle can also be seen. How far should their sense of familial loyalty oblige Sue and her children to go to help her husband's son's wife clarify her reproductive risks? Most of them are willing to be tested, and to endure the inconvenience of out-of-state testing. By agreeing to be tested, Sue even risked exposing the secret of Karen's paternity and the familial harm that might have resulted. Should Dale now also agree to be tested in turn, despite his initial reluctance? How hard should Angela and David press to convince him to help them? Familial fractures along risk status lines are already observed in other disease contexts, in which at-risk members become overly zealous in protecting their branches from the disease of the afflicted. In the analogous context of HIV testing, see levine (38). Genetic testing in Blended families will impose the additional tension of tracing health risks along lines already marked by parentage.

Virtual Families

Somewhere between the extended and the blended family are families that retain the identity of a biological extended family but no longer live together. Here is the incarnation of what some call the "abstract" concept of families: "an idea or ideal that refers to a family name or genetic line, the extended family in the largest sense, whose boundaries or members extend over both space and time" (36, p. 47). They are "virtual families" in that except for high holidays, most of their interactions are mediated by communication technology. Like organic families, these families may be motivated to seek testing, since they define themselves in terms of their biological connections. However, both distance and increased control over communication serve to make extended family members more inaccessible and more autonomous. As sociologists point out, for today's virtual families, "contact with relatives outside the nuclear family depends not only on geographical proximity — not to be taken for granted in our mobile society — but also on personal preference. Even relations between parents and children are matters of individual negotiation once children have left home" (39).

Clinicians involved in genetic testing already describe cases that suggest how the familial virtues might be challenged in families that fit this description as well. For example, consider the following case from the Indiana study:

> Paul, a healthy twenty-five year old man, contacts a testing center to request genetic testing for Huntington disease. Paul's father has HD. The father and his family live in another part of the country and are not known to the center. Paul has an identical twin brother, Michael, who does not wish to have the predictive test. Paul has moved away from his family for work reasons, but he maintains regular telephone contact with his parents, twin brother, and older sister. . . . (35, p. 21).

In discussing the management of this case, the authors frame its professional ethical challenge as a conflict between Paul's "right to know" his HD risk status, and Michael's "right not to know" his own identical risk. Interestingly, however, Paul apparently uses no "rights" language at all in framing the issue he faces. The case notes simply report that:

> Paul is planning to get married; he and his partner Linda are determined to find out whether he will develop HD before they start a family. Michael is single, and, according to Paul, does not think he could cope with the knowledge if he knew that he carried the HD mutation. Paul understands that his test result will reveal Michael's gene status. He states that he will not tell Michael or other family members that he is undergoing the predictive test. If he receives a positive result, he says that he will not tell Michael. However, he might tell Michael if he were to receive a negative result.

Against the background of his loyalty to his future family, Paul is wrestling with the tension between his existing obligation to protect his vulnerable brother and the costs to his natal family's intimacy such secrecy will exact. While he thinks he could succeed in keeping the test a secret from his natal family in order to protect Michael from bad news, how could he not share good news with Michael if he got it?

It is interesting that the authors of this case study concur with Paul's reasoning in this case and recommend accepting him for testing, despite their strong defense of obligate carriers' "right not to know" in more organic familial circumstances. As a result of the structural loosening of family ties in virtual families, it is also often easier to keep secrets within these families, increasing local control over both decisions to be tested and the privacy of test results. On the other hand, virtual families also face a structural challenge in preserving their intimacy for the same reasons, and that challenge poses its own risks. Paul and his wife will find themselves having to conceal increasingly more important aspects of their life if his test is negative, or else attempt to reveal their secret to Michael without the benefit of the mental preparation their own pre-test counseling provided them. Moreover efforts to recruit dispersed families for studies of hereditary mental health problems have shown that as familial communication becomes attenuated, it is possible for the dispersed branches to lose familiarity with the health problems to which they are prone, making it potentially more stigmatizing to reveal genetic risks across the family tree and reducing the motivation for distant members to contribute to clinical studies of the extended family (40).

CONCLUSION

From the current discussions of professional ethical problems in clinical genetics, it appears that at least three traditional virtues of family life are being severely tested by the increasing availability of predictive genetic risks assessment tools: loyalty, intimacy, and security. For blended and virtual families, the prospect of genetic testing can threaten our commitment to the familial virtues by focusing us on what separates us from each other. Within organic families, however, genetic testing can serve to reinforce our sense of solidarity. In that context the risk is not that the traditional virtues will be undermined but that we will be encouraged to go too far in pursuit of what it means to be a good family and our black sheep will suffer for it. To that extent, ironically, it may be that it is the most traditional form of the family, the sessile organic family, in which the challenge to our traditional familial virtues will be most difficult to anticipate and address.

Clearly, these claims can only be suggestions for further empirical research at this point: They assume that the categories of families employed here are useful within the context of genetic testing and that the associated familial virtues are the values that are at stake. To the extent that these ideas are worth exploring further, however, they do serve to suggest one important conclusion: as important as the professional ethical and public policy challenges of human genome research are, they may ultimately be eclipsed by the impact on how we, the public, understand our relations with our closest kin and the obligations those relations entail.

BIBLIOGRAPHY

1. President's Commission for the Study of the Ethical Problems in Medicine and Biomedical and Behavioral Research, *Screening and Counseling for Genetic Conditions*, U.S. Government Printing Office, Washington, DC, 1983, p. 55.
2. E.T. Juengst, *Am. J. Hum. Genet.* **54**, 121–128 (1994).
3. E. Thomson, *Ethical, Legal and Social Implications Program: Grants, Contracts, and Related Activities*, National Human Genome Research Institute, Bethesda, MD, 1996.
4. S. Post, *Spheres of Love: Toward a New Ethics of the Family*, Southern Methodist University Press, Dallas, TX, 1994.
5. H. Lindemann Nelson and J. Lindeman Nelson, *The Patient in the Family: An Ethics of Medicine and Families*, Routledge, New York, 1995.
6. S. Mintz and S. Kellogg, *Domestic Revolutions: A Social History of American Family Life*, Free Press, New York, 1988.
7. R. Shinn, in M. Frankel and A. Teich, eds., *The Genetic Frontier: Ethics, Law and Policy*, AAAS, Washington, DC, 1993, pp. 9–25.
8. R.M. Green and A.M. Thomas, *J. Genet. Couns.* **6**, 245–254 (1997).
9. B. Biesecker and L. Brody, *J. Am. Med. Women's Assoc.* **52**, 22–27 (1997).
10. B. Biesecker et al., *J. Am. Med. Assoc.* **269**, 1970–1974 (1993).
11. M. Yarborough, J. Scott, and L. Dixon, *Theor. Med.* **10**, 139–149 (1989).
12. L. Parker and C. Lidz, *IRB: Rev. Hum. Subj. Res.* **16**, 6–12 (1994).
13. T.M. Powell, D. Cox, S. Jeffery, and D. Cua, *Am. J. Hum. Genet.* **63**(4), A204 (1998).
14. R.J. Pokorski, *Am. J. Hum. Genet.* **60**, 205–216 (1997).
15. M. Huggins et al., *Am. J. Hum. Genet.* **47**, 4–12 (1990).
16. A. Wexler, *Mapping Fate: A Memoir of Family, Risk and Genetic Research*, Random House, New York, 1995.
17. J.K.M. Gevers, *Med. Law* **7**, 161–165 (1988).
18. P.R. Winter et al., *Am. J. Med. Genet.* **66**, 1–6 (1996).
19. L. Andrews, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, pp. 255–281.
20. E.W. Clayton, *Am. J. Med. Genet.* **57**, 630 (1995).
21. M. Shaw, *Am. J. Hum. Genet.* **26**, 243–246 (1987).
22. R. Wachbroit, *Suffolk Univ. Law Rev.* **27**, 1391–1410 (1993).
23. T.H. Murray, *The Worth of a Child*, University of California Press, Berkeley, 1996.
24. D. Callahan, *Hastings Cent. Rep.* **15**, 32–37 (1985).
25. A. Buchanan and D. Brock, *Deciding for Others: The Ethics of Surrogate Decision-making*, Cambridge University Press, Cambridge, UK, 1989, p. 136.
26. A. Buchanan, *Soc. Philo. Policy* **13**, 18–47 (1996).
27. T. Marteau and H. Drake, *Soc. Sci. Med.* **40**, 1127–1132 (1995).
28. A. Charo and K. Rothenberg, in E. Thomson and K. Rothenberg, eds., *Women and Prenatal Testing: Facing the Challenges of Genetic Technology*, Ohio State University Press, Columbus, 1994, pp. 105–131.
29. R. Faden, in E. Thomson and K. Rothenberg, eds., *Women and Prenatal Testing: Facing the Challenges of Genetic Technology*, Ohio State University Press, Columbus, 1994, pp. 88–98.
30. E. Clayton, *J. Med. Philos.* **22**, 233–251 (1997).
    A. Fryer, *Arch. Dis. Child.* **73**, 97–99 (1995).
    ASHG/ACMG, *Am. J. Hum. Genet.* **57**, 1233–1241 (1995).
31. C. Cohen, *Kennedy Inst. Ethics J.* **8**, 111–130 (1998).
32. B. Gottleib, *The Family in the Western World*, Oxford University Press, Oxford, UK, 1993.
33. R. Macklin, *Hastings Cent. Rep.* **21**, 5–12 (1991).
34. G.M. Petersen and P.A. Boyd, *J. Natl. Cancer Inst.* **17**, 67–71 (1995).
35. D. Smith et al., *Early Warning: Cases and Ethical Guidance for Presymptomatic Testing in Genetic Diseases*, Indiana University Press, Bloomington, 1998, p. 72.
36. P. Smith, in D. Meyers, K. Kipnis, and C. Murphy, eds., *Kindred Matters: Rethinking the Philosophy of the Family*, Cornell University Press, Ithaca, NY, 1993, p. 47.
37. J. Maley, *An Ethics Casebook for Genetic Counselors*, Center for Biomedical Ethics and Division of Medical Genetics, University of Virginia, Charlottesville, 1994, pp. 36–41.
38. C. Levine, *Milbank Q.* **68**(Suppl. 1), 37–50 (1990).
39. R. Bellah et al., *Habits of the Heart: Individualism and Commitment in American Life*, University of California Press, Berkeley, 1985, p. 89.
40. D. Shore et al., *Am. J. Med. Genet.* **48**, 17–22 (1993).

See other entries GENETIC COUNSELING; GENETIC INFORMATION, ETHICS, AND INFORMATION RELATING TO BIOLOGICAL PARENTHOOD; GENETIC INFORMATION, ETHICS, PRIVACY AND CONFIDENTIALITY: OVERVIEW; GENETIC INFORMATION, LEGAL, FDA REGULATION OF GENETIC TESTING; GENETIC INFORMATION, LEGAL, GENETIC PRIVACY LAWS.

GENETIC INFORMATION, ETHICS, INFORMED CONSENT TO TESTING AND SCREENING

KIMBERLY A. QUAID
Indiana University School of Medicine
Indianapolis, Indiana

OUTLINE

INTRODUCTION

The rapid advance of genetic knowledge in the past decade has led to dramatic increases in our ability both to diagnose individuals with a variety of genetic disorders as well as to identify individuals at increased risk for developing genetic disorders at some time in the future. The relative newness of this technology, as well as its rapid dissemination into general medical practice, has raised numerous issues regarding the role of genetic testing in public health, in medical care, and in the lives of individuals at risk for or affected by diseases of genetic origin. Debates on these issues are clouded by the fact that our ability to treat many of these disorders lags far behind our ability to identify those at high risk, the treatments that are available are often expensive and less than entirely satisfactory, and the cost of testing remains fairly high. In the absence of general societal consensus on the larger issues surrounding genetic testing such as who should be tested, for what indications, what type of treatment and counseling should be available prior to testing, who pays and who should have access the information, the issue of informed consent has gradually emerged as a central focus in the testing debate. This article will discuss the general issue of informed consent, delineate the differences between genetic screening and genetic testing, and present the salient informed consent issues to be considered.

DEFINITION OF INFORMED CONSENT

An in-depth analysis of the issue of informed consent is beyond the scope of this article, and the reader is referred to other articles that address this topic at length. For our purposes it is important simply to understand the basic concept of informed consent. Informed consent has been defined as an *autonomous authorization* by a subject or patient that allows a professional either to involve the subject in research or to initiate a medical plan for the patient, or both (1). Informed consent occurs if and only if a patient or subject with a substantial understanding and in a substantial absence of control by others intentionally authorizes a professional to do something (2).

Traditionally medical ethics has been dominated by a commitment to the principle of medical beneficence — the principle that obligates physicians to further the medical best interests of their patients (1). Within this tradition and over time, the general consensus has emerged that "informed consents should be obtained for research or experimental interventions (therapeutic as well as non-therapeutic), for innovative interventions, for interventions that carry with them substantial or unknown risks and in situations of choice between substantially different medical plans. Also, all other things being equal, informed consents need not be obtained for routine interventions that pose little or no risk to the patient and that are not part of a research protocol" (1, p. 39). This level of general agreement as to when informed consent is required provides little guidance for specific medical treatments, interventions, or innovations such as genetic testing, although one author has advanced the idea that most applied genetic technology — screening, diagnosis, counseling, or treatment — should be characterized as therapeutic clinical research (3) requiring informed consent. The history and development of the concept of informed consent for genetic testing and screening is relatively recent, and it has been affected by myriad historical events, public policy and economic shifts, and technological innovations. These variables have resulted in two divergent approaches to informed consent such that consent is usually required for genetic testing but not always for genetic screening. The development of these two approaches to consent will be discussed below.

GENETIC SCREENING AND THE PUBLIC HEALTH MODEL

Definition of Genetic Screening

The term "genetic screening" is generally reserved for the public health context in which interventions that can detect disease or risk of disease are employed on a populationwide basis without regard to family history. A committee of the National Academy of Sciences (NAS) has defined genetic screening as "a search in a population of apparently healthy individuals for those genotypes which place them or their off spring at high risk for disease" (4). With genetic screening, carriers of deleterious genes who have never had an affected child can be identified as can fetuses, newborns, or adults who do not have an affected relative (5). For this reason, in most of the populations targeted for genetic screening — African-Americans for sickle cell, Jews of Ashkenazi descent for Tay-Sachs, virtually all newborns for PKU — parents or individuals will have no first-hand knowledge of the disorder, its

symptoms, severity, course, or risk for recurrence due to a lack of family history.

Public health is concerned with the prevention and reduction of morbidity and mortality in a given community. Within the public health framework, the principle of beneficence, which focuses on considerations of human welfare or well-being, is the guiding ethical principle. The principle of beneficence asserts a duty to confer benefits and to work actively to prevent and remove harms (2). Equally important is the duty to balance possible benefits that might accrue as a result of screening against possible harms. In the public health context, the benefit to be sought is the health or welfare of the community as a whole rather than that of the individual. The goal of genetic screening therefore is to reduce the incidence of morbidity and mortality related to genetic diseases in the community (6). If the benefit to be sought is the welfare of the community, what might some of the harms be? In general, harms that may occur during screening programs will occur to individuals and include labeling (7), stigmatization (8), misunderstanding (9), and false reassurance (10).

The degree to which a screening program, once implemented, is successful in reducing morbidity and mortality depends on several variables. These include the prevalence of the condition in the population to be screened, the sensitivity and specificity of the screening tool, the availability of a treatment or intervention for the condition, and the follow-up plans for those identified as positive.

Criteria for Screening

Principles to be followed with regard to genetic screening have been suggested by several authors (3,11,12) These principles include the following:

1. The disorder should be of high burden to the affected individual.
2. The inheritance and pathogenesis of the disorder should be understood.
3. The disorder should be preventable and practical therapy available, including genetic counseling and reproductive alternatives.
4. Patient's right to informed consent, voluntary participation, and confidentiality should be protected.
5. The benefit-to-cost ratio to the patient (public) should be greater than one.
6. The laboratory screening method should minimize false positive and exclude all false negative results.
7. A diagnostic test should be available.
8. Both screening and diagnostic tests should be available to all who require it.

In practice, however, these recommendations are rarely followed. A short history of our experience with newborn, prenatal, and carrier screening may prove instructive.

Experience with Screening

Newborn Screening for PKU. The first widespread experience with newborn screening emerged in the early 1960s with the invention by Robert Guthrie of the Guthrie assay to detect phenylketonuria (PKU). PKU is a hereditary metabolic disorder in which a deficiency of an intracellular enzyme results in the accumulation of the amino acid phenylalanine. PKU has an incidence in the United States and Europe of approximately 1 in 10,000 live births with the primary manifestation of the disorder being severe mental retardation (13). The Guthrie assay could be performed on apparently healthy newborns using a blood sample obtained by heel prick and was simple, inexpensive, and relatively painless. These characteristics were seen to weigh in favor of testing as many individuals as possible relative to the cost of not screening. At the time, however, the ability of a low-phenylalanine diet to prevent retardation had not been proved. Despite this lack of a clear medical benefit to screening, lobbying by Guthrie and the National Association for Retarded Children soon resulted in laws in most states that mandated the screening of newborns for increased concentrations of phenylalanine in the blood (14) even in the face of opposition to mandatory screening on the part of prominent groups including the American Academy of Pediatrics (15). In the rush to initiate widespread screening, some normal children with transient hyperphynylalaninemia were incorrectly labeled as having PKU and prescribed diets deficient in phenylalanine with deleterious outcomes including death (4,16).

Today, in most states, parental consent for PKU screening is not sought (17). In fact genetic screening, especially for newborns, is often carried out without informed consent or parental permission (18). Current common newborn screening tests performed without express informed consent include tests for hypothyroidism, PKU, sickle cell, and/or other hemoglobinopathies, although parents may object to screening on religious grounds (17). This exception to mandatory screening, however, may be seen as meaningless because parents seldom learn about the test until after it has been performed (19).

Some have concluded that informed consent ought not to be necessary for a procedure that offers great benefit and little risk (20). According to this argument, parental autonomy in decision making about newborn screening is grounded in the principle of beneficence, which holds that parents are the people best suited to act in the best interests of the infant. If it is generally agreed that it is in the best interest of the child to be screened, parental consent is superfluous and need not be sought. While this argument has clear appeal in the case where treatment is available (i.e., PKU), it is less compelling in situations in which the benefits of screening are less clear-cut. A second justification for not seeking informed consent in the context of newborn screening has been lodged in the notion of *therapeutic privilege* according to which a physician may legitimately withhold information based on sound medical judgment that to divulge the information would be potentially harmful. Harmful outcomes that would qualify for consideration of therapeutic privilege include endangering life, causing irrational decisions, and producing anxiety or stress (2). Physicians have justified the lack of informed consent by "maintaining that they do not wish to alarm women unnecessarily, since the tests will be normal in the vast majority of cases" (5, p. 184).

Finally, a third reason underlying the sentiment favoring lack of informed consent is that women might actually refuse to have the test done (5). Empirical investigation of this question for PKU, however, has shown that parents can be educated about this test at little cost (21) and that, if asked, less than 0.05 percent of parents would refuse testing (19). Some commentators feel that the scope of therapeutic privilege should be severely circumscribed and that at the very least, the privilege should not apply in situations when the potential harm to the patient from full disclosure would result not from the disclosure itself but from the decision that the practitioner fears that the patient might make as a result of the information disclosed (22).

Prenatal Screening. In addition to screening tests performed on newborns, screening tests are also often performed on pregnant women without providing any explanation or obtaining informed consent (23). This screening includes conditions that are primarily of interest because of the risk they pose to the developing fetus, such as rubella and Rh factor, as well as diseases with medical implications for both mother and fetus, such as tuberculosis, gonorrhea, gestational diabetes, and herpes simplex. Justification for this practice lies in the fact that for each of these disorders, effective interventions are available to prevent or markedly reduce harmful consequences (20).

Carrier Screening. The history of carrier screening in the United States might be characterized as one of mixed success. One notable failure is the experience with screening for sickle cell disease in the 1970s (23). Sickle cell disease is an autosomal recessive hemolytic anemia occurring most frequently in blacks but also in persons of Mediterranean, Asian, Caribbean, Middle Eastern, and South and Central American origin. Sickle cell disease is estimated to occur in as many as 1 in 400 African-American newborns (24). The public health implications of sickle cell disease, especially for African-Americans were brought to national attention in the early 1970s. By April 1973, 10 states had laws requiring mandatory screening for sickle cell disease despite the fact that no effective treatment was available. This rush to mandatory screening was not accompanied by provisions or funding for either genetic counseling or education of those screened (14). As a result there was great confusion between sickle cell trait or being a carrier for sickle cell and actually having the disease, on both the part of the general public and those who were screened. Identification of having sickle cell trait or sickle cell disease resulted in documented cases of job discrimination most notably in the military where for years the U.S. Air Force did not train black recruits with sickle cell trait to become pilots (25). Life insurance companies charged higher premiums for individuals with sickle cell trait or refused to insure them at all (26). Other reported hazards of screening for sickle cell included inappropriate medication and treatment of individuals whose symptoms were falsely attributed to sickle cell disease, delays in the adoption of children suspected to have the disease or trait, and the exposure of nonpaternity (27).

This experience made clear to what extent hastily planned, poorly executed, and underfunded screening programs mounted in the name of public health had the potential to cause great harm to individuals. When coupled with the growing importance of the principle of autonomy which advanced the right of those approached for testing to be fully informed as to the profound effects that testing may have on their lives as well as the lives of other family members, this experience led some to question the wisdom of the old approach. The genetics community, if not the wider medical community, also became sensitized to the concept that psychosocial risks, in contrast to medical/physical risks, were most likely to occur as a consequence of these programs and others like them. More recently our experience with screening for the evidence of human immunodeficiency virus (HIV) infection has further dramatized the psychosocial risks of diagnostic or predictive testing (coerced diagnosis, anxiety, loss of privacy, stigmatization, and discrimination). For the individuals most intimately involved, these risks came to be seen as important as physical risks and highlighted the potential downside of being identified as having or being at risk for developing a disorder that was poorly understood or considered highly undesirable by the society at large (14). These lessons have become apparent in the relatively recent handling of genetic screening for another autosomal recessive disorder, cystic fibrosis.

As early as 1983 the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research saw cystic fibrosis (CF) as a likely candidate for widespread genetic screening (12). CF is the most common autosomal recessive disorder among northern Europeans, with 1 in 25 being known carriers (10). The President's Commission came out in favor of explicit informed consent for CF screening and emphasized the fact that in their view, the "fundamental value of genetic screening and counseling lies in its potential for providing individuals with information they consider beneficial for autonomous decision-making" (12, p. 97).

In 1985 identification of DNA probes linked to the CF gene allowed prenatal diagnosis for families who had a previous child with CF and for whom specimens for the affected child and both parents could be analyzed. In 1989 the DNA sequence of the gene associated with CF was published (28) making widespread screening for carrier status possible. Mass screening was widely advocated for several reasons, with the most compelling being the ability to counsel heterozygote carriers of reproductive age. Arguments against mass screening were equally compelling. The initial test was able to detect only 75 percent of carriers (the frequency of the first detectable mutation ΔF_{508}). In addition the cost of screening was estimated to be as high as 2.2 million for each cystic fibrosis birth avoided (29). Shortly after the announcement of the CF gene, the board of directors of the American Society of Human Genetics (ASHG) issued a statement that "Routine CF carrier testing of pregnant women and other individuals is not yet the standard of care in medical practice" (30).

Continuing research showed that CF was genetically heterogeneous with over 100 different mutations giving

rise to the same symptoms. As more mutations were found and the ability of the test to detect carriers improved to over 90 percent, CF came to be seen as a test case for the application of future genetic technologies. Funding for pilot projects was made available by the National Center for Human Genome Research with the express research goals of gathering information to identify clinical practices that best increase patient understanding of disease–gene carrier testing and test results, and best protect individuals and families from test-related psychological harm, stigmatization, and discrimination (31).

In contrast to the previous experience with sickle cell screening, two important developments had taken place. The first was that changes in technology had spawned the development of numerous biotechnology companies eager to capitalize on the potential market for genetic tests and poised to market tests directly to physicians and consumers. The prospect of a test that could potentially be marketed to every Caucasian of child-bearing age was an enormous incentive to tout the benefits of testing. The second development was a major shift in perception of what the proper goals of routine carrier screening ought to be. Rather than promoting the public health model, wherein the goal of a carrier screening program would be a reduction in the incidence of genetic disease, there now appeared to be some level of agreement that the appropriate goal of carrier screening is to help clients make more informed decisions (32). With this change in emphasis came a concomitant change in the definition of what would count as the mark of a successful program. If the public health context dictated that a successful program should result in the reduction of morbidity and mortality in a population, the new emphasis on informed consent suggested that the effectiveness of screening programs should be assessed by whether the participants in the screening program have become informed (29) regardless of the ultimate decision made and regardless of that decision's ultimate effect on morbidity and mortality.

The CF pilot studies established an important precedent by incorporating the assessment of psychosocial impact into clinical studies usually dominated by concerns of medical safety, reliability and efficacy (31). This change reflected the now-common assumption that genetic "screening is a medical intervention with serious psychosocial risk" (29). The feeling was now that mass genetic screening programs should be considered experimental public health programs, implemented only after a favorable assessment of a program design that effectively achieves its goal while minimizing the potential medical, ethical, legal, and social problems (29). For that reason informed consent should be obtained to ensure that the client understands the risks of taking the test and is willing to accept them (33). Pre-test counseling or education has become the most accepted means to promote patient autonomy by providing accurate and understandable information designed to allow patients to decide for themselves whether the benefits outweigh the risks. In the context of screening for CF, the information to be presented should include information about CF, including its prognosis, treatment, and costs. The fact of screening raises the possibility that both members of a couple may be identified as carriers of the CF gene. In that case such a couple would have a one in four chance of having a child affected with CF. This possibility raises the issue of alternative reproductive options. Because one or more of these options, such as abortion, may not be acceptable to some families, these options need to be discussed in detail. Finally, the risks of testing, including testing errors and the possibility of stigmatization or discrimination, should be disclosed (29). This advancement of patient autonomy as the proper goal of screening programs may serve as a potent counterweight to commercial and other pressures in favor of mass screening.

Future of Informed Consent for Genetic Screening

There is a clear precedent for performing prenatal, newborn, and carrier screening for a variety of disorders without obtaining informed consent, although recent experience with screening for cystic fibrosis has placed informed consent clearly on the table as desirable if not required. As the Human Genome Project proceeds apace, the number of screening tests that can be performed on a single blood sample will increase rapidly. The proliferation of genetic screening tests coupled with a perceived shortage of personnel trained in genetics (29) has raised the question of whether obtaining informed consent, or even attempting educational efforts about the tests and their potential findings, is practical (34). George Annas and Sherman Elias have suggested that "generic consent" may be a more useful approach because it may not be feasible to explain the risks and benefits of each test. This approach is based on the presumption that "it will soon be impossible to do meaningful prescreening counseling about all available carrier tests" (34, p. 1611). The authors warn against the dangers of "information overload" and envision a situation in which patients would be informed of broad concepts and common denominator issues in genetic screening rather than specific information about each test and each condition that may be detected. The authors stress that the concept of generic consent is not a waiver of the individual patient's right to information. Rather, "it would reflect a decision by the genetics community that the most reasonable way to conduct a panel of screening tests to identify carriers of serious conditions is to provide basic, general information to obtain consent for the screening and much more detailed information on specific conditions only after they have been detected" (34, p. 1612). This view raises the issue of who will decide for what disease, symptoms, or conditions the screening will be done. Ultimately a single blood test may determine not only the risk to infants for Down syndrome, trisomy 18 and 13, but also sex chromosome abnormalities such as Turner's syndrome and Klinefelter's syndrome and conditions not present at birth but manifest in middle age including breast cancer, Huntington's disease, and Alzheimer's disease. This approach needs to be evaluated in view of the fact that now, and in the foreseeable future, we have few cures or treatments to offer for these genetic disorders. Individuals identified as affected, at risk, or carriers face difficult choices involving mating, reproduction, and abortion. The continued subtle but

steady erosion of a woman's right to abortion should be factored into the prospect of potentially identifying hundreds of thousands of individuals as affected, or at risk for, serious untreatable disorders.

This model has been criticized as driven primarily by technological capabilities rather than one based on established ethical principles (35). However, this model is perfectly in keeping with the historical framing of genetic screening as a public health issue whereby a group of influential experts decides what is in the best interest of the health of the population. A more sober note regarding this practice, and one that we might do well to ponder, is the observation that public health, rather than individual therapy, was the driving force behind the Nazi medicalization of eugenics that brought about the horrors of the holocaust (36).

GENETIC TESTING AND THE PSYCHOSOCIAL MODEL

Definition of Genetic Testing

The term genetic testing encompasses a number of technologies used to detect genetic traits, changes in chromosomes, or changes in DNA. Genetic testing is normally performed for two purposes; (*1*) to confirm a diagnosis in a person with symptoms of a particular disease, or (*2*) to clarify risk in an asymptomatic person who is known to be at risk for a specific disorder. This risk clarification in asymptomatic individuals can be further divided into presymptomatic or predictive testing in those individuals where disease genes have been identified that have high sensitivity and specificity and penetrance, and so-called susceptibility testing where the presence of a particular disease gene may be a necessary but not sufficient precondition for the development of the disorder. Examples of diseases for which predictive genetic testing are available would include Huntington disease, caused by an expanded trinucleotide repeat on chromosome 4 (37), and early-onset Alzheimer's disease associated with mutations on chromosomes 1 (38), 14 (39), and 21 (40). Examples of diseases for which susceptibility testing is available (although not necessarily recommended) include colon cancer (41) and breast cancer associated with mutations in BRCA1 (42) or BRCA2 (43).

In contrast to genetic screening, individuals, whether affected or at risk, for whom genetic testing might be appropriate usually have first-hand knowledge of the disease in question by virtue of having an affected relative. One advantage with regard to informed consent is that these individuals will often have a deep and personal familiarity with the disease, its symptoms and its course. They may not, however, have a clear understanding of the risk to their own offspring or of their options for avoiding the conception or birth of an affected child (5).

Genetic testing is typically initiated by a patient and/or his or her family and is performed in consultation with his or her medical care provider. Until very recently almost all genetic testing was performed in specialized medical centers primarily by professionals trained in the various subdisciplines of genetics including clinical genetics, genetic counseling, molecular genetics, and cytogenetics. The issue of context is important in this discussion. Although medical genetics may be viewed as a subspecialty of medicine as evidenced by the recent formation of the American College of Medical Genetics and the recognition of medical genetics as a primary subspecialty by the American Board of Medical Subspecialties in 1991, medical genetics has evolved with a very different moral tradition. Because of a previous history of eugenics, and for many years its almost exclusive involvement in highly personal and difficult issues of human reproduction, medical genetics almost from its inception has adopted as its primary moral principle respect for autonomy of the client or patient (44). "Respect for autonomy obligates professionals to disclose information, to probe for and to insure understanding and voluntariness, and to foster adequate decision-making" (2, p. 127).

In part, this commitment to respect for autonomy in genetics can be explained by the fact that most of what medical genetics has had to offer is information regarding recurrence risks for specific disorders rather than the more traditional medical benefits such as treatment or cure. What has evolved in genetics is a commitment to the idea of nondirective counseling in the provision of genetic information (45) and in the convention of obtaining informed consent prior to all diagnostic genetic testing (46). The goal of genetic testing is seen as to improve the ability of people affected or at risk to make informed personal and reproductive decisions in light of their genetic status (47). It has been argued therefore that before an individual agrees to be tested for a genetic condition, pre-test education and informed consent are necessary, and post-test counseling must be provided (29).

Genetic Testing for Huntington's Disease

This approach to genetic testing is most clearly seen in the experience with predictive testing for Hunting's disease (HD). In 1983 a linked marker to the HD gene was found (48). Ensuing discussions among genetic researchers, clinicians working with HD patients and their families as well as interested laypersons resulted in preliminary guidelines for predictive testing including eligibility criteria and testing protocols. These protocols included neurological examination, psychiatric and psychological screening, pretest counseling and follow-up (49–51). National and international guidelines for testing were soon issued by the Huntington's Disease Society of America (HDSA) (52–53) and the World Federation of Neurology Research Group on Huntington's chorea (54–55).

These protocols were guided by two main ethical principles, nonmaleficence and autonomy. Nonmaleficence is termed a negative duty to avoid doing harm. Many argue that predictive genetic testing is a complex endeavor and despite ten years of experience the jury is still out on the overall benefits of testing (56). Many of those most intimately involved in the testing process, either as persons at risk or experienced clinicians continue to urge caution (57–59). Mandatory pre-test counseling was, and remains to this day, a major element of testing for HD. The purpose is to ensure, to the extent possible, that individuals considering tested have clearly understood the

risks and benefits of testing to them, and have made a decision regarding testing that is consistent with their own personal goals and values. In other words, mandatory pretest counseling was seen as a means by which to promote autonomous decision-making regarding this predictive test. While this approach has been criticized as being overly paternalistic (60), there is some evidence that this approach is useful. A recent survey of 12 testing centers following recommended testing procedures in Canada has indicated a small number of catastrophic events, defined as a completed suicide, suicide attempt, or psychiatric hospitalization, have occurred as a result of testing (61). A recent survey of all testing centers in the United States has indicated that one advantage to the recommended approach to testing is the ability to identify the estimated 3 percent of individuals for whom testing is inappropriate or who would benefit from further counseling prior to testing (62). Debate continues regarding how much a person needs to know prior to testing and whether the standard approach to informed consent (33) can be maintained as the number and complexity of potential genetic tests increase.

The experience with HD and other disorders, as well as thoughtful commentary by a number of people, have alerted us to a number of other ethical issues that need to be considered in the context of informed consent for genetic testing. One of these is the issue of confidentiality with regard to genetic information. In truth, the results of a genetic test often have important ramifications beyond the individual patient including children or other blood relatives. For example, if an individual at risk for HD by virtue of having an affected parent tests positive for the mutation that causes HD, each of his or her children now has a 50 percent of having inherited the HD mutation from his or her parent. While confidentiality is basic to all health care relationships, limits to confidentiality have been accepted in the case of communicable diseases or intent to harm another person. In the HD testing, a common example is the situation in which an individual who has tested positive refuses to inform his or her children about their risk for HD. Situations such as these have led to discussions that at least consider the justification of breaching confidentiality, although the issue remains a thorny one (63).

A different, but related issue is that of privacy of genetic information. Should results of genetic tests be placed in medical records whereby they might be easily accessed by employers or insurance companies. Concern for confidentiality and privacy of genetic information is highest in those situations were the results of a genetic test may affect an individuals ability to obtain or to keep health insurance. This concern has been heightened by reports of individuals denied insurance, especially those stories based on misunderstanding of the genetic condition, confusions between being affected with a particular disease and being a carrier, or between having a positive genetic test and actually being affected with a particular condition (64). These concerns also extend into the arena of employment where individuals fear being asked to undergo genetic testing as a prerequisite for employment.

Genetic Testing for Cancer

As discussed above, genetic testing has traditionally been performed for well-defined, monogenetic syndromes that are inherited in a strict Mendelian fashion such as cystic fibrosis and HD. For these diseases, inheritance of the disease gene leads to the expression of the disease phenotype. While raising interesting issues in their own right, these diseases are not the best models for the future of genetic testing and the challenges that are likely to arise. For a better view of the future of genetic testing, we must look to the inherited cancers to understand the full complexity of genetic testing and counseling.

Although chromosomal locations for several cancer-predisposing genes, including those for hereditary retinoblastoma (Rb) gene, the WT1 gene for Wilms's tumor, neurofibromatosis type 1 gene, the APC gene for familial polyposis coli, and the p53 gene for Li-Faumenia syndrome, have been mapped, cancers due to these mutations are rare or are usually preceded by distinctive clinical manifestations. For these reasons, discussions concerning genetic testing for these disorders are rare (65). With the discovery of a major breast–ovarian cancer susceptibility gene on chromosome 17, the situation has changed dramatically.

In 1997 alone, approximately 180,000 American women developed breast cancer and 44,000 of these women died of this disease (66). The first breast cancer predisposition gene (BRCA1) was located to chromosome 17q in 1990 (67) and cloned in 1994 (42). BRCA1 is a large tumor suppressor gene with 22 exons, and more than 100 mutations have been identified. A second breast cancer predisposition gene on chromosome 13 was identified in 1995 (43). While mutation analysis is possible in high risk families, it is not yet feasible in the general population (68). It is estimated that as many as 1 in 300 women may carry germ-line mutations in one or more breast cancer susceptibility genes (69). In contrast to the Mendelian disorders, studies of inherited breast cancer susceptibility provide evidence of incomplete penetrance suggesting that the inheritance of an altered breast cancer gene is not sufficient to produce the disease. Other factors, including additional genetic and/or environmental factors may be necessary. The fact of incomplete penetrance presents the potential for prevention, a factor that may weigh heavily in favor of testing unaffected individuals from families with known hereditary breast cancer (70).

The prospect of widespread screening for cancer susceptibility provides the opportunity to take a close look at the process of informed consent, not only for cancer but for any clinical context in which there is inherent uncertainty about the benefits and risks of a specific test (71). From the beginning, those involved in genetic testing for breast cancer took the approach used in HD as their model and structured testing protocols based on the principles of autonomy, beneficence, confidentiality, and equity or justice (71). The research protocols through which testing for breast cancer was first offered included pretest counseling and education and follow-up (73). In virtually all discussions of genetic testing for breast cancer susceptibility, the issue of informed consent is considered of crucial importance. Reasons for this emphasis may be

the influence of the HD model, the fact that most testing involves adults and/or the fact that the benefits of testing are unclear while the risks and limitations of testing are becoming better known.

Three key limitations of genetic testing for breast cancer susceptibility have been identified (74). First, genetic testing for breast cancer susceptibility may not be informative. Second, the genetic information is probablistic in nature indicating an increased or decreased risk but not certainty. A negative genetic test result for BRCA1 leaves an individual with the population lifetime risk of breast cancer due to environmental and other factors and which is estimated at approximately 11 percent (75). Third, there are very real limitations to the current methods for cancer prevention and early detection. It is unclear whether current prevention methods of treatment (surgery, radiation, chemotherapy) or surveillance practices (mammography, clinical breast examination, and self-examination) actually reduce morbidity and mortality or what role genetic testing should play. In addition individuals must consider the risk of genetic discrimination in employment or in obtaining or maintaining insurance (64) and the potential for negative psychological consequences of learning one's genetic status (74,76–77).

The major benefit of genetic testing for susceptibility is the possibility of reassuring information if one is found not to carry a particular genetic mutation associated with increased risk for cancer (78). For those who test negative, one risk may include survivor guilt and shame (77). For those who test positive, testing may prove beneficial if this information facilitates actions to help prevent cancer or to aid in its early detection. This benefit will not be realized, however, if increased worry about cancer interferes with adherence to surveillance measures such as breast self examination, clinical breast examinations, and mammograms (79). Data suggest that a substantial proportion of women identified as gene carriers experience some level of psychological distress including persistent worries, depression, confusion and sleep disturbance (76).

Several studies have been designed specifically to determine what people want to know about susceptibility testing for cancer. In one study the authors concluded that at a minimum, disclosure statements needed to include information on (*1*) risk factors for cancer, (*2*) practical details of testing, (*3*) current limitations of testing, (*4*) available follow-up options, (*5*) known benefits of testing, and (*6*) known risks of testing (80). Other groups have recommended that informed consent be obtained through full disclosure of all the risks and benefits of testing including information about test accuracy, importance of correct cancer diagnosis, laboratory error rate, physical risks of blood drawing, potential problems with insurance or employment, psychological risks, and benefits of health care planning (68). A third study focused on two essential goals of informed consent, assuring that patients have substantial understanding and that their decisions whether to accept or to reject interventions are substantially voluntary. Two specific concerns arose with regard to these two goals, the relationship between patients' backgrounds, beliefs, and their understanding and the role of provider recommendations in voluntary decision making (71). One conclusion reached was that patients often expect and request that their providers play an active role in decision making. This approach is somewhat at odds with the traditional stance of nondirective counseling as held by genetic counselors and is time intensive, presenting a major challenge to our current system of health care delivery.

Genetic Testing for Children and Adolescents

The genetic testing of minors, especially for late-onset disorders, has long been a subject of much controversy due to questions about the ability of minors to give informed consent and because of the sometimes conflicting views of lawyers and ethicists. Anglo-American common law has long granted almost absolute authority to parents (81), and in the context of medical treatment, parents are allowed to do almost anything that is not harmful in and of itself or not clearly against the best interests of the child (82). From the beginning, however, general testing protocols and guidelines for HD prohibited the testing of those under 18 years of age at their parents' request (52–55). This prohibition was based on a desire to preserve the autonomy of the child and the ability to make his or her own decision about testing, and the desire to avoid causing harm. Several professional societies have published statements upholding the presumption against testing asymptomatic children for late-onset disorders for which there is no treatment or cure (83–85). Research has shown, however, that many health professionals would test children at their parents' request including 53 percent who would test for HD (86), and the issue remains controversial (87–88).

Potential harms of testing are seen to include misconceptions about the future, stigmatization, feelings of unworthiness, fear of intimacy or interpersonal relationships, harm to self-concept, guilt, and blame. Depending on the condition, potential benefits may include increased medical surveillance in those at high risk, relief in those found to be at low risk, the opportunity to obtain accurate information from a trained professional rather than reliance on parental knowledge, and allowing more time for adaptation to the possibility of future illness (89).

In general, genetic testing of minors can be seen to fall into four general categories based on utility where (*1*) testing offers immediate medical benefits for the minor, (*2*) there are no medical benefits but testing may be useful in making reproductive decisions, (*3*) the parents or the minor requests testing, and (*4*) testing is done solely for the benefit of another family member (90). The general consensus on the acceptability of testing varies widely among these categories as does the discussion of what might consitute proper counseling, how to assess minors' ability to give informed consent, and what are the actual outcomes of such testing. While the debate goes on, evidence suggests that more and more children are being tested (91).

Future of Informed Consent for Genetic Testing

The traditional role of informed consent for genetic testing is facing many challenges. One major challenge to the role of informed consent is the rapid movement of

genetic testing out of specialized genetics centers and into general medical practice. The increasing use of heavily marketed, commercially available tests by physicians who may not have the necessary training or background in genetics to either inform patients properly about the risks and benefits of testing or to interpret the results of testing accurately may seriously undermine the current practice of obtaining informed consent. One recent study examining genetic testing for the germ-line mutation of the adenomatous polyposis coli (APC) gene that causes colon cancer found that patients who underwent APC testing often received inadequate counseling, did not provide informed consent, and would have been given incorrectly interpreted results (92).

Evidence exists that biotechnology companies are also offering tests directly to consumers raising questions of whether patients and families are adequately informed about the tests, their limitations, and the risks and benefits of choosing to be tested (91). As the number of tests increases, especially for disorders that are relatively common in the population, the pressure on the part of biotechnology companies to market these tests will increase as well as the pressure on health care professionals to make them available. It remains to be seen whether and to what extent the fragile adherence to the notion of informed consent will survive.

BIBLIOGRAPHY

1. R.R. Faden, in J. Rodin and A. Collins, eds., *Women and New Reproductive Technologies: Medical, Psychosocial, Legal and Ethical Dilemmas*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1991, p. 39.
2. T.L. Beauchamp and J.A. Childress, *Principles of Biomedical Ethics*, 4th edn., Oxford University Press, New York, 1994.
3. L.J. Elsas, *Emory Law J.* **39**, 811–853 (1990).
4. Committee on Inborn Errors of Metabolism, *Genetic Screening: Programs, Principles and Research*, National Academy of Sciences, Washington, DC, 1975.
5. N.A. Holtzman, *Proceed with Caution: Predicting Genetic Risks in the Recombinant DNA Era*, Johns Hopkins University Press, Baltimore, MD, 1989.
6. R.R. Faden, N.E. Kass, and M. Powers, in R.R. Faden, G. Geller, and M. Powers, eds., *AIDS, Women and the Next Generation: Towards a Morally Acceptable Public Policy for HIV Testing of Pregnant Women and Newborns*, Oxford University Press, New York, 1991, pp. 3–26.
7. A.B. Bergman and S.J. Stamm, *N. Engl. J. Med.* **176**, 1008–1013 (1967).
8. G. Stamatoyannopoulos, in S.G. Motulsky and W. Lenz, eds., *Birth Defects*, Excepta Medica, Amsterdam, The Netherlands, 1974, pp. 268–276.
9. M.T. Pugilese et al., *Pediatrics* **80**, 175–182 (1987).
10. H. Bekker et al., *J. Med. Genet.* **31**, 364–368 (1994).
11. L.J. Elsas, in A. Rudolph and J. Hoffman, eds., *Pediatrics*, 18th ed., Appleton and Lange, Norwalk, CT, 1987, pp. 222–225.
12. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Genetic Screening and Counseling: The Ethical, Social and Legal Implications of Genetic Screening and Education Programs*, U.S. Government Printing Office, Washington, DC, 1983.
13. F.D. Ledley, A.G. DiLella, and S.L.C. Woo, *Trends Genet.* **I**, 309–313 (1985).
14. B.S. Wilfond and K. Nolan, *J. Am. Med. Assoc.* **270**(24), 2948–2954 (1993).
15. American Academy of Pediatrics, *Pediatrics* **39**, 623–624 (1967).
16. S.P. Bessman and J.P. Swazey, in E. Mendelsohn, J.P. Swazey, and I. Traviss, eds., *Human Aspects of Biomedical Innovation*, Harvard University Press, Cambridge, MA, 1971, pp. 49–76.
17. K.L. Acuff, in R.R. Faden, G. Geller, and M. Powers, eds., *AIDS, Women and the Next Generation: Towards a Morally Acceptable Public Policy for HIV Testing of Pregnant Women and Newborns*, Oxford University Press, New York, 1991, pp. 121–165.
18. L.B. Andrews, *Medical Genetics: A Legal Frontier*, American Bar Foundation, Chicago, IL, 1987.
19. R.R. Faden et al., *Am. J. Pub. Health* **72**, 1347–1352 (1982).
20. R.R. Faden, N.A. Holtzman, and A.J. Chwalow, *Am. J. Pub. Health* **72**, 1396–1400 (1982).
21. N.A. Holtzman et al., *Pediatrics* **72**, 807–812 (1983).
22. A.M. Capron, *Univ. of Pennsylvania Law Rev.* **123**, 364–376 (1974).
23. K.L. Acuff and R.R. Faden, in R.R. Faden, G. Geller, and M. Powers, eds., *AIDS, Women and the Next Generation: Towards a Morally Acceptable Public Policy for HIV Testing of Pregnant Women and Newborns*, Oxford University Press, New York, 1991, pp. 59–93.
24. D.L. Rucknagle, *Arch. Int. Med.* **133**, 595–606 (1974).
25. L. Andrews, *State Laws and Regulations Governing Newborn Screening*, American Bar Foundation, Chicago, IL, 1985.
26. P. Reilly, *Genetics, Law and Social Policy*, Harvard University Press, Cambridge, MA, 1977.
27. E. Beutler et al., *N. Engl. J. Med.* **285**(26), 1486–1487 (1971).
28. J.R. Riordan et al., *Science* **245**, 1066–1073 (1989).
29. B. Wilfond and N. Fost, *J. Am. Med. Assoc.* **263**, 2777–2783 (1990).
30. C.T. Caskey, M.M. Kaback, and A.L. Beaudet, *Am. J. Hum. Genet.* **46**, 393 (1990).
31. K.E. Hanna, in *Society's Choice: Social and Ethical Decision-Making in Biomedicine*, Institute of Medicine, National Academy Press, Washington, DC, 1995, pp. 432–457.
32. M. Lappe, J. Gustafson, and R. Roblin, *N. Engl. J. Med.* **286**, 1129–1132 (1972).
33. R.R. Faden and T.L. Beauchamp, *A History and Theory of Informed Consent*, Oxford University Press, New York, 1986.
34. S. Elias and G.J. Annas, *N. Engl. J. Med.* **330**, 1611–1613 (1994).
35. L.G. Biesecker and B.S. Wilfond, *N. Engl. J. Med.* **331**, 1024 (1994).
36. A.L. Caplan, in G.J. Annas and S. Elias, eds., *Gene Mapping: Using Law and Ethics as Guides*, Oxford University Press, New York, 1992, pp. 128–141.
37. Huntington's Disease Collaborative Research Group, *Cell* **72**, 971–983 (1993).
38. E. Levy-Lehad et al., *Science* **269**, 970–973 (1995).
39. R. Sherrington et al., *Nature* **375**, 754–760 (1995).
40. A.M. Goate et al., *Nature* **349**, 704–706 (1991).
41. I. Nishisho et al., *Science* **253**, 665–669 (1991).
42. Y. Miki et al., *Science* **266**, 66–71 (1994).
43. R. Wooster et al., *Nature* **378**, 789–792 (1995).

44. K.L. Garver and B. Garver, *Am. J. Med. Genet.* **49**, 1109–1118 (1991).
45. B.A. Fine, in D.N. Bartels, B.S. LeRoy, and A.L. Caplan, eds., *Prescribing Our Future: Ethical Challenges in Genetic Counseling*, Aldine de Gruyter, New York, 1993, pp. 101–117.
46. B.S. LeRoy, in D.N. Bartels, B.S. LeRoy, and A.L. Caplan, eds., *Prescribing Our Future: Ethical Challenges in Genetic Counseling*, Aldine de Gruyter, New York, 1993, pp. 101–117.
47. K. Nolan and S. Swenson, *Hastings Center Rep.* **18**(5), 40–46 (1988).
48. J.F. Gusella et al., *Nature* **306**, 234–238 (1983).
49. G.J. Meissen et al., *N. Engl. J. Med.* **318**, 535–542 (1988).
50. S. Fox et al., *Am. J. Med. Genet.* **32**, 211–216 (1989).
51. J. Brandt et al., *J. Am. Med. Assoc.* **261**, 3108–3114 (1989).
52. Huntington's Disease Society of America, *Guidelines for Predictive Testing for Huntington Disease*, Huntington's Disease Society of America, New York, 1989.
53. Huntington's Disease Society of America, *Guidelines for Geneitc Testing for Huntington's Disease*, Huntington's Disease Society of America, New York, 1994.
54. L. Went, *J. Med. Genet.* **27**, 34–38 (1990).
55. L. Went, *Neurol.* **44**, 1533–1536 (1994).
56. A.M. Codori and J. Brandt, *Am. J. Med. Genet.* **54**, 174–184 (1994).
57. K.A. Quaid, *J. Genet. Counsel.* **1**, 277–302 (1992).
58. C.V. Hays, *N. Engl. J. Med.* **327**, 1449–1451 (1992).
59. M.A. Chapman, *Am. J. Med. Genet.* **42**, 491–498 (1992).
60. D. DeGrazia, *J. Clinic Ethics* **2**, 219–228 (1991).
61. M. Bloch et al., *Am. J. Hum. Genet.* **49**(4), A33, #1939 (1996).
62. M.A. Nance et al., *Am. J. Hum. Genet.* **61**(Supp.), A225 (1997).
63. D. Smith et al., *Early Warning: Cases and Ethical Guidance for Presymptomatic Testing in Genetic Diseases*, Indiana University Press, Bloomington, IN, 1998.
64. P.R. Billings et al., *Am. J. Hum. Genet.* **50**, 476–482 (1992).
65. F.P. Li et al., *J. Nat. Cancer Inst.* **84**, 1156–1160 (1992).
66. American Cancer Society, *Cancer Facts and Figures — 1997*, American Cancer Society, Atlanta, GA, 1997.
67. J.M. Hall et al., *Science* **250**, 1684–1689 (1990).
68. B.J. Baty et al., *J. Genet. Counsel.* **6**, 223–244 (1997).
69. E.B. Claus, N. Risch, and W.D. Thompson, *Am. J. Hum. Genet.* **48**, 232–242 (1991).
70. J.E. Garber and A.F. Patenaude, *Cancer Surveys* **25**, 381–397 (1995).
71. G. Geller et al., *Hastings Center Rep.* (March–April) **27**, 28–33 (1997).
72. K.M. Kash, *Ann. N.Y. Acad. of Sci.* 41–52 (1995).
73. J.R. Boukin et al., *J. Nat. Cancer Inst.* **88**, 872–882 (1996).
74. C. Lerman and R. Croyle, *Arch. Intern. Med.* **154**, 609–616 (1994).
75. K. Offit, *Clinical Cancer Genetics: Risk Counseling and Management*, Wiley-Liss, New York, 1998.
76. H.T. Lynch et al., *Arch. Intern. Med.* **53**, 1979–1987 (1993).
77. B.B. Biesecker et al., *J. Am. Med. Assoc.* **269**, 1970–1974 (1993).
78. K.G. MacDonald et al., *Can. Med. Assoc. J.* **154**, 457–464 (1996).
79. C. Lerman et al., *Health Psychol.* **10**, 259–267 (1991).
80. B.A. Bernhardt et al., *J. Genet. Counsel.* **6**, 207–222 (1997).
81. P. Aries, *Centuries of Childhood: A Social History of Family Life*, Vintage Books, New York, 1965.
82. R.H. Nicholson, *Medical Research with Children: Ethics, Law and Practice*, Oxford University Press, Oxford, UK, 1986.
83. Clinical Genetics Society, *J. Med. Genet.* **31**, 785–797 (1994).
84. American Medical Association, Council on Ethical and Judicial Affairs, *Code of Medical Ethics Reports*, vol. 6, No. 2, Rep. 66. American Medical Association, Chicago, IL, 1995.
85. American Society of Human Genetics Board of Directors and the American College of Medical Genetics Board of Directors, *Am. J. Hum. Genet.* **57**, 1233–1241 (1995).
86. A. Clarke, *The Genetic Testing of Children: Report of a Working Party of the Clinical Genetics Society*, Birmingham Maternity Hospital, Edgbaston and Birmingham, UK, 1993.
87. N.F. Sharpe, *Am. J. Med. Genet.* **46**, 250–253 (1993).
88. K.A. Quaid, *Am. J. Med. Genet.* **49**, 354 (1994).
89. J.H. Fanos, *Am. J. Med. Genet.* **71**, 22–28 (1997).
90. D.C. Wertz, J.H. Fanos, and P.R. Reilly, *J. Am. Med. Assoc.* **272**, 875–881 (1994).
91. D.C. Wertz and P.R. Reilly, *Am. J. Hum. Genet.* **61**, 1163–1168 (1997).
92. F.M. Giardiello et al., *N. Engl. J. Med.* **226**, 823–827 (1997).

See other entries GENETIC INFORMATION, ETHICS, PRIVACY AND CONFIDENTIALITY: OVERVIEW; GENETIC INFORMATION, LEGAL, ERISA PREEMPTION, AND HIPAA PROTECTION; GENETIC INFORMATION, LEGAL, FDA REGULATION OF GENETIC TESTING; GENETIC INFORMATION, LEGAL, GENETIC PRIVACY LAWS; GENETIC INFORMATION, LEGAL, REGULATING GENETIC SERVICES.

GENETIC INFORMATION, ETHICS, PRIVACY AND CONFIDENTIALITY: OVERVIEW

MADISON POWERS
Georgetown University
Washington, District of Columbia

OUTLINE

INTRODUCTION

Research into the genetic contribution to disease holds out the promise of scientific discoveries that will revolutionize the diagnosis, prevention, and treatment of numerous medical conditions, thereby reducing premature mortality and excess morbidity. However, the expected benefits from genetic research and the incorporation of its techniques into clinical practice are accompanied by risks to individual privacy. The importance of protecting individual privacy, especially in circumstances involving highly personal,

often sensitive information such as that revealed by the analysis of an individual's genome, finds justification in a number of closely related moral arguments.

A major defense of stringent privacy protection is the claim that the ability to limit the access others have to personal, highly sensitive information is an essential element in any social policy committed to the preservation of individual autonomy (1). The ability to limit third party access to personal information also is necessary for establishing trust and intimacy within personal and professional relationships, for making reproductive and other medical decisions without the undue influence and interference of others, and for preserving valuable social and economic opportunities for persons whose life prospects may be diminished unfairly by the disclosure of information to persons with competing interests (2,3).

Genetic testing reveals just the sort of information that those concerned with the preservation of individual privacy are most anxious to protect. It may reveal a person's current medical condition, increased susceptibility to particular illnesses, or status as a carrier of genes that affect offspring. Such information can be used in ways highly detrimental to many of an individual's most fundamental opportunities and greatly limit his or her range of available economic and social choices. The potential adverse consequences of unwanted disclosures include limitations on access to health, life and disability insurance, lowering of medical treatment priority within public or private health care programs, loss of employment and educational opportunities, restriction of or added burdens placed upon available reproductive options, and social stigma and discrimination (3).

Although the present or future impact of the proliferation of genetic information regarding individuals remains uncertain and not easily quantifiable, persons having first hand experience with genetic testing within their immediate families view the threats to privacy as substantial, and many within the scientific community are concerned that public policies and clinical practices should be designed to mitigate or prevent the kinds of harms that can result from unwanted disclosure of personal genetic data (4). Indeed, many favoring increased genetic research and the rapid translation of genetic knowledge into clinical and public health benefit also favor enhanced efforts to protect the privacy, confidentiality, and anonymity interests of patients and research subjects.

PRIVACY, CONFIDENTIALITY, AND ANONYMITY

Although the precise contours of privacy and related concepts remain the subject of ongoing debate within both legal and philosophical circles a few basic definitions are needed to better understand the issues associated with the collection and use of genetic information. Privacy is often regarded as a state or condition in which personal information about an individual is inaccessible to others (3,5). A person can therefore be said to enjoy genetic privacy when there are limitations or constraints on the access third parties have to information about a person's individual genetic makeup. The most stringent genetic privacy protection policies, accordingly, are those that allow individuals to decide for themselves what if any individualized genetic information will be collected and analyzed.

Complete privacy, however, is both impossible and often undesirable for a variety of reasons. Genetic privacy, or the inaccessibility of others to genetic or any other type of information about individuals, is often a matter of degree. In some instances, what matters most to individuals is not complete privacy — namely where no one has access to personal information — but control over the access particular individuals or institutions have to information regarding one's genetic makeup. Often the central moral concern focuses on the identities of those who have access to their genetic information and the purposes for which that information will be used. For example, individuals often have prudential medical or reproductive reasons to make information about genetic susceptibility to some disease available to their spouses, children, physicians, or genetic counselors, but they may not want to reveal that information to their employers, insurers, or specific family members. In short, the paramount concern in many circumstances is for confidentiality: persons may not want information disclosed in confidence to health professionals to be shared with third parties for whom subsequent disclosure is not authorized.

Complete genetic privacy is impossible to assure for more basic scientific reasons as well. One reason is that sophisticated techniques for collecting and analyzing DNA make it impossible to guarantee that individual genetic information will not be obtained without his or her consent and subsequently used to his or her detriment. Blood, hair and other bodily materials from which DNA can be extracted are available to many third parties in numerous medical settings, non-medical institutional settings such as prisons and the military, as well as the conduct of criminal investigations (6,7). The opportunities third parties have for gaining unwanted and unconsented access to individual genetic information thus extend well beyond the clinic and research institutions.

Another reason that complete genetic privacy is impossible to assure is that many conclusions about genotype can be inferred from phenotype. For example, medically knowledgeable persons can make reliable inferences about the existence of genetic conditions such as Marfan's syndrome from visual observation of physical attributes. Moreover genetic privacy is more difficult to assure because often it is not information over which any individual always can retain exclusive control. Third-party knowledge of the genetic makeup of some individuals can be obtained from testing conducted on an individual's relatives. Researchers and health care providers will therefore learn about the genetic makeup of persons who are neither research subjects nor patients and will have access to genetic information about those who had no opportunity to refuse disclosure.

In other instances, a central concern is the prevention of unwanted third party access to genetic information through the preservation of anonymity. For example, individuals may be motivated by altruistic reasons to make DNA samples available to researchers studying inheritance of a particular disease within a family

or population geneticists studying genetic variation. Apart from any concern about subsequent disclosures of confidential information there are, in addition, risks that others may inadvertently obtain genetic knowledge about an individual who provided DNA samples. In such cases those persons providing the sample reasonably may insist on complete anonymity, or the assurance that once DNA is made available for research the researchers will not retain any personally identifying links (8,9).

Even the term anonymity can be misleading in what it implies about protection from loss of privacy. Genuinely anonymous collection of genetic information occurs when samples are not linked to specific persons. Studies using genetic information where specific personal identifiers have been collected but later discarded have been "anonymized" and thus direct identification of individuals is impossible (9). In instances in which personal identifiers have been stripped but a researcher retains the possibility of linking individuals to samples through a code persons who contributed the samples are not directly identifiable. The possibility of loss of privacy, however, is not eliminated entirely for anonymized (or even anonymous) samples. Under some circumstances (discussed below) researchers or members of the general public may be able to infer the identity of individual participants in studies, and a prudent privacy protection policy may include additional precautions to prevent publicly available or published information about the population from whom DNA samples were collected to be used to gain access to information about specific persons (10).

Privacy thus appears to be the most central conceptual category of moral analysis. Confidentiality, anonymity, and other policies may be largely strategic in achieving the goal of limiting access others have to personal information.

UNIQUENESS OF GENETIC INFORMATION

Those responsible for public policies, research protocols, or clinical practices designed to ensure the privacy, confidentiality, or anonymity of individuals must address the question of whether genetic information is uniquely sensitive and subject to a greater degree of abuse than other types of information. Such claims are often referred to under the rubric of "genetic exceptionalism," or the view that because the risk of privacy loss and its potential for adverse consequences may be greater for genetic information than for other types of medical information, separate and more stringent genetic privacy standards and policies are sometimes warranted (11).

There are three distinguishable questions within the genetic exceptionalist debate. First, is genetic information sufficiently distinct from other types of medical information? Second, does the intimate nature and sensitivity of such information represent a greater threat to individual privacy interests, thus justifying a heightened degree of protection? Third, are there feasible mechanisms for treating genetic information differently than other medical information?

The issue of distinctiveness is a threshold concern. Unless some workable definition of genetic information can be found, neither the question of its special sensitivity nor the issue of appropriate policies for its protection can be resolved. Genetic information clearly includes knowledge that may be gained from genetic testing of individual DNA samples. However, knowledge of an individual's genome also can be obtained by biochemical tests designed to detect gene products such as the production of enzymes or proteins that result from a particular genetic mutation (11,12). Moreover, inferences about an individual's genetic makeup can be made from family histories. For example, hemophilia A is an X-linked, recessive condition, and a diagnosis of a male child with hemophilia A reveals that it was inherited from the mother, and that the mother's female siblings also have a 50 percent probability of being carriers of the disease-causing gene (13).

To the extent that the aim is restriction of access third parties have to genetic information about specific individuals, a focus on limiting access to the results of genetic tests may be too narrow. However, the case for employing a broader conception of genetic information presents a dilemma; the fact that matters of family history and the results of biochemical tests are integral parts of patients's medical records undermines efforts to afford more stringent protection for genetic information than other types of medical information. Moreover the goal of improved clinical care will necessitate increased availability of genetic information, including the incorporation of results of individual tests into the patient record. Thus effective protection of genetic privacy often will be dependent upon the development of effective medical privacy protection policies generally (14).

The special privacy concerns about genetic information mirror similar concerns for other highly sensitive medical information that, if made available to third parties, may undermine individual autonomy and result in arguably unfair economic and social burdens on those whose genes have been identified. There is already a fairly well-established social consensus regarding the special sensitivity of medical information bearing on sexual behavior, mental health, sexually transmitted diseases, alcoholism, and drug use. These matters are deemed especially sensitive because they reveal intimate details of an individual's life and often carry a degree of social stigma or may result in discrimination in employment or insurance. Even without stigma or loss of social opportunities, many persons will view public knowledge of intimate aspects of an individual's life as an assault on human dignity and the ability to exercise autonomous control over matters at the heart of individual personality (3).

Genetic information can be seen as analogous to these already established categories of sensitive medical information because of similar risks associated with their disclosure to others, and accordingly it is not unreasonable to argue that genetic information should be added to the list of medical information for which added privacy precautions are warranted. Moreover, to the extent that stigmatized medical conditions such as alcoholism and mental illness have a genetic component, the case for more stringent privacy protection of these types of genetic information is strengthened.

There are other aspects of genetic information that enhance the arguments for the view that the collection of genetic information poses even greater privacy risks than other similarly sensitive medical information. One such argument is that genetic information, unlike cholesterol screening or testing for treatable sexually transmitted diseases, has more significant familial and reproductive implications. Because testing of one person can reveal information about an individual's siblings, parents, and children the privacy interests at stake go well beyond the risks any unwanted disclosure poses for a single individual (14). Hence, even if the genetic information arguably is no more sensitive or intimately revealing than some other types of nongenetic information, the privacy interests of many more persons may be implicated with each instance of privacy breach.

Another argument for the special moral importance of genetic privacy is the fact that much genetic information is merely probabilistic, and that learning the existence of some genetic susceptibility to disease often does not predict a particular individual's risk of acquiring a disease, its severity, or its date of onset. Thus the disclosure of genetic information is said to carry an inherent risk of being used to falsely label an individual on the basis of an increased statistical risk of having a mental or physical disease that has not and indeed may not result in any symptoms or disability (5,15). Of course, many other kinds of medical tests are merely probabilistic indicators of increased risk of disease as well, and genetic information may not be unique in this respect (11). However, the deeper concern seems to be that genetic predictors often are mistakenly seen as deterministic, immutable guides to assessing an individual's mental and physical capabilities (15).

A further argument for the uniqueness of genetic information and the need for more stringent genetic privacy protection lies in the observation that our DNA contains hidden information that may in the future reveal more about us than either we or those who collect and analyze DNA samples now realize. DNA has been called a "future diary," "a code not yet cracked," holding currently undecipherable but highly personal, sensitive information bearing on one's medical and social prospects (6,7,16). Two points are of special concern in this claim. First, the availability of DNA samples that have been analyzed for one gene may be subject to later analysis for another gene whose existence and function are not known at the time of consent to testing. It may not be reasonable to suppose that those who have consented to tests revealing one type of genetic information would have consented to the disclosure of genetic information of another type.

A second worry is that even testing for an allele with a known function can in the future reveal knowledge about an individual that no one could have anticipated at the time of testing. Genes typically have many functions, only some of which are known. A single gene may contribute to some disease burden such as sickle cell anemia but that same gene also may play a protective role such as improving one's resistence to malarial infection. Because there are no inherently "good" or "bad" genes, even testing for a "good" gene today may reveal the same gene as a "bad" gene tomorrow. Because persons who consent to testing may not be able to anticipate all that is revealed by a single test, adherence to the duty of confidentiality is a crucial moral component of any comprehensive genetic privacy protection approach.

It remains an important empirical question whether genetic information is inherently more sensitive than other types of information, and whether there is likely to be any consensus on what types of information are more important to protect than others. Certainly, all types of genetic information are not on par in their sensitivity, given the differences in what genetic tests may reveal about individuals and the potential consequences of their disclosure. Some persons may be more concerned with limiting the information others have about their sexual behavior and drug use than their increased genetic susceptibility to beryllium disease. Or a well-known trial lawyer may have great concern about the public disclosure of a heart condition for fear that potential clients will doubt his capability for aggressive advocacy, while being indifferent to what others know about his sexual conduct.

Individual differences in privacy attitudes are of tremendous moral significance. To the extent that the moral arguments for privacy protection are grounded chiefly in the protection of individual autonomy and dignity, matters that ought to be accorded the greatest protection are largely for the individual to decide for herself. However, to the extent that privacy interests are grounded in more general concerns for the protection of valuable social opportunities, then the issue of whether genetic information is more sensitive and therefore deserving of greater protection will depend on the extent to which the configuration of the major social and economic institutions leave persons more vulnerable to unfair disadvantage based on individual genotype.

Given uncertainties of whether genetic information is uniquely sensitive, its increasing integration with other types of medical information, and the competing moral rationales offered for its protection, debate about how best to deal with the protection of genetic privacy is likely to continue for some time.

RESEARCH CONTEXTS

Many who have considered the moral demands of privacy within research contexts argue for individual control over the creation of individually identifiable genetic information. Concerns about individual privacy take on great significance when there is a more than minimal risk of privacy loss for the subject and the contribution of DNA samples by research subjects offers little or no realistic expectation of any personal benefit. Such protections are embodied in current guidelines for federally funded medical research in the United States, and others also argue that this should be the norm in clinical practice, especially when collection of samples involves subsequent research uses of genetic information (9). These policy guidelines and suggested moral norms recognize the fundamental importance of obtaining informed voluntary consent to the collection of blood or other DNA samples for a variety of moral reasons, including but not limited to privacy concerns. Thus there is broad consensus that

individually identified stored blood or tissue samples ought not be made available to others for subsequent analysis without the explicit consent of the persons providing the samples, and that in instances where additional testing not contemplated at the time of obtaining the initial informed consent is proposed researchers should recontact subjects to obtain appropriate consent for these new uses (8,9).

Moreover many argue that in addition to the routine disclosure of medical risks associated with research, risks of privacy loss should be included among the core elements of the informed consent process (17,18). When links to identifiable persons who contributed genetic material to a study are retained, researchers cannot guarantee that confidentiality will not be compromised, but they can and should provide subjects with information about the steps they plan to take to minimize the risks to confidentiality.

One issue that continues to divide knowledgeable commentators is when informed consent for analysis of DNA samples is necessary. Some have concluded that identified or identifiable subjects should be given the option of restricting the use of their samples to the study for which samples were initially collected (8). Others also recommend that since many samples are obtained in nonresearch settings, all individuals should be given the opportunity to decide whether their samples may be used in research and whether they are willing to have those samples used in identifiable or linked research (9).

An alternative approach is to ask research subjects to give a blanket consent to additional testing for other genes or other genetic conditions not examined initially and perhaps not possible for researchers to contemplate at the time of obtaining informed consent (6–8,17,18). The main benefit of such a proposal would be one of efficiency. Researchers would have ready at hand a supply of DNA samples, perhaps in the form of immortalized cell lines, which can be used as a resource for many research studies, and in the process save both time and money associated with collecting and preserving DNA samples.

The major drawback in such a proposal is a concern over whether is it plausible to suppose that genuinely informed consent is possible for types of research not within the contemplation of the research community. In particular, informed consent to subsequent research uses may pose privacy risks (as well as other moral concerns about use or purpose of the research itself) that differ in kind or in magnitude from those that might be recognized initially. Even if the probability of privacy loss is not increased with re-testing for a different genetic condition, the magnitude of harm from an unwanted identification of any individual may be much greater than the harm the subject would have had in mind at the time of contributing the DNA samples. For example, agreeing to be tested for a remote risk of unwanted disclosure of some genetic mutation with no obvious adverse implications for a subject's ability to obtain insurance is quite different from testing for a different condition that does have insurance implications. The upshot of any policy of obtaining broad, open-ended consent for genetic testing for research purposes is that the subject may end up waiving any rights to protect his or her privacy interests without any feasible basis for knowing how significant those privacy interests might be.

Another area of continuing debate concerns the use of anonymous and anonymized samples and how best to deal with the risk of inferential identification of individual subjects mentioned previously. The recommendations of some commentators restrict the requirement of informed consent for research to instances in which actual personal identifiers are linked to the samples and are retained by the researchers performing the genetic tests. They therefore conclude that no informed consent is necessary either for prospective or retrospective tests when samples are collected anonymously or later are anonymized (8). Where genetic analysis is performed and the identities of subjects are removed from the samples but individual subjects remain directly identifiable through means of a code maintained by researchers, informed consent for the use of such samples is claimed to be necessary for prospective studies and usually necessary for retrospective studies, unless there is no more than minimal risk to privacy and the research could not be carried out practicably otherwise.

Others, however, suggest that in some instances, even anonymized samples may pose nonnegligible risks to individual privacy when it is possible for researchers or others to identify individual persons having a genetic mutation putting them at increased risk for some specific disorder. In such cases they conclude that there may be a need for recontacting individuals to obtain informed consent for subsequent use of samples originally contributed for a different purpose, and that only genuinely anonymous samples — where it is impossible under any circumstances to identify the individual source — clearly qualifies as exempt from any need for informed consent or other institutional efforts to safeguard individual privacy (9). Critics, however, point out that conducting research without specific informed consent has been the norm for most retrospective studies using patient medical records not involving genetic analysis, and while the critics do not reject categorically the stricter recommendations, they do worry that the proper balance between the legitimate aims of public health research and individual privacy requires further reflection and public discussion (19).

Inferential identification of individual subjects may occur in a variety of ways. Suppose that a group of researchers obtains DNA samples for a longitudinal study of several hundred members of families with a high prevalence of some type of cancer (e.g., cervical cancer) and that as a part of that study researchers also collect detailed medical histories of each study participant. The researchers therefore need to retain identifying links to each individual that are necessary for follow-up and for purposes of contacting the subjects with new information that might be of medical benefit to those subjects. Researchers therefore promise only confidentiality because the needs of the study as well as the medical needs of the subjects make anonymous collection of DNA samples undesirable. Suppose then that a later date a second group of researchers propose to conduct a different study of another medical condition

also suspected of greater than average prevalence in those families, and they therefore seek permission to reanalyze the DNA samples for a second gene mutation. Even if the second group of researchers receive the samples without individual identifiers — an anonymized study — publication or other dissemination of study results may enable family members to infer the identities of specific individuals who were tested and thereby reveal previously undisclosed information about who among those families had been affected either by cervical cancer or the second medical condition. Moreover publication of detailed family pedigrees showing family lineages and various characteristics — including additional medical data such as miscarriages or abortions potentially relevant to findings regarding cervical cancer risk — can magnify both the likelihood of inferential identification of specific individuals and the harm flowing from inadvertent disclosure of other sensitive medical information in the process (10).

The risk of inferential identification may be increased, for example, when the initial pool of subjects is small, subjects are drawn from a patient population publicly associated with some particular medical or research institution, the second medical condition tested for is rare, or some discernible phenotypic characteristics are described in the publication of the study findings (9). The publication of study results may allow individual researchers, and perhaps more significantly, other family members, to more easily infer the identities of individual subjects.

Genetic research poses other privacy risks not ordinarily raised by many other types of medical research. It is not only patients whose medical records may be examined by researchers and research subjects who voluntarily contribute DNA samples for analysis that risk having others learn information about mutations predisposing them to genetic disease or their status as carriers of mutations that may affect their offspring. Personal genetic information about individuals with no personal relationship to researchers or medical caregivers may appear in another family member's medical records used by researchers or become part of a data bank or genetic registry developed and maintained by researchers from information obtained from probands (5,14,17). Detailed, comprehensive, and accurate family histories may be essential to genetic linkage studies, and thus sensitive and perhaps inaccurate information about family members not participating in the research may be obtained without their permission or knowledge.

Some risks of breach of confidentiality can be minimized through efforts to build a strong wall of separation between these specialized research records and other medical data banks and individual patient records. While the risks may be small that persons will be harmed through third-party access to these records, individuals concerned about privacy on the ground that respect for human dignity requires an ability to control the kinds of sensitive personal information others can maintain in data banks will not agree that confidentiality protections are sufficient. As the boundaries between research and clinical practice erode the feasibility of more stringent privacy protections for genetic information is diminished and the burdens of privacy protection will fall increasingly on the shoulders of medical practitioners, genetic counselors, and researchers alike (20).

CONFIDENTIALITY IN CLINICAL CONTEXTS

The fact that genetic testing of one individual can reveal information relevant to the health status and medical care of other family members is the source of one of the most hotly debated moral issues in clinical genetics. When patients refuse to disclose genetic information that would benefit other family members, health care professionals are presented with a dilemma: They must choose whether to preserve patient confidentiality or to breach that duty of confidentiality for the benefit of third parties. The dilemma is often especially keenly felt when information regarding increased familial risk for some disease can be used by other family members to seek medical interventions that may prevent or delay the onset of disease or reduce its likely severity (e.g., familial breast or colon cancer), make informed reproductive decisions (e.g., CF, thalassemia, or Tay-Sachs), or make social and financial plans to cope with an anticipated physical or mental impairment when no treatment is available (e.g., Huntington's disease).

The dilemma, of course, is not unique to cases involving physicians, genetic counselors, and the family members of patients; researchers without any clinical connection to the person tested may obtain information bearing on the well-being of others, and interests other than those of family members may be at stake (21). For example, a widely discussed case involves researchers who learn that an airline pilot has the gene for Huntington's disease, which ultimately will impair his ability to perform his job safely.

Nonetheless, the dilemma is often felt more acutely in contexts involving clinical caregivers and those who are genetically related to their patients. The prospect of breaching patient confidentiality runs against the grain of the clinician's deepest professional commitments (22). That commitment in its strongest form is reflected is reflected in the fifth-century B.C. Hippocratic Oath: "What I may see or hear in the course of the treatment or even outside the treatment in regard to the life of men, which on no account one must spread abroad, I will keep to myself holding such things shameful to be spoken about."

More recent formulations of a physician's duties of confidentiality are more permissive. A 1980 statement of the American Medical Association (AMA), for example, states that a physician shall "safeguard patient confidences within the constraints of the law." The implication of the AMA statement is that legal requirements may override moral duties of confidentiality, and in fact court decisions in the famous *Tarasoff* case and in similar cases subsequently have provided much of the grist for moral discussions of the morally permissible exceptions to the ordinary duties of confidentiality (22,23). Although the *Tarasoff* case involved the issue of whether a psychologist must breach confidentiality in order to protect a third party from impending threats of bodily harm learned in the course of a therapeutic relation, both courts and the

numerous legal and ethical commentators have extended the reasoning to other contexts as well. For example, analogous duties to disclose information to persons at risk of harm have been argued for in the context of contagious diseases such as HIV. Predictably lawyers and ethicists have pondered the extension of these duties to cases involving genetic information.

The legal and ethical issues are, of course, distinct, but the reasoning relied upon in both realms is strikingly similar and the answers given by lawyers rely heavily on ethical concerns (22–25). Both raise the fundamental question, Under what conditions, if at all, a health care professional's duty to protect third parties from harm outweighs the duty to maintain the confidentiality of genetic information obtained from a patient? A secondary question is, To what extent ought these moral duties be reflected in the law?

Several positions on these questions can be found in the literature. Some, such as the members of the President's Commission on Biomedical Issues, the Nuffield Council on Bioethics, and the authors of the recent Institute of Medicine report on genetic testing counsel against disclosure of any medical information, including genetic test results, except in very narrowly defined circumstances (3,23,26). The analytic framework of the President's Commission provides the starting point for most of the discussions the exceptions to confidentiality for genetic information. They recommended four necessary conditions that should be met before disclosure is justified:

1. There must have be an unsuccessful attempt to persuade the patient to disclose the information to the relevant party.
2. There must be a high probability of harm without disclosure and a high probability that the disclosure itself will avert the anticipated harm.
3. The potential harm must be serious.
4. Only the degree of informational detail necessary to avert harm should be disclosed.

These criteria restrict significantly the instances in which confidentiality may be overridden. For example, in cases where no effective treatment is available, no disclosure would be warranted. Similarly, where third parties have other opportunities through which they may learn of their increased genetic risk for some specific disease, the duty of confidentiality would not be overridden.

Other commentators appear sympathetic to a more liberal attitude toward disclosure in a broader array of cases. They appear more concerned, for example, about the role of health professionals in promoting the health of the entire family and often speak of the family unit rather than any individual as the patient. When conceptualized in this fashion, strictly speaking no duty of confidentiality is broken. If, as Dorothy Wertz and John Fletcher claim, "hereditary information is a family possession rather than simply a personal one" the dilemma of confidentiality versus the duty to protect others is dissolved, and the issue is recast as a question about how best to discharge the health care professional's duty to make decisions for the best interests of their patients collectively (27).

The shift to viewing the family as patient nonetheless fails to resolve what may be the central issues lurking behind claims on behalf of patient confidentiality in the context of these familial disagreements. Important questions include whether health care professionals should be presumed to possess superior knowledge and insight about what is best in these circumstances, and whether they should be seen as having the moral authority to act as enforcers of a view of family communication or of an ideal of familial loyalty and connectedness which everyone may not share. Indeed, part of the moral justification given by the defenders of patient privacy is the claim that individuals, not governments or members of professional elites, ought to control highly personal, sensitive information, even in cases where failure to disclose such information may make it harder for others to take steps to prevent harm to themselves.

Others cite a second reason for their reluctance to leave it to health care professionals to decide when to breach patient's confidentiality. They argue that genetic harm is different from some other cases in which the duty to protect from harm is thought to outweigh the duty of confidentiality. They note that the patient, whose interest in confidentiality is sacrificed, is not normally the direct causal vector of potential harm to third parties (24). Disclosure of patient information in the genetic context imposes burdens on persons who simply fail to assist others, not on persons who, for example, through their own violent or risky behaviors impose risks on innocent third parties. Thus, unlike the psychiatric or HIV cases, where it can be claimed that individuals, through their own behavior, may have forfeited their privacy rights, no such claims can be made for genetic information.

Whether a particular factual situation meets the criteria set out by the President's Commission will vary considerably. To the extent that there is any agreement on these thorny issues, most concede that case for breach of patient's confidentiality is strongest when it involves a disclosure of information obtained from a patient who is the direct and morally culpable cause of another's harm, and the harm to third parties is grave, not otherwise preventable, and immediate.

NONMEDICAL USES OF GENETIC INFORMATION

The range of potential nonmedical uses of genetic information is lengthy and limited only by available technology and imagination. Moreover, the loss of genetic privacy may result from the generation or disclosure of genetic information gathered in nonmedical contexts, many of which are not governed by established ethical and legal norms of privacy governing conduct in medical and research arenas. A few widely discussed examples illustrate the extent to which genetic privacy concerns arise outside of traditional clinical and research contexts.

Two of the nonmedical uses of genetic information receiving the most critical attention arise within employment and insurance contexts. The use of such information to deprive a person of employment or health, life or disability insurance, are among the most economically consequential implications of a loss of genetic privacy (28,29).

The connection between genetic privacy issues related to employment and various forms of insurance is especially close in the United States, since a great majority of working persons obtain their insurance as a benefit of employment (30). The consequence, for example, of an employee having some genetic predisposition for increased susceptibility to some substance found in the industrial workplace are powerful economic incentives for employers to avoid hiring persons who may be both expensive to maintain as a member of its group health insurance plan and may in the future expose the employer to some additional risk of greater tort liability for employee illnesses developed after exposure to workplace hazards.

Many other potential nonmedical uses of genetic information also may arise as a consequence of their relevance in judicial proceedings. One such use now in the public spotlight is the collection of DNA samples for criminal forensic use (31). While the use of what is popularly known as DNA fingerprinting may aid in the identification of criminals and the exoneration of persons falsely accused of crimes, the widespread and systematic collection of DNA samples from convicted felons has some attendant privacy risks. The use of DNA in forensic analysis relies on a pattern of DNA sequences often referred to as junk DNA because they do not reveal specific portions of the genome believed to be associated with any medical condition, increased susceptibility to disease, or a person's carrier status. However, state and federal law enforcement agencies that collect DNA samples for identification purposes often retain the original DNA sample, and it can be analyzed and sensitive genetic information about individual persons may be revealed. In addition to the routine ways genetic information may be used to the detriment of any individual, such information could further be used as a basis for subsequent judicial decisions about sentencing, parole, or postconviction confinement to mental health or sex offender treatment programs (32).

The potential uses of genetic information in civil litigation are limitless as well. In addition to the defensive use employers might make of such information in exclusionary hiring and other workplace decisions, employers may also assert genetic predisposition as a partial defense against lawsuit by employees, consumers, and members of the general public who are exposed to substances in their products found to contribute to illnesses that also have a genetic component in their development. The scenario some envision is one in which genetic privacy may be routinely sacrificed for the sake of apportioning legal liability for harm suffered by workers and consumers (33). Beyond the strictly legal aspects of such uses of genetic information the use of genetic information in judicial contexts raises profound ethical worries about the fairness or justice of how risks and benefits are distributed within society's major social institutions.

Other morally problematic uses of genetic information in judicial contexts raise related issues of how individual responsibility for conduct and wrongdoing are assessed and the fitness of individuals to occupy professional roles involving a high degree of trust and fiduciary integrity. For example, in a California disbarment proceeding, DNA evidence was used to show that a lawyer who had embezzled client funds was genetically disposed to alcoholism (34). Although the California case resulted in the lawyer being placed on probation on the grounds that his genetic predisposition should be viewed as a mitigating factor in assessing his responsibility for his ethical lapses, other equally controversial uses of genetic information can be imagined as well. For example, some might argue that genetic information should be relevant in the assessment of an individual's fitness for entry in to professions such as law, medicine, or nursing. The moral objection this prospect raises — apart from the obvious worries that statistical associations between certain genes and behavior may not survive subsequent scrutiny by researchers — is that such information may be both a poor predictor of any specific individual's behavior and too intrusive into the private lives of individuals.

A few final examples reflect the potential uses of genetic information by courts, schools, and parents in making decisions affecting the custody and education of children. If detailed genetic information is added to the existing body of health information prospective adoptive parents can receive about children, problems of hard-to-place children will be exacerbated. If couples come to view genetic information as relevant to their decisions about adding an abandoned child to their family the risk is that more of the already unfortunate children will be put in an even greater disadvantage (35).

Similarly a parent, grandparent, or other party to a divorce or custody battle may make information regarding an opposing litigant's genetic illness or predisposition to illness a factor in determining the child's best interest, which is the touchstone of how judges make these difficult decisions (34). In addition schools and day care centers may assert a need to know genetic information about a child in order to make assessments of potential learning disabilities or judgments regarding the behavioral traits bearing on their suitability for enrollment (36).

The use of genetic information by those who resort to judicial proceedings to resolve disputes related to family life is already a fact of modern life, and the extent that role such information may hasten greater governmental intrusion into family life remains an open question.

CONCLUSION

In each context discussed in this article, we have seen how the increased collection and use of genetic information adds layers of complexity to old and familiar ethical problems, takes us into novel and poorly charted moral terrain in balancing privacy claims against other legitimate social goals, and often requires a fresh look at the moral foundations and value assumptions underlying the practices of many of our most basic social, medical, legal, and economic institutions.

BIBLIOGRAPHY

1. J. Rachels, *Philos. Public Affairs* **4**, 323–333 (1975).
2. J. Reiman, *Philos. Public Affairs* **6**, 27–44 (1976).

3. L. Andrews et al., *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994.
4. J. Benkendorf et al., *Am. J. Med. Genet.* **73**, 296–303 (1997).
5. M. Powers, in M.S. Frankel and A. Teich, eds., *The Genetic Frontier: Ethics, Law and Policy*, AAAS Press, Washington, DC, 1994, pp. 77–100.
6. G.J. Annas, *JAMA, J. Am. Med. Assoc.* **270**, 2346–2350 (1993).
7. G.J. Annas, in T.F. Murphy and M.A. Lappé, eds., *Justice and the Human Genome Project*, University of California Press, Berkeley, 1994, pp. 75–90.
8. American Society of Human Genetics, *Am. J. Hum. Genet.* **59**, 471–474 (1996).
9. E.W. Clayton et al., *JAMA, J. Am. Med. Assoc.* **274**, 1786–1792 (1995).
10. M. Powers, *IRB: Rev. Hum. Subj. Res.* **15**, 7–11 (1993).
11. T. Murray, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 60–73.
12. M. Yesley, *Microb. Comp. Genom.* **2**(1), 9–35 (1997).
13. R. Wachbroit, *Rep. Inst. Philos. Public Policy* **9**, 9–11 (1989).
14. L.O. Gostin, *J. Law, Med. Ethics* **23**, 320–330 (1995).
15. E. Juenst, *Genome Sci. Technol.* **1**(1), 21–30 (1995).
16. G.J. Annas, *Am. J. Public Health* **85**, 1196–1197 (1995).
17. P.R. Reilly et al., *Nat. Genet.* **15**, 16–20 (1997).
18. R.F. Weir and J.R. Horton, *IRB: Rev. Hum. Subj. Res.* **17**, 1–4 (1995).
19. B.M. Knoppers and C. Laberge, *JAMA, J. Am. Med. Assoc.* **274**, 1806–1807 (1995).
20. G. Geller et al., *J. Law, Med. Ethics* **21**, 238–240 (1993).
21. P. Harper, *Br. Med. J.* **306**, 1391–1394 (1993).
22. D. Orentlicher, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 77–91.
23. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Screening and Counseling for Genetic Conditions*, U.S. Government Printing Office, Washington, DC, 1992.
24. S. Suter, *Mich. Law Rev.* **91**, 1854–1908 (1993).
25. M.Z. Pelias, *Am. J. Med. Genet.* **39**, 347–354 (1991).
26. Nuffield Council on Bioethics, *Genetic Screening: Ethical Issues*, Nuffield Council on Bioethics, London, 1993.
27. D.C. Wertz and J.C. Fletcher, *Bioethics* **5**, 212–232 (1991).
28. *Genetic Information and Insurance*, Task Force Report, Genetic Information and Health Insurance, National Institutes of Health, National Center for Human Genome Research, Bethesda, MD, 1993.
29. N. Kass, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 299–316.
30. M.A. Rothstein and B.M. Knoppers, *Eur. J. Health Law* **3**, 143–161 (1996).
31. J.E. McEwen, *Am. J. Hum. Genet.* **56**, 1487–1492 (1995).
32. B. Scheck, *Am. J. Hum. Genet.* **54**, 931–933 (1994).
33. M.A. Rothstein, *J. Law Health* **9**, 109–120 (1994–1995).
34. R.C. Dreyfuss and D. Nelkin, *Vanderbilt Univ. Law Rev.* **45**, 313–348 (1992).
35. L. Andrews, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 255–280.
36. L. Rothstein, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 317–331.

See other entries GENETIC INFORMATION, ETHICS, AND INFORMATION RELATING TO BIOLOGICAL PARENTHOOD; GENETIC INFORMATION, ETHICS, ETHICAL ISSUES IN TISSUE BANKING AND HUMAN SUBJECT RESEARCH IN STORED TISSUES; GENETIC INFORMATION, ETHICS, FAMILY ISSUES; GENETIC INFORMATION, ETHICS, INFORMED CONSENT TO TESTING AND SCREENING; GENETIC INFORMATION, LAW, LEGAL ISSUES IN LAW ENFORCEMENT DNA DATABANKS.

GENETIC INFORMATION, LAW, LEGAL ISSUES IN LAW ENFORCEMENT DNA DATABANKS

PAUL GIANNELLI
SHARONA HOFFMAN
WENDY WAGNER
Case Western Reserve University, School of Law
Cleveland, Ohio

OUTLINE

INTRODUCTION

Jean Ann Broderick was sexually assaulted and murdered on November 17, 1991, in Minneapolis. There were no suspects, and the possibility of another unsolved crime loomed large. The police, however, discovered semen at the crime scene, extracted a DNA profile from this evidence, and entered the profile into the state DNA databank.

The computer responded with what is known as a "cold hit" — a match that in an electronic second transformed a "no suspect" case to one with overwhelming prosecutorial merit. It was the "first case in American history in which the new tool of DNA databanking was used to solve a rape or murder case" (1). The prosecutor would later remark, "Without a DNA pool, there is no way we would have been able to identify the suspect. And we certainly would not have been able to get the conviction" (1).

As the Broderick case illustrates, DNA databanks are a significant advancement in crime solution. As of January 1999, over 400 "hits" have been recorded. These databanks are similar to the computerized fingerprint databank, called AFIS (Automated Fingerprint Identification System), which has been operational during the last decade. In some ways DNA databanks may be of greater utility than AFIS. While the wearing of gloves prevents the leaving of fingerprints, it is more difficult to prevent the deposition of some type of evidence that contains the perpetrator's DNA — especially in rape cases. Indeed, Virginia officials claim that "material susceptible to DNA analysis, including blood, skin tissue, hair follicles, and semen, may be found at thirty percent of all [violent] crime scenes" (2). Saliva and sweat should be added to this list (3). Furthermore, fingerprints cannot be dated; they can place a suspect at a specific location but cannot, by themselves, establish when the suspect was there, a significant limitation in cases in which the suspect has innocent access to the crime scene location. In contrast, semen found in a rape victim eliminates the "dating" problem in cases where the suspect claims mistaken identification.

An understanding of the role of DNA databanks in the criminal justice system requires some appreciation of the impact of DNA evidence in criminal prosecutions.

DNA EVIDENCE

In 1985 Dr. Alec Jeffreys of the University of Leicester, England, recognized the utility of DNA profiling in criminal cases. Its first use in American courts came the following year. The initial appellate case, *Andrews v. State* (4), was reported in 1988. By January 1990 forensic DNA evidence had been admitted in at least 185 cases by 38 states and the U.S. military (5). Today DNA evidence, in one form or another, is admissible in every state and federal circuit (6). These developments are remarkable. No other scientific technique had gained such widespread acceptance so quickly. No other technique had been as complex or evolved so rapidly. DNA profiling raised issues at the cutting edge of modern science (7). New DNA technologies were introduced even as cases litigating the older procedures worked their way through the court system; there have already been three generations of tests — Restriction Fragment Length Polymorphism (RFLP), Polymerase Chain Reaction (PCR) for discrete alleles, and the current state of the art, Short Tandem Repeats (STRs). In addition, testimony based on mitochondrial DNA has been admitted in evidence (8). Moreover, nonhuman DNA has proved useful in litigation, "ranging from homicide prosecutions to patent infringement litigation, with organisms as diverse as household pets, livestock, wild animals, insects, plants, bacteria, and viruses" (9).

Finally, no other technique has been as potentially valuable. One court commented that DNA evidence may be the "single greatest advance in the 'search for truth' ... since the advent of cross-examination" (10). Even its critics acknowledge that "[a]ppropriately carried out and correctly interpreted, DNA typing is possibly the most powerful innovation in forensics since the development of fingerprinting in the last part of the 19th Century" (11). For instance, in the World Trade Center bombing prosecution, an FBI expert matched saliva on an envelope sent by the terrorists to the *New York Times* with the DNA of one of the defendants (12). The next crime tool on the horizon is a credit-card-size device that would permit the analysis of DNA at a crime scene (3).

DNA evidence's power to convict is matched by its power to exculpate. This was underscored by a Department of Justice report that discussed the exoneration of 28 convicts through the use of DNA technology — some of whom had been sentenced to death (13). By mid-1999, more than 70 convicts had obtained postconviction relief based on exculpatory DNA test results. This development has already resulted in a change in some legal procedures. For example, the basis for motions for new trials has historically been quite limited; after a trial with the full panoply of constitutional protections, "finality" becomes a significant, if not paramount, interest. Thus, courts are skeptical of witnesses who subsequently "change" their testimony or post-trial "confessions" by unavailable third parties. Due to its reliability DNA evidence alters the calculus between "finality" and justice. Consequently, New York and Illinois have statutorily extended the time period for post-trial challenges to convictions based on DNA evidence, and the Department of Justice's Commission on the Future of DNA has advocated adoption of similar provisions in other jurisdictions (14).

Similarly, legislatures are reconsidering time limits on statutes of limitation in criminal cases (15). Ohio, for instance, has increased its statute of limitations for felonies from 6 to 20 years. DNA evidence reduces, to some extent, the danger of conviction based on evidence that is unreliable because it is stale. Indeed, to toll the Wisconsin statute of limitations, one creative prosecutor indicted a "John Doe" rapist based solely on his DNA code (16).

DNA DATABANKS

There are a variety of organizations that collect samples for DNA analysis — the Department of Defense collects blood and tissue samples from every U.S. service member, states authorize laboratories to secure dried blood samples (and often other tissue samples) from newborns, and reproductive laboratories and blood banks store samples from patrons. Each of these DNA systems raises a number of important privacy concerns (17). This discussion will focus, however, only on DNA databanks created by state or federal law for criminal enforcement purposes. These databanks present the full range of privacy and related issues in the much more dramatic context of criminal

enforcement, with the government, rather than a private employer or insurer, poised as the entity that threatens an individual's right to genetic privacy.

The first DNA databank used for criminal enforcement purposes was established by the Virginia legislature in 1989 (18). Today every state has enacted databanking legislation (19). The DNA Identification Act of 1994 provides federal funds to assist in this endeavor (20). Although each state legislates the conditions under which DNA samples are taken, the FBI has established a national databank system, called CODIS (Combined DNA Index System), into which the state profiles can be entered (21,22). Now states can search the databases of other states (23).

The state databank statutes vary widely with respect to their coverage. Some states require only sex offenders to provide samples for databank use (24). Other states also include different crimes of violence (25). Still others reach all convicted felons (26). One statute extends to those arrested for felony sex offense crimes or other specified offenses (27). Some states include juvenile offenders (28,29) and others cover probationers as well as parolees (30). Several databases also contain DNA profiles of missing persons and victims of mass disasters (31). The method of collection differs; some statutes require the collection of blood (sometimes a finger prick) (32) while others collect cheek swabs (33). Some statutes contain expungement procedures, under which a person's profile may be removed from the database if that person's conviction is reversed on appeal (19,34).

The state databanks also vary in other respects. Some states have legislated the circumstances under which their database can be searched. For example, two states allow their databases to be searched only for the investigation of sex-related or violent crimes (19). States also vary in the resources dedicated to DNA collection and analysis. Some jurisdictions have made considerable headway entering samples into their databases, while others face a tremendous backlog of samples yet to be analyzed (19,35,36). One report notes: "So while a new national FBI databank and state databanks now hold a total of 270,000 DNA profiles, there is also a backlog of roughly 500,000 unanalyzed DNA specimens. And the DNA of an estimated 1 million more people is supposed to be added by law, but some jurisdictions are already so far behind they're not even bothering to collect new samples" (37).

While variations in the coverage and procedures for state databases produce inconsistencies, state databanks do share important similarities as well. First, DNA profiles are generally kept in a database that identifies them by a coded identification number. To determine the identity of the person, a separate database must be accessed that decodes the identification number and links the profile to a specific individual. These security measures help to ensure that the DNA profile does not provide readily usable information about the identity of a particular individual (38). Second, databases generally contain one set of DNA profiles that have been taken from identified individuals, and a second set of profiles, usually taken from crime scenes, for which a match is sought (39). If a crime scene profile does not result in a match, it remains in the system. Some time in the future it may be matched with the profile of a subsequently convicted offender (39). Or, it may be matched with another crime scene profile, alerting the police that they are looking for a serial offender (40).

The present success and future potential of these databases for determining the identity of criminal perpetrators is clear. Commentators have suggested that dramatic deterrence and criminal enforcement benefits will be gleaned from the enhanced enforcement potential (19,36,39). Much of the criticism of these databases, in fact, focuses not on the lack of benefits to criminal enforcement but rather on deficiencies in confidentiality assurances that protect individuals' remaining rights to privacy.

PRIVACY AND RELATED CONCERNS

The privacy issues associated with DNA profiling were recognized from the beginning. In 1990 Congress's Office of Technology Assessment highlighted this issue: "Citing the inherent intimacy of genetic information, the current and developing ability to test for personal information other than unique identity, and the difficulties of maintaining the confidentially in a computer network, experts raise concerns that genetic information could be used unfairly to deny future benefits to persons with criminal records, and that genetic profiling within the criminal justice sphere could lead to wider testing and broader threats to privacy" (5, p. 35).

The National Academy of Science's 1992 DNA report also took note of privacy concerns, citing developments in both molecular biology and computer technology. "Molecular geneticists are rapidly developing the ability to diagnose a wide variety of inherited traits and medical conditions. The list already includes simply inherited traits, such as cystic fibrosis, Huntington's disease, and some inherited cancers. In the future, the list might grow to include more common medical conditions, such as heart disease, diabetes, hypertension, and Alzheimer's disease. Some observers even suggest that the list could include such traits as predispositions to alcoholism, learning disabilities, and other behavioral traits (although the degree of genetic influence on these traits remains uncertain)" (41). The report goes on to state: "Even simple information about identity requires confidentiality. Just as fingerprint files can be misused, DNA profile identification information could be misused to search and correlate criminal-record databanks or medical record databanks. Computer storage of information increases the possibilities for misuse. For example, addresses, telephone numbers, social security numbers, credit ratings, range of incomes, demographic categories, and information on hobbies are currently available for many of the citizens in our society from various distributed computerized data sources" (41).

Privacy concerns also arise from inadequate confidentiality protections for the DNA profiles, and more importantly, for stored samples in many state DNA databanks (17,42). Only some of the states provide a meaningful penalty for the unauthorized use of DNA samples or profiles by private parties (43–45). There also appear to

be few, if any, common law remedies available to deter this avenue for abuse, although there has been some discussion about legislative reforms that would provide individuals with protected privacy rights in their genetic information (45). It is not clear whether the unauthorized use of DNA profiles will ultimately present serious privacy intrusions, since the profiles are done with "junk DNA" that is currently believed to reveal little if any information about personal traits (46). Inadequate protection of the original biological samples is a completely different matter, however, and there is unanimity that these samples contain very private information about an individual (38,47,48). Yet in some state databank laws, the original biological samples receive less protection than the DNA profiles (19,48). The danger that DNA information contained in original samples may be disseminated in unauthorized ways becomes even more worrisome, since states may retain samples indefinitely in order to adapt to possible future changes in the profiling system (48).

In sum, DNA databanking will be a powerful tool in solving, and perhaps deterring, crime, but the possibility for misuse exists — probably in ways that cannot be anticipated. Some prior precedents demonstrate the possibility of abuse. For example, a former assistant U.S. Attorney "recalls an incident from his days as a prosecutor in the 1970s in which police officers were caught selling confidential police records to private investigators" (39). The harm that can result from unauthorized release of genetic information could be still greater. Because of privacy concerns, bioethicist Eric Juengst argues that "any DNA taken for identification purposes should only be typed for information-free markers" and "no physical DNA samples should be banked" (44, p. 64).

LEGAL CHALLENGES

Databanks have been challenged on a wide range of constitutional grounds — for example, freedom of religion (49) and the right to privacy (50). Also statutory attacks under the Religious Freedom Restoration Act have been advanced (51). Several attacks have been quite creative. For example, the Tenth Circuit has rejected arguments that taking DNA samples violates the Ninth Amendment and deprives offenders of a property interest in their blood without due process (52). None of these challenges has prevailed; often well-accepted legal principles foreclosed many of these attacks (53).

Six constitutional grounds are discussed in this article: (*1*) self-incrimination, (*2*) ex post facto, (*3*) equal protection, (*4*) due process, (*5*) cruel and unusual punishment, and (*6*) unreasonable search and seizure. In addition, states may provide greater protection under state constitutions or statutes than the U.S. Supreme Court has recognized under the federal constitution; independent state grounds have been raised, but no challenge has yet prevailed on this basis (54–57).

SELF-INCRIMINATION CLAUSE

Challenges to the collection of blood or saliva grounded in the Self-incrimination Clause of the Fifth Amendment have been quickly dismissed based on well-established precedent. The leading case is *Schmerber v. California* (58). While being treated at a hospital for injuries sustained in an automobile collision, Schmerber was arrested for driving under the influence of alcohol. At the direction of the investigating police officer, a physician obtained a blood sample from Schmerber. Although the defendant objected to this procedure on the advice of counsel, his blood was extracted and analyzed for alcoholic content. Before the Supreme Court, Schmerber argued that the extraction of blood violated the privilege against self-incrimination. Rejecting this argument, the Court held that the privilege covers only *communicative or testimonial evidence*, not *physical or real evidence*. According to the Court:

> It is clear that the protection of the privilege reaches an accused's communications, whatever form they might take.... On the other hand, both federal and state courts have usually held that it offers no protection against compulsion to submit to fingerprinting, photographing, or measurements, to write or speak for identification, to appear in court, to stand, to assume a stance, to walk, or to make a particular gesture. The distinction which has emerged, often expressed in different ways, is that the privilege is a bar against compelling "communications" or "testimony," but that compulsion which makes a suspect or accused the source of "real or physical evidence" does not violate it (58).

Subsequent Supreme Court cases reaffirmed the testimonial-physical evidence distinction recognized in *Schmerber*. In *United States v. Wade* (59), the Court held that compelling an accused to exhibit his person for observation was compulsion "to exhibit his physical characteristics, not compulsion to disclose any knowledge he might have" (59) and thus not proscribed by the privilege. In *Gilbert v. California* (60), the Court concluded that the compelled production of a "mere handwriting exemplar, in contrast to the content of what is written, like the voice or body itself, is an identifying physical characteristic outside [the Fifth Amendment's] protection" (60). Similarly, in *United States v. Dionisio* (61), the Court ruled that compelling a defendant to speak for the purpose of voice analysis did not violate the Fifth Amendment because the "voice recordings were to be used solely to measure the physical properties of the witnesses' voices, not for the testimonial or communicative content of what was to be said" (61,62). Cheek swabbing falls into the same category.

Courts addressing the Fifth Amendment arguments in the databank context have applied these precedents when rejecting such arguments (63–65).

EX POST FACTO CLAUSE

The United States Constitution prohibits the retroactive application of criminal laws. Article I provides that neither Congress nor any state shall pass an "ex post facto Law" (66). According to the Supreme Court, this prohibition means that "[l]egislatures may not retroactively alter the definition of crimes or increase the punishment of criminal acts" (67). The ex post facto argument is limited to convicts who were already

incarcerated at the time the databank legislation took effect; prospective application does raise this issue (68). Nevertheless, the efficacy of the databanking program would be severely undercut if only the profiles of persons convicted of sex offenses in the future were in the databank; many of these new defendants would not be released for years, while previously convicted inmates would be released into the community without inclusion in the databank.

Some courts have ruled that the ex post facto prohibition does not apply because databanking statutes are not *penal* in nature (69–71). For example, the Ninth Circuit rejected such a challenge to the Oregon statute because its "obvious purpose is to create a DNA data bank to assist in the identification, arrest, and prosecution of criminals, not to punish convicted murderers and sexual offenders" (72).

The ex post facto issue, however, does not necessarily disappear merely because a statute is labeled "non-penal." Ex post facto principles apply when punishment is retroactively increased (73), and that may occur if a sanction for refusal to provide a DNA sample is the denial of parole or the forfeiture of good time credits (credits awarded for a period of good behavior in prison). Much depends on how a parole or good time statute is written. Of course, many states have eliminated both parole and good-time credit.

Parole Release

If parole is purely discretionary, a parole board may consider a refusal to comply with a valid prison regulation, such as one requiring a DNA sample, in determining the appropriateness of parole. In contrast, an increase in the length of a sentence caused by new conditions in a mandatory parole jurisdiction is suspect. For example, the Virginia parole statute mandated parole six months before the sentence release date, and the Fourth Circuit ruled that withholding release for failure to provide DNA samples would be unconstitutional (74). This does not necessarily mean that these inmates can escape providing a sample; a state may make it a new crime to refuse to provide a sample (75).

Good-Time Credit

Reduction of good-time credit raises somewhat different issues. In *Weaver v. Graham* (76), the Supreme Court ruled that the elimination of good-time credit constituted an increase in punishment because "a prisoner's eligibility for reduced imprisonment is a significant factor entering into both the defendant's decision to plea bargain and the judge's calculation of the sentence to be imposed." *Weaver*, however, involved inmates whose good-time credit was legislatively reduced across the board, even if they had *not* violated any prison regulation. Several courts have distinguished databank statutes on this basis, finding that at the time of sentencing good-time credits were known to be contingent on compliance with legitimate prison regulations and the nature of those regulations may be amended while the prisoner is serving penitentiary time (77–79).

EQUAL PROTECTION CLAUSE

The Fourteenth Amendment establishes that no state may "deny any person within its jurisdiction the equal protection of the laws." Several inmates have asserted equal protection grounds as a basis for striking down databank statutes. They claim, for example, that sex offenders are treated differently from other offenders in violation of the equal protection mandate.

The Supreme Court has developed a multi-tiered classification for reviewing equal protection claims. A state statute is subjected to "strict" scrutiny if it adversely affects a suspect class. Utilizing strict scrutiny analysis, a court will require the state to prove that it has a *compelling* governmental interest and that it is employing the *least restrictive means* to achieve its compelling goal. Suspect classifications that warrant strict scrutiny under the Equal Protection Clause are race, alienage, and national origin (80).

A statute can be challenged under the Equal Protection Clause even if it does not adversely affect a suspect classification. Thus, a databank statute that requires DNA sampling only of sex offenders and violent felons may be attacked on the ground that it treats those particular criminals unequally in violation of the equal protection requirement. If no suspect classification is involved, however, courts use a lower level of scrutiny, namely, what is known as the "rational basis" test.

The rational basis test is derived from a long line of Supreme Court decisions (81). Under this type of judicial review, a "statute is presumed to be valid and will be sustained if the classification is rationally related to a legitimate state interest" (82). In *Boling v. Romer* (83), the Tenth Circuit rejected the argument that taking DNA samples only from sex offenders violated the Equal Protection Clause. The court held that there was a "rational relationship" between the "government's decision to classify inmates as convicted sex offenders and the government's stated objective to investigate and prosecute unsolved and future sex crimes" (83, p. 1341).

In *State v. Olivas* (84), the Washington Supreme Court considered a challenge to the state statute that required a DNA sample from anyone convicted of a sexual or violent offense. The court held that "[t]here is a rational relationship between the interest of the government in law enforcement and the application of the statute to this class of persons" (84, p. 1087). The statute's purpose of facilitating the investigation and prosecution of sex offenses and violent crimes was sufficiently important to defeat the equal protection challenge.

DUE PROCESS

Both the Fifth and the Fourteenth Amendments forbid the denial of life, liberty, or property "without due process of law." Inmates have asserted two different due process arguments: substantive due process and procedural due process.

Substantive Due Process

The Supreme Court has stated that "[d]ue process of law is a summarized constitutional guarantee of respect for those personal immunities which ... are 'so rooted in the traditions and conscience of our people as to be ranked as fundamental' ... or are 'implicit in the concept of ordered liberty'" (85). Recognized fundamental rights include the right to bodily integrity (86), which is arguably violated when the state conducts a medical procedure over an individual's objection. State action that infringes a fundamental right protected by the Constitution is subject to the strict scrutiny test (87).

In the 1952 case of *Rochin v. California* (85), the Supreme Court held that the forcible stomach pumping of a suspect to recover narcotic pills "shock[ed] the conscience" and did not comport with traditional ideas of fair play and decency, thereby violating due process. By contrast, the Court, faced with a due process challenge in the 1957 case of *Breithaupt v. Abram* (88), upheld the involuntary extraction of blood from an unconscious suspect after an automobile accident in order to determine whether he was intoxicated. In distinguishing *Rochin*, the Court emphasized that unlike the extraction of stomach contents, the extraction of blood was performed "under the protective eye of a physician" and was a routine and scientifically accurate method that did not involve the "brutality" and "offensiveness" present in *Rochin* (88, pp. 435–437).

The *Rochin* and *Breithaupt* decisions predated the applicability of the Fourth Amendment to the states through the Due Process Clause of the Fourteenth Amendment in 1961 (89,90), and thus the continued validity of an independent substantive due process analysis in these cases is questionable. Challenges to databank statutes no longer need be addressed in terms of due process, but rather as possible violations of specific constitutional guarantees enumerated in the Bill of Rights, such as the right to be free of unreasonable searches and seizures (91). The Supreme Court specifically held in *Schmerber v. California* (92) that the manner in which evidence is obtained from a suspect is subject to the reasonableness requirement of the Fourth Amendment (92, p. 771), and in *Winston v. Lee* (93), the Court applied the Fourth Amendment to the surgical removal of a bullet from a suspect. Thus, virtually all DNA databank and other cases that are potentially subject to attack on substantive due process grounds are better analyzed under the Fourth Amendment (94–96).

Procedural Due Process

Procedural due process mandates that a person cannot be deprived of "life, liberty, or property" without a hearing and attendant procedural safeguards, although the nature of the safeguards differs depending on the interest involved (97). Some inmates have challenged DNA databank statutes on the ground that the taking of a DNA sample without a hearing deprives them of a liberty or a property interest in their genetic material without due process of law. These challenges have uniformly failed.

In *Rise v. Oregon* (98), the plaintiffs argued that the Due Process Clause required prison officials to provide an opportunity for a hearing before requiring felons to submit a blood sample in accordance with Oregon's databank statute. The court held that "[t]he extraction of blood from an individual in a simple, medically acceptable manner, despite the individual's lack of an opportunity to object to the procedure, does not implicate the Due Process Clause" (98, pp. 1562–1563; 99). Consequently, the felons did not have a liberty or property interest at stake.

Similarly, in *Boling v. Romer* (100), the plaintiff challenged a Colorado statute that required inmates convicted of sexual assault offenses to submit a DNA sample as a condition of release on parole. Without providing the sample, inmates could not regain their liberty. The court nevertheless found that plaintiff's argument that the state "unconstitutionally deprived him of a property interest in his blood without due process" was "unpersuasive" (100, p. 1340). The court explained that parole in Colorado was discretionary and that convicts have no constitutional right to be conditionally released before the expiration of their valid sentences.

CRUEL AND UNUSUAL PUNISHMENT

Several challenges to DNA databanks focused on the Eighth Amendment, which proscribes cruel and unusual punishment (101). In *Sanders v. Coman* (102), inmates argued that the use of force to obtain blood samples violated the amendment; they alleged: "The uses of force have included instances of several officers surrounding an inmate while one held his arm still, the spraying of mace, and bending inmates' wrists in a painful manner to induce compliance." An Eighth Amendment violation, however, occurs only if force is applied for the purpose of causing harm (103), or if the force is excessive (104). Neither theory, in the district court's view, applied in this context. Here, force was used to compel compliance with a valid prison regulation (105,106).

Courts have also held that placement in solitary confinement for failing to comply with an order to provide a blood sample does not violate the Clause (107). In *Boling v. Romer* (108), the plaintiff argued that DNA sampling in a prison constituted cruel and unusual punishment because it exposed him to potential abuse from fellow inmates. He claimed that when prison authorities indicated in front of other prisoners that he was required to submit to a DNA test, they identified him as a sex offender and thus made him vulnerable to possible physical harm from inmates who were apparently particularly hostile toward sex offenders (108, p. 1341). The Tenth Circuit rejected this argument, finding it insufficient to support an Eighth Amendment claim.

Another prisoner asserted that DNA testing violated the Eighth Amendment because the blood test itself was painful (109). Not surprisingly, the district court found that the argument lacked merit, noting that the blood was withdrawn by a trained technician in accordance with medically accepted procedures.

The Cruel and Unusual Punishment Clause also prohibits deliberate indifference to an inmate's serious medical needs (110). Inmates asserting that they were

injured because of the blood test, however, must show more than mere negligence in withdrawing blood (111).

Consequently, the Eighth Amendment does not present an obstacle to databanking. Moreover, DNA samples need not be blood. Profiles can be created from cheek swabs, which inflict no pain and are extremely unlikely to cause injury.

SEARCH AND SEIZURE

The most significant legal challenge to databanks is based on the Fourth Amendment's prohibition of unreasonable searches and seizures. Although the U.S. Supreme Court has yet to address the issue, its decisions in other areas provide a framework for analysis.

The Fourth Amendment is intended to ensure "privacy, dignity, and security of persons against certain arbitrary and invasive acts by officers of the Government or those acting at their direction" (112). There are three distinct Fourth Amendment issues raised in this context. First, there is a "seizure" of the person, which brings that person under the control of the government agents. Second, there is a subsequent search for and seizure of a biological sample or trace evidence from this person (113). Third, the use to which the genetic information in the sample is put raises a final Fourth Amendment issue.

A finding that the Fourth Amendment applies does not mean that a procedure is unconstitutional. That is merely the first step in the analysis. As the Supreme Court has often remarked: "[T]he Fourth Amendment does not proscribe all searches and seizures, but only those that are unreasonable" (114).

Applicability of Fourth Amendment

Seizure of the Person. In the databanking context the first issue — seizure of the person — is not problematic because convicts are already incarcerated. The seizure would be an issue for parolees, probationers, or previously released convicts. Nevertheless, notifying such persons to report and provide DNA samples would be a reasonable seizure. Indeed, it is probably not a "seizure" within the meaning of the Fourth Amendment (113). As for arrestees, probable cause is required, but an arrest warrant is not mandated if the arrest takes place in a public place (115).

Search to Obtain Samples. The leading case on defining which governmental activities are "searches" within the meaning of the Fourth Amendment is *Katz v. United States* (116). *Katz* substituted a privacy approach for the traditional property approach to this issue. According to the Supreme Court: "[T]he Fourth Amendment protects people, not places. What a person knowingly exposes to the public, even in his own home or office, is not a subject of Fourth Amendment protection. . . .But what he seeks to preserve as private, even in an area accessible to the public, may be constitutionally protected" (116, p. 351).

There is little dispute that taking blood samples is a search. In *Schmerber* the Supreme Court held that the extraction of blood for the purpose of scientific (blood/alcohol) analysis "plainly constitutes searches of the 'persons'" within the meaning of the Fourth Amendment. In *Skinner v. Railway Labor Executives' Ass'n* (117), which involved a drug testing program, the Court wrote that "it is obvious that this physical intrusion, penetrating beneath the skin, infringes an expectation of privacy that society is prepared to recognize as reasonable" (117, p. 616). In addition to blood samples, lower courts have generally treated the taking of hair (118–120) and saliva (121) samples as searches.

In contrast, the taking of fingerprints (122), voice exemplars (113), or handwriting samples (123) do not constitute searches because such physical characteristics are constantly exposed to the public. (Note the difference between fingerprints and blood or cheek swabbings; it will be important in discussing arrestees later in this article.)

Use of Genetic Information. In *Skinner* the Supreme Court also ruled that the subsequent chemical analysis of the blood sample to obtain physiological data "is a further invasion" of privacy interests — informational privacy (124). This point was further refined when the Court considered the collection of urine samples. Even though this procedure did not involve a bodily intrusion, the Court held that it was a search. Like blood, the chemical analysis of urine can "reveal a host of private medical facts," including whether a person is epileptic, pregnant, or diabetic (124, p. 617).

The courts addressing the constitutionality of databank statutes have acknowledged the applicability of the Fourth Amendment to the taking of a sample (125–127) as well as its subsequent analysis (127). Consequently, the databanking litigation has focused on the second step in Fourth Amendment analysis — the reasonableness of these programs.

Reasonableness of Search

As noted above, the Fourth Amendment does not prohibit all searches, only unreasonable ones. Traditionally, reasonable searches are those conducted pursuant to a warrant issued by a neutral and detached magistrate and based on probable cause. Moreover, search warrants must describe the place to be searched and the items to be seized with "particularity." The particularity requirement circumscribes the police's discretion in executing a search warrant. Nevertheless, exceptions to these traditional requirements have been recognized, and courts have cited several in upholding DNA databank statutes.

The databank cases can be grouped around three lines of Supreme Court precedents: (*1*) administrative searches, (*2*) "special needs" searches, and (*3*) prisoner searches (127). These categories, however, are not mutually exclusive — and they all involve a balancing of interests in determining the reasonableness of the procedure. The next sections focus on sex offenders, the most common category in DNA databank statutes. Later sections discuss persons convicted of other crimes and arrestees.

Administrative Searches. Originally, the phrase "administrative search" was used to describe non-law enforcement searches. For example, the landmark case, *Camara v. Municipal Court* (128), involved housing inspections. The

purpose of these inspections was not to gather evidence of criminal conduct but rather to ensure compliance with health and safety standards. Housing inspectors rather than police officers conducted these searches, although violation of the regulations could result in criminal prosecution.

In *Camara*, the Court held that the reasonableness of an administrative search is determined by balancing the governmental interest against the nature and extent of the intrusion on privacy.

> The ... argument is in effect an assertion that the area inspection is an unreasonable search. Unfortunately, there can be no ready test for determining reasonableness other than by balancing the need to search against the invasion which the search entails. But we think that a number of persuasive factors combine to support the reasonableness of area code-enforcement inspections. First, such programs have a long history of judicial and public acceptance. Second, the public interest demands that all dangerous conditions be prevented or abated, yet it is doubtful that any other canvassing technique would achieve acceptable results. Many such conditions — faulty wiring is an obvious example — are not observable from outside the building and indeed may not be apparent to the inexpert occupant himself. Finally, because inspections are neither personal in nature nor aimed at the discovery of evidence of crime, they involve a relatively limited invasion of the urban citizen's privacy (128, pp. 536–537).

The Court found the inspection system "of indispensable importance to the maintenance of community health" (128, p. 537). Thus, in *Camara*, the Court concluded that housing inspection programs were supported by the compelling government interest of avoiding dangerous living conditions and maintaining housing stock and that the inspection programs were a reasonable means for achieving these societal interests.

Later cases involved the inspection of gun dealerships (129), mines (130), and the workplace pursuant to the Occupational Safety and Health Act (OSHA) (131). Perhaps the most familiar administrative search is the metal detector procedures at airports (132).

New York v. Burger (133), decided in 1987, is a transitional case. It involved a New York statute authorizing warrantless administrative searches of automobile junkyards, which the Supreme Court upheld. The *key* point is that the statute was aimed specifically at finding evidence of *crime*. In contrast, prior administrative searches had focused on governmental interests such as health and safety. Moreover, the junkyard inspections were conducted by the *police*. In a later case the Court employed the balancing test to uphold sobriety roadblock checkpoints (134).

While the balancing approach provides flexibility in achieving significant government objectives, such as airline passenger safety, the danger exists that this approach will result in the "balancing" away of constitutional rights. Therefore, this analysis demands rigor. For example, while the Supreme Court upheld the drug testing of railroad employees after an accident and custom's officers involved in drug interdiction operations, it has struck down the drug testing of political candidates as mandated by a Georgia statute (135). The Court found that the justification for the latter procedure was simply not compelling. In one case, which involved jail searches, the Supreme Court explained the "balancing" analysis as follows:

> The test of reasonableness under the Fourth Amendment is not capable of precise definition or mechanical application. In each case it requires a balancing of the need for the particular search against the invasion of personal rights that the search entails. Courts must consider the scope of the particular intrusion, the manner in which it is conducted, the justification for initiating it, and the place in which it is conducted (136).

Roe v. Marcotte (137), a Second Circuit case decided in 1999, can be used to illustrate this approach. In this case the court reviewed the Connecticut databank statute, which is limited to sex offenders. First, the court correctly found the government interest — solving past and future violent sex crimes — both legitimate and significant. Moreover, the databank system "may" deter future crimes by those whose profile is in the system. Second, the means selected to accomplish these objectives were reasonable. The state cited studies showing a high rate of recidivism for sexual offenders and DNA evidence is "particularly useful" in investigating these crimes "because of the nature of the evidence left at the scenes of these crimes and the demonstrated reliability of DNA testing" (138). Third, the blanket testing of all sex offenders eliminated the need for discretionary decisions, an historical concern in Fourth Amendment jurisprudence. Fourth, the intrusion — the extraction of blood — was slight ["minimal" in the Supreme Court's view (124)] and did not raise a health risk. In these circumstances, the court held that the balance tipped in favor of the databanking statute.

Three other aspects of the Connecticut scheme are noteworthy. First, trained medical personnel are required to take the blood sample. Second, the identifying information associated with the DNA profile remains anonymous until a match is made. Third, procedures limiting access to and dissemination of information in the system are specified.

"Special Needs" Searches. Over time, the rationale underlying administrative searches was extended to other procedures, commonly called "special needs" searches. The Supreme Court in *New Jersey v. T.L.O.* (139) applied this rationale to searches of public school children by teachers; the "special need" was the maintenance of a safe, orderly, and contraband-free school environment in order to create a healthy learning atmosphere. To achieve the desired environment, the Court recognized that "the school setting requires some easing of the restrictions to which searches by public authorities are ordinarily subject" (139, p. 340).

Similarly, in *Griffin v. Wisconsin* (140), the Supreme Court upheld a Wisconsin regulation that permitted a warrantless search of a probationer's home if there existed "reasonable grounds" to believe that the probationer possessed contraband. The Court observed that "[a] State's operation of a probation system, like its operation of a school, government office or prison, or its supervision of a regulated industry, likewise presents 'special needs' beyond normal law enforcement that may justify

departures from the usual warrant and probable-cause requirements" and that "in certain circumstances government investigators conducting searches pursuant to a regulatory scheme need not adhere to the usual warrant or probable-cause requirements as long as their searches meet 'reasonable legislative or administrative standards'" (140, pp. 873–874).

Subsequently, the Court applied this rationale in cases involving government-required alcohol and drug testing for railroad employees (124) and customs agents involved in drug interdiction (141). In the school and probationer cases, the special need resulted in a lesser standard (reasonable suspicion instead of probable cause) to justify an invasion of privacy (142), while the drug testing cases upheld regulatory schemes that did not require any quantum of proof.

A number of courts have used the "special needs" rationale to uphold databank statutes (143–145). In contrast, other courts have balked at applying the "special needs" rationale in this context, noting that this category is limited to governmental objectives "beyond normal law enforcement" (146,147). These courts note that DNA databanks are intended *only* for law enforcement purposes. Other courts point out, however, that "special needs" searches, such as probationer searches, are also associated with law enforcement but do not involve the investigation of a specific crime (148).

More important, as noted above, the administrative search and "special needs" categories are not mutually exclusive — indeed, they often overlap. This is because the "special need" beyond normal law enforcement is typically some administrative objective. For example, an inventory search of the personal belongings of arrestees prior to placement in a jail cell is reasonable, whether classified as a "special need" or an administrative search (149). Similarly, this procedure could also be considered a prisoner search, the next category to be considered. The important point is the "balancing" rationale employed in determining reasonableness. There may, however, be a tendency in some opinions to use "special needs" as a talismanic incantation, curtailing further inquiry.

Fourth Amendment Rights of Prisoners. In *Jones v. Murray* (150), the Fourth Circuit adopted a third type of analysis. In upholding the Virginia statute, the Fourth Circuit relied on several Supreme Court decisions that had held that prisoners had reduced expectations of privacy under the Fourth Amendment. In *Bell v. Wolfish* (151), for example, the Supreme Court upheld the constitutionality of body cavity inspections of pretrial inmates following "contact visits," even in the absence of probable cause. The Court's rationale in determining the reasonableness of the procedure focused on the significant security dangers inherent in this environment: "A detention facility is a unique place fraught with serious security dangers. Smuggling of money, drugs, weapons, and other contraband is all too common an occurrence. And inmate attempts to secrete these items into the facility by concealing them in body cavities are documented in this record" (151, pp. 558–560).

In a later case, *Hudson v. Palmer* (152), the Supreme Court upheld cell searches ("shakedown" inspections) for the purpose of discovering contraband in a prison. The Court, in a 5–4 decision, ruled that a prisoner did not have a reasonable expectation of privacy in a cell. Yet, this holding (like *Wolfish*) was justified on institutional security needs. The Court wrote: "The recognition of privacy rights for prisoners in their individual cells simply cannot be reconciled with the concept of incarceration and the needs and objectives of penal institutions" (152, p. 526).

There are no institutional security needs in the databanking context, and thus this rationale is simply inapplicable. Indeed, some statutes apply even in the absence of incarceration (146). Moreover, both *Wolfish* and *Hudson* acknowledged that the Court's jurisprudence in prisoner rights cases recognizes the applicability of constitutional protections: "There is no iron curtain drawn between the Constitution and the prisons of this country" (153,154). Similarly, in another case the Court recognized that although lawful incarceration "brings about the necessary withdrawal or limitation of many privileges and rights," a retraction had to be "justified by the considerations underlying our penal system" (155).

Finally, these cases are, in the words of one court, "nothing more than special needs cases" (146). In sum, the administrative search rationale provides the best approach; it does not torture the "special needs" rationale, nor misapply the Fourth Amendment prison cases.

Inadequate Remedies. Under any rationale the most troublesome aspect of the databank statutes is the lack of meaningful remedies. As discussed previously, in many states there is little to no deterrent for unauthorized dissemination of DNA profiles and samples. The Virginia statute makes unauthorized dissemination of databank information a misdemeanor, but other statutes do not. By contrast, unauthorized disclosure of information in the federal databank system is punished by a $100,000 fine. While significant criminal penalties should be enacted, criminal prosecution may be insufficient. It often requires proof of intentional conduct, a standard that may be difficult to establish beyond a reasonable doubt. More important, prosecutors have enormous discretion in charging crimes, including the power not to charge at all — a distinct possibility considering the close relationship between prosecutors and the police.

Civil remedies should thus also be included in all databank statutes. Because databank statutes involve constitutional rights, civil rights suits under Section 1983 are possible, although not without impediments. Appropriate remedial models are not hard to find, however. For example, the federal eavesdropping and wiretap act provides for civil damages and injunctive relief in addition to felony sanctions (156). For some violations a plaintiff may recover actual damages or statutory damages of $100 a day or $10,000, whichever is greater. In addition, punitive damages are permitted as well as reasonable attorney's fees and other litigation costs (157). The Privacy Protection Act of 1980 provides for civil (actual) damages but not less than liquidated damages of $1000, reasonable attorney's fees, and other litigation costs (158).

Comparable provisions should be added to the databanking statutes. Proponents of databanking could not

object to such provisions because they assert that violations will be few.

Expansion of Coverage Beyond Sex Offenders

Most databank statutes are limited to sex offenders. These provisions are supported by empirical research on recidivism; more than 50 percent of brutal and violent crimes, e.g., rape and murder, are carried out by repeat offenders (3). Recidivism is noted by several courts in upholding databank schemes (159). The nature of these offenses — their brutality and their often serial nature (3, p. 48) — is a critical point. However, some statutes also encompass homicides and other crimes of violence. Still others include all felons. The justification for including prisoners who have been convicted of white-collar felonies is difficult to discern. Even the sex offender category is problematic if it includes prostitution and public indecency as some statutes do.

Jones v. Murray (160), the Fourth Circuit case discussed above, addressed this issue because all felons are included in the Virginia DNA databank system. To buttress its position, the court cited recidivism studies encompassing all felons (161). The inmates, however, argued that the statistics on nonviolent felons undercut the state's position. The inmates' "statistics indicate[d] that 97% of the cases in which DNA evidence was used to link a defendant with a crime involved murder or rape, and further, less than 1% of all nonviolent offenders are later arrested on murder and rape charges" (162). In response, the majority merely noted that the percentages need not be high where the objective is significant and the privacy intrusion is limited.

The dissent in *Jones* believed that the distinction between violent and nonviolent felons was critical: "The only state interest offered by the Commonwealth for including non-violent felons is administrative ease," but such an interest does not suffice "to outweigh a prisoner's expectation of privacy in not having blood withdrawn from his body when that prisoner is not significantly more likely to commit a violent crime in the future than a member of the general population" (162, pp. 313–314). Indeed, the Virginia senate report concluded that the recidivism data *only* "supported the inclusion of plaintiffs convicted for felony sex offenses, assault, capital murder, first and second degree murder, voluntary manslaughter, larceny and burglary" (162, p. 314). All felons were added to make the databank "more efficient and cost effective." The dissent also pointed to other statistics in the record: "United States Justice Department statistics provided in the record show that only %0.4 of non-violent felons are later arrested on rape charges, and only %0.8 are later arrested on murder charges. One might assume non-violent drug offenders would be more likely to commit violent crime subsequent to release than other non-violent felons; yet, only %0.4 of them are later arrested for rape, and %0.3 for murder." The dissenting judge concluded: The lack of justification "leads me to a deep, disturbing, and overriding concern that, without a proper and compelling justification, the Commonwealth may be successful in taking significant strides toward the establishment of a future police state, in which broad and vague concerns for administrative efficiency will serve to support substantial intrusions into the privacy of citizens" (162, p. 314).

The British experience, which commenced earlier than that of the United States, may be instructive. The British initially focused on sex offenses but later included burglaries and car theft because of the high number of matches. They found cross-over among offenses. According to one official, "People who commit serious crime very often have convictions for petty crime in their history" (163). While the cross-over concept is significant, the scope of the British system is breathtaking; they expect to "eventually include a third of all English men between 16 and 30, the principal ages for committing crimes."

The category of crimes subject to databanking should be supported by empirical data or persuasive reasons. There is apparently some support for including some nonviolent crimes (23), such as burglary, but each offense should be specified. For example, historically, burglary was not considered a "property" crime; it was a crime against habitation, intended to protect people in their dwellings (164). The crime that is the objective of the burglary need not be larceny or theft; it could be any felony including murder or rape. Burglars must anticipate what action they will take if surprised by an occupant, including the use of force. Therefore an argument to include burglary could be made, but felony tax evasion would be a different issue.

Expansion of Coverage to Arrestees

The Louisiana statute applies to sex offender and other specified *arrestees*. New York Police Commissioner Howard Safir has proposed that DNA be collected from *all* arrestees (23). Not satisfied with that proposal, New York Mayor Rudy Giuliani suggested that all newborns be tested (39). These proposals raise significant legal and policy problems.

Unlike a conviction, which is either based on a jury verdict and the "beyond a reasonable doubt" standard or a guilty plea with its attendant constitutional safeguards (including the right to counsel), *one* police officer can make an arrest based on his or her own view of probable cause, which is not a high standard. If the arrest occurs in a public place, an arrest warrant is not required (165). There is a requirement for judicial review of the probable cause determination within 48 hours (166,167), but the DNA sample will have been taken by then. In any event, this judicial review occurs in an ex parte procedure — that is, without the presence of the arrestee or defense counsel. The difference between an arrest and conviction is immense. In 1994, 65.3 percent of murder arrests resulted in conviction, but the conviction rates for some other crimes were much lower: robbery (39.3 percent), aggravated assault (14.1 percent), and burglary (38.8 percent) (168). Moreover, the FBI has reported that one-third of the initial rape suspects identified by the police are exonerated by DNA profiling (169); this statistic further underscores the difference between arrest and conviction. In short, the expansion of databank coverage to arrestees raises significant constitutional issues. Several theories that may

be used to justify such expansion are discussed next and rejected.

Search Incident to Arrest. One possible theory for obtaining blood samples or cheek swabs would be a search incident to apprehension, a well-established exception to the warrant requirement (170,171). Under this exception, once a suspect has been arrested based on probable cause, a search of the arrestee's person and the area within her immediate control is permitted. In *Chimel v. California* (172), the Supreme Court set forth a twofold justification for this exception: (*1*) protection of the arresting officer and (*2*) prevention of the destruction of evidence. The search is automatic once there is an arrest; no additional showing is required.

The Supreme Court, however, has shown a greater concern about searches involving bodily intrusions than about other types of incident searches. In *Schmerber v. California* (173), the Court considered the constitutionality of extracting blood for the purpose of blood-alcohol analysis. The Court rejected the notion that the extraction of blood would automatically be encompassed by the search incident to arrest doctrine. According to the Court, the justifications underlying the search incident to arrest rule

> have little applicability with respect to searches involving intrusions beyond the body's surface. The interests in human dignity and privacy which the Fourth Amendment protects forbid any such intrusions on the mere chance that desired evidence might be obtained. In the absence of a *clear indication* that in fact such evidence will be found, these fundamental human interests require law officers to suffer the risk that such evidence may disappear unless there is an immediate search (173, pp. 769–770).

The Court further considered the necessity of securing a warrant based on probable cause as a prerequisite to the extraction of blood. It found the purpose underlying the warrant requirement — the intervention of a neutral detached magistrate between the police and the citizen — applicable to bodily intrusions: "The importance of the informed, detached and deliberate determinations of the issue whether or not to invade another's body in search of evidence of guilt is indisputable and great" (174). Nevertheless, because the alcohol content of blood diminishes with the passage of time, the Court recognized an "emergency" exception to the warrant requirement, which was necessary to preclude the destruction of evidence. This emergency exception, however, does not apply in other contexts — for example, when blood is sought for the purpose of genetic testing, including DNA profiling, a physical characteristic that remains constant (175,176).

Search for Blood. There may be grounds to take a DNA sample for testing *outside* the databank context. For example, if the crime that is the basis for an arrest involves blood, semen, or other evidence, there may be probable cause to issue a search warrant in that *specific case*. Often the probable cause to arrest also provides probable cause to search the arrestee for a DNA profile — for example, to compare the suspect's DNA with that from a semen stain in a rape case. A search warrant, as distinguished from an arrest warrant, requires probable cause that (*1*) the subject committed a crime and (*2*) blood analysis results would be evidence of that crime (177). In other words, the search is not automatic upon arrest, and this type of search differs from an administrative search (databanking), which is based on future or past crimes. Once a sample is obtained for this purpose, a search of a databank (for past offenses) would involve only a slight incremental privacy invasion (178).

Identification Rationale. Another rationale that has been suggested is an "identification" exception, applicable at a stationhouse "booking" after arrest. *Jones v. Murray* (179) alluded to this rationale: "[W]hen a suspect is arrested upon probable cause, his identification becomes a matter of legitimate state interest and he can hardly claim privacy in it. . . .[T]he identification of suspects is relevant not only to solving the crime for which the suspect is arrested, but also for maintaining a permanent record to solve other past and future crimes" (179, p. 306). While *Jones* did not involve an arrestee, it went on to cite an analogy to fingerprinting — as have other databank cases (180,181).

Unquestionably, the proper identification of a person arrested is a legitimate governmental objective (182); it is not unusual for fugitives to use an alias (183,184). Moreover, fingerprinting arrestees is a reasonable method to accomplish this goal, as a number of courts (but not yet the U.S. Supreme Court) have recognized (185–187). Prior to the time that fingerprinting became routine, photographing (188) and the Bertillion system (based on physical measurements) were used for this purpose (189).

Nevertheless, the "identification" rationale is problematic for several reasons. First, there are significant differences between fingerprinting and DNA sampling. The former does not involve the kind of privacy issues raised by DNA samples (as opposed to profiles). As the Supreme Court has noted, "[f]ingerprinting involves none of the probing into an individual's private life and thought that marks an interrogation or search" (190). The fingerprint system is less intrusive and not as subject to abuse as some present methods of DNA sample collection. Second, the availability of fingerprinting undercuts the need to use DNA for the purpose of identification; every person whose DNA profile is in the database has fingerprints in AFIS. There are 226 million fingerprint cards in the FBI's Criminal Justice Information Services Division.

Third, fingerprinting arrestees as a means of identification developed before computers automated the process in AFIS. Prior to AFIS, the FBI used the Henry classification system, which was based on friction ridge patterns (e.g., arches, loops, whorls, ridge counting) and required prints from all *ten* fingers to identify arrestees. This took weeks if not months. In contrast to the classification of fingerprints, the identification of a partial crime scene print was based on ridge detail (e.g., ridge endings, bifurcations, enclosures) and required *suspects* because prints of all ten fingers are rarely left at a crime scene (191). In short, a single crime scene print could not be matched to the FBI's central depository. For example, a serial rapist, known as

the "Westside Rapist" terrorized Cleveland in the 1980s. He was eventually convicted of raping 29 women (192). The police had partial fingerprints from several crime scenes, and the perpetrator had a record of prior convictions, but there was no way at that time to connect the prints without a suspect or suspects. Stated another way, there were no "cold hits" before AFIS. Accordingly, when the identification "exception" was judicially recognized, stationhouse fingerprinting did not solve past or future crimes. Consequently, the fingerprint "precedent" cannot be cited without further analysis.

Other Uses: Medical and Administrative Purposes

Some databank statutes do not limit use of DNA profiles or stored samples to criminal identification. For example, one statute authorizes databank samples to be used in medical research, even though informed consent has not been obtained from any of the subjects (48, p. 1491). Somewhat similarly, the Massachusetts statute permits disclosure for "advancing other humanitarian purposes" (193). Although many research uses of criminal databanks may ultimately prove to be ethically and legally acceptable, added safeguards such as mandatory approval by medical review boards are essential (17,19,44,47). The greater possibility that individual genetic information will be identified or released in the course of research also underscores the need for enhanced security for DNA samples and profiles described above, particularly civil remedies.

Military and Medical Records

In the future, one can expect criminal enforcement officials to turn to other repositories of DNA, such as military and medical databanks, to search for matches with crime scene profiles (47,195). Commentators have only begun to analyze the privacy issues raised by government use of DNA for law enforcement when the original biological samples were collected by other entities for entirely different purposes. While such a dramatic expansion of criminal DNA databanks seems susceptible to the legal challenges outlined above, there is currently neither judicial guidance nor academic consensus on whether this type of dramatic expansion to the criminal enforcement artillery will survive constitutional challenge (17,19,21,45,163). Additional legislative prohibitions or limitations seem inevitable if DNA databanks expand in this way (195).

CONCLUSION

DNA databanking offers a powerful tool for crime solution, especially in violent crime cases such as rape. Unfortunately, possible infringements of essential rights to privacy may also be made possible by the collection and centralization of individuals' genetic information. While the courts have begun to consider the constitutionality of DNA databank programs, other privacy concerns presented by the databank statues are likely to evade judicial scrutiny. There is thus need for more anticipatory legislation to ensure that the best of DNA databanks is harnessed and used for the good, without eroding the privacy rights that remain for those whose genetic material is stored in state DNA databanks. As bioethicist Eric Juengist has remarked, "It is ... individuals' 'informational privacy' that is at stake in the prospect of widespread [state DNA databank laws], and it is in those terms that the policy challenge ... should be framed. What should society be allowed to learn about its citizens in the course of attempting to identify them?" (44, p. 63). While important steps have been taken over the past few years to improve state databank laws, work remains to be done.

BIBLIOGRAPHY

1. H. Levy, *And the Blood Cried Out: A Prosecutor's Spellbinding Account of the Power of DNA*, Basic Books, New York, 1996, p. 128.
2. *Jones v. Murray*, 962 F.2d 302, 304 (4th Cir.), *cert. denied*, 506 U.S. 977 (1992).
3. *Popular Science* **48**, 49 (1999).
4. 533 So. 2d 841 (Fla. Dist. Ct. App. 1988), *rev. denied*, 542 So. 2d 1332 (Fla. 1989).
5. *Office of Technology Assessment, U.S. Congress* 99, *Genetic Witness: Forensic Uses of DNA Tests*, U.S. Government Printing Office, Washington, DC, 1990.
6. P.C. Giannelli and E.J. Imwinkelried, *Scientific Evidence*, 3rd ed., Lexis Law Publishing, Charlottesville, VA, 1999.
7. P.C. Giannelli, *J. Crim. L. Criminol.* **88**, 380 (1997).
8. *State v. Council*, 515 S.E.2d 508, 518 (S.C. 1999).
9. G. Sensabaugh and D.H. Kaye, *Jurimetrics J.* **39**, 2–3 (1998).
10. *People v. Wesley*, 533 N.Y.S.2d 643, 644 (Co. Ct. 1988), *aff'd*, 633 N.E.2d 451 (N.Y. 1994).
11. R.C. Lewontin and D.L. Hartl, *Science* **254**, 1745–1746 (1991).
12. *United States v. Salameh*, 152 F.3d 88 (2d Cir. 1998); A. Blum, *Nat. Law J.* **8** (October 25, 1993).
13. E. Connors et al., *Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence after Trial*, National Institute of Justice, U.S. Department of Justice, Washington, DC, 1996.
14. G.W. O'Reilly, *Judicature* **81**, 114 (1997).
15. J.W. Diehl, *Jurimetrics J.* **39**, 431 (1999).
16. B. Dedman, *N.Y. Times*, October 7, 1 (1999).
17. G.J. Annas, *J. Am. Med. Assoc.* **270**, 2346 (1993).
18. Va. Code Ann. §19.2–310.2.
19. M. Hibbert, *Wake Forest Law Rev.* **34**, 767, 775 (1999).
20. 42 U.S.C. §1370.
21. J.S. Deck, *Vermont Law Rev.* **20**, 1057, 1065–1067 (1996).
22. R. Hoyle, *Nat. Biotechnol.* **16**, 987 (1998).
23. M. Hansen, *Am. Bar Assoc. J.* **85**, 26 (1999).
24. Colo. Rev. Stat. §17-2-201(5)(g).
25. Wash. Rev. Code §43.43.754; Mo. Stat. §650.055.
26. Ala. Code §36-18-24; N.M. Stat. Ann. §29-16-6; Va. Code Ann. §19.2-310.2; Wyo. Stat. Ann. §7-19-403.
27. La. Rev. Stat. Ann. §15-609.
28. Maricopa County Juvenile Action, 930 P.2d 496 (Ariz. Ct. App. 1996).
29. In re Mitchell, 880 P.2d 958 (Or. Ct. App. 1994), *cert. denied*, 522 U.S. 1004 (1997).
30. *Landry v. Attorney General*, 709 N.E.2d 1085 (Mass. 1999).
31. N.C. Gen. Stat. §§15A-266.1.

32. Code Mass. Regs. 55, §1.03-1.05.
33. *Shelton v. Gudmanson*, 934 F. Supp. 1048, 1050 (W.D. Wisconsin 1996).
34. Mass. Gen. Laws ch. 22E, §15.
35. J.E. McCain, *Am. J. Hum. Genet.* **56**, 1487–1489 (1995).
36. R.W. Schumacher, II, *Fordham Urb. Law J.* **26**, 1635, 1667–1668 (1999).
37. *U.S. News & World Rep.*, October 25, 33 (1999).
38. Y.H. Yee, *Am. J. Crim. Law* **22**, 461, 483 (1995).
39. M. Higgins, *Am. Bar Assoc. J.* **85**, 64, 66 (1999).
40. FBI, *Crime Lab. Dig.* **20**, 51 (1993).
41. National Research Council, *DNA Technology in Forensic Science*, National Academy Press, Washington, DC, 1992, p. 114.
42. R. Murch and B. Budowle, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality*, Yale University Press, New Haven, CT, 1997, pp. 212–231.
43. U.S. General Accounting Office, *National Crime Information Center: Legislation Needed to Deter Misuse of Criminal Justice Information*, GAO/T-GGD-9341, U.S. Government Printing Office, Washington, DC, 1993.
44. E.T. Juengst, *Chicago-Kent Law Rev.* **75**, 64–67 (1999).
45. M.J. Markett, *Suffolk Univ. Law Rev.* **30**, 185, 208–215 (1996).
46. D.L. Burk, *Univ Tol. Law Rev.* **24**, 87, 91–92 (1992).
47. D. Kaye and E. Imwinkelried, *Forensic DNA Typing: Selected Legal Issues*, Report to the Working Group on Legal Issues, National Commission on the Future of DNA Evidence, Washington, DC, 2000, pp. 53–54.
48. J.E. McEwen and P.R. Reilly, *Am. J. Hum. Genet.* **54**, 941, 955–956 (1994).
49. *Ryncarz v. Eikenberry*, 824 F. Supp. 1493, 1502 (E.D. Wash. 1993).
50. *Jones v. Murray*, 962 F.2d 302 (4th Cir. 1992); *State v. Olivas*, 856 P.2d 1076 (Wash. 1993).
51. *Shaffer v. Saffle*, 148 F.3d 1180 (10th Cir. 1998).
52. *Boling v. Romer*, 101 F.3d 1336, 1340 (10th Cir. 1996).
53. *State v. Olivas*, 856 P.2d 1076, 1086–1067 (Wash. 1993).
54. *People v. Calahan*, 649 N.E.2d 588, 592 (Ill. App. 1995).
55. *People v. Wealer*, 636 N.E.2d 1129, 1137 (Ill. App. 1994).
56. State *ex rel. Juv. Dept. v. Orozco*, 878 P.2d 432 (Oregon 1994).
57. *State v. Olivas*, 856 P.2d 1076, 1080 (Wash. 1993).
58. 384 U.S. 757 (1966).
59. 388 U.S. 218 (1967).
60. 388 U.S. 263 (1967).
61. 410 U.S. 1 (1973).
62. *Fisher v. United States*, 425 U.S. 391, 408 (1976).
63. *Shaffer v. Saffle*, 148 F.3d 1180, 1181 (10th Cir.)(Oklahoma statute), *cert. denied*, 119 S.Ct. 520 (1998).
64. *Boling v. Romer*, 101 F.3d 1336, 1340 (10th Cir. 1996)(Colorado statute).
65. *Cooper v. Gammon*, 943 S.W.2d 699, 705 (Mo. App. 1997).
66. U.S. Const. art. I, §9, CL. 3 and §10.
67. *Collins v. Youngblood*, 497 U.S. 37, 42 (1990).
68. *Weaver v. Graham*, 450 U.S. 24, 29 (1981).
69. *Shaffer v. Saffle*, 148 F.3d 1180, 1182 (10th Cir. 1998), *cert. denied*, 119 S.Ct. 520 (1998).
70. *Gilbert v. Peters*, 55 F.3d 237, 239 (7th Cir. 1995).
71. *Jones v. Murray*, 962 F.2d 302, 309 (4th Cir. 1992).
72. *Rise v. Oregon*, 59 F.3d 1556, 1562 (9th Cir. 1995), *cert. denied*, 517 U.S. 1160 (1996).
73. *Calder v. Bull*, 3 U.S. Dall. 386, 390 (1798).
74. *Jones v. Murray*, 962 F.2d 302, 310 (4th Cir. 1992).
75. Mass. Gen. L. ch. 22E, §11.
76. 450 U.S. 24, 32 (1981).
77. *Gilbert v. Peters*, 55 F.3d 237, 239 (7th Cir. 1995).
78. *Ewell v. Murray*, 11 F.3d 482, 486 (4th Cir. 1993), *cert. denied*, 511 U.S. 1111 (1994).
79. *Jones v. Murray*, 962 F.2d 302, 309–310 (4th Cir. 1992).
80. *Cleburne v. Cleburne Living Ctr., Inc.*, 473 U.S. 432, 440 (1985).
81. *Chapman v. United States*, 500 U.S. 453, 465 (1991).
82. *Bankers Life and Cas. Co. v. Crenshaw*, 486 U.S. 71, 81 (1988).
83. 101 F.3d 1336 (10th Cir. 1996) (Colorado statute).
84. 856 P.2d 1076 (Wash. 1993).
85. *Rochin v. California*, 342 U.S. 165, 169 (1951).
86. *Washington v. Glucksberg*, 117 S.Ct. 2258, 2267 (1997).
87. *San Antonio Independent School District v. Rodriguez*, 411 U.S. 1, 17 (1973).
88. 352 U.S. 432 (1957).
89. *Mapp v. Ohio*, 367 U.S. 643 (1961) (exclusionary rule).
90. *Wolf v. Colorado*, 388 U.S. 25 (1949) (core value).
91. *Yanez v. Romero*, 619 F.2d 851, 854 (10th Cir.), *cert. denied*, 449 U.S. 876 (1980).
92. 384 U.S. 757 (1966).
93. 470 U.S. 753 (1985).
94. *County of Sacramento v. Lewis*, 523 U.S. 833, 849 n. 9 (1998).
95. *People v. Bracamonte*, 540 P.2d 624, 631 (Cal. 1975) (en banc).
96. *Yanez v. Romero*, 619 F.2d 851, 853 (10th Cir 1980).
97. Compare *Gagnon v. Scarpelli*, 411 U.S. 778 (1973)(probation revocation), with *Goss v. Lopez*, 415 U.S. 912 (1974) (school disciplinary hearing).
98. 59 F.3d 1556 (9th Cir. 1994).
99. *Cooper v. Gammon*, 943 S.W.2d 699, 706 (Mo. App. 1997).
100. 101 F.3d 1336 (10th Cir. 1997).
101. U.S. Const. amend. VIII.
102. 864 F.Supp. 496, 498 (E.D. N.C. 1994).
103. *Hudson v. McMillian*, 503 U.S. 1 (1992).
104. *Whitley v. Albers*, 475 U.S. 312, 319 (1986).
105. 864 F. Supp. at 500.
106. W.R. LaFave and A. Scott, *Substantive Criminal Law*, §2.14(f), West Publishing, St. Paul, MN, 1986.
107. *Cooper v. Gammon*, 943 S.W.2d 699, 707 (Mo. App. 1997).
108. 101 F.3d 1336 (10th Cir. 1996).
109. *Kruger v. Erickson*, 875 F. Supp. 583, 588 (D. Minn. 1995).
110. *Estelle v. Gamble*, 429 U.S. 97, 104–105 (1976).
111. *Ryncarz v. Eikenberry*, 824 F. Supp. 1493, 1502 (E.D. Wash. 1993).
112. *Skinner v. Ry. Labor Executives' Association*, 489 U.S. 602, 613–614 (1989).
113. *United States v. Dionisio*, 410 U.S. 1, 8 (1973) (grand jury subpoena for voice exemplar not a seizure).
114. *Skinner v. Ry. Labor Executives' Ass'n*, 489 U.S. 602, 619 (1989).
115. *United States v. Watson*, 423 U.S. 411 (1979).
116. 389 U.S. 347 (1967).

117. 489 U.S. 602 (1989).
118. *Bouse v. Bussey*, 573 F.2d 548, 550 (9th Cir. 1977).
119. *United States v. D'Amico*, 408 F.2d 331, 333 (2d Cir. 1969).
120. *State v. Sharpe*, 200 S.E.2d 44, 47–48 (N.C. 1973).
121. *United States v. Nicolosi*, 885 F. Supp. 50, 55 (E.D. N.Y. 1995).
122. *Cupp v. Murphy*, 412 U.S. 291, 294 (1973).
123. *United States v. Mara*, 410 U.S. 19, 21 (1973).
124. Skinner, 489 U.S. at 616.
125. Schlicher (NFN) Peters, I and II, 103 F.3d 940, 942 (10th Cir. 1996).
126. *Landry v. Attorney General*, 709 N.E.2d 1085, 1090 (Mass. 1999).
127. *People v. Wealer*, 636 N.E.2d 1129, 1132 (Ill. App. 1994).
128. 387 U.S. 523, 538 (1967).
129. *United States v. Biswell*, 406 U.S. 311 (1972).
130. *Donovan v. Dewey*, 452 U.S. 594 (1981).
131. *Marshall v. Barlow's Inc.*, 436 U.S. 307 (1978).
132. W.R. LaFave, *Search and Seizure: A Treatise on the Fourth Amendment*, §9.6(c), 3rd ed., West Publishing, St. Paul, MN, 1996.
133. 482 U.S. 691 (1987).
134. *Michigan Dept. of State Police v. Sitz*, 496 U.S. 444, 455 (1990).
135. *Miller v. Chandler*, 502 U.S. 305 (1997).
136. *Bell v. Wolfish*, 441 U.S. 520, 559 (1979).
137. 193 F.3d 72 (2d Cir. 1999).
138. 193 F.3d at 79.
139. 469 U.S. 325 (1985).
140. 483 U.S. 868 (1987).
141. *National Treasury Employees Union v. Von Raab*, 489 U.S. 656, 665–666 (1989).
142. *Terry v. Ohio*, 392 U.S. 1 (1968).
143. *Roe v. Marcotte*, 193 F.3d 72 (2d Cir. 1999).
144. *Shelton v. Gudmanson*, 934 F. Supp. 1048, 1051 (W.D. Wisconsin 1996).
145. *State v. Olivas*, 856 P.2d 1076, 1086 (Wash. 1993).
146. *People v. Wealer*, 636 N.E.2d 1129, 1135 (Ill. App. 1994).
147. *State v. Olivas*, 856 P.2d 1076, 1092 (Wash. 1993).
148. *Shelton v. Gudmanson*, 934 F. Supp. 1048, 1050–1051 (W.D. Wis. 1996)(Wisconsin statute).
149. *Illinois v. LaFayette*, 462 U.S. 640, 643 (1983).
150. 962 F.2d 302 (4th Cir. 1992).
151. 441 U.S. 520 (1979).
152. 468 U.S. 517 (1984).
153. *Wolff v. McDonnell*, 418 U.S. 539, 555–556 (1974).
154. P.C. Giannelli and F.A. Gilligan, *Va. Law Rev.* **62**, 1045 (1976).
155. *Price v. Johnston*, 334 U.S. 266, 285 (1948).
156. 18 U.S.C. §2521.
157. 18 U.S.C. §2520(b)(2) & (3).
158. 42 U.S.C. §2000aa.
159. *Shelton v. Gudmanson*, 934 F. Supp. 1048, 1051 (W.D. Wis. 1996).
160. 962 F.2d 302 (4th Cir. 1992).
161. A.J. Beck and B.E. Shiplely, *Recidivism of Prisoners Released in 1983*, U.S. Dept. of Justice, Office of Justice Programs, Bureau of Justice Statistics, Washington, DC, 1989, p. 1.
162. 962 F.2d at 308.
163. N. Wade, *N.Y. Times*, October 12, A1 (1998).
164. W.R. LaFave and A. Scott, *Substantive Criminal Law*, West Publishing, St. Paul, MN, 1986.
165. *United States v. Watson*, 423 U.S. 411 (1976).
166. *County of Riverside v. McLaughlin*, 500 U.S.44 (1991).
167. *Gerstein v. Pugh*, 420 U.S. 103 (1975).
168. *N.Y. Times, 1998 Almanac* 323 (1997)(citing Bureau of Justice Statistic Bull. 1994).
169. FBI, *The Application of DNA Testing to Solve Violent Crimes* (1993).
170. W. LaFave, *Search and Seizure: A Treatise on the Fourth Amendment*, West Publishing, St. Paul, MN, 1996.
171. E. Imwinkelried et al., *Courtroom Criminal Evidence*, 3rd ed., Michie, Charlottesville, VA, 1998.
172. 395 U.S. 752 (1969).
173. 384 U.S. 757 (1966).
174. 384 U.S. at 770.
175. *Graves v. Beto*, 301 F. Supp. 264, 265 (E.D. Tex. 1969), *aff'd*, 424 F.2d 524 (5th Cir.), *cert. denied*, 400 U.S. 960 (1970).
176. *Mills v. State*, 345 A.2d 127, 132 (Md. App. 1975), *aff'd*, 28 Md. 262, 363 A.2d 491 (1976).
177. In re Lavigne, 641 N.E.2d 1328, 1331 (Mass. 1994).
178. H.J. Krent, *Tex. Law Rev.* **74**, 49, 53 (1995) (for a different view).
179. 962 F.2d 302 (4th Cir. 1992).
180. *Landry v. Attorney General*, 709 N.E.2d 1085, 1092 (Mass. 1999).
181. *Olivas*, 856 P.2d 1076, 1093 (Wash. 1993).
182. *United States v. Laub Baking Co.*, 283 F. Supp. 217, 221 (N.D. Ohio 1968).
183. *United States v. Valencia-Lucena*, 925 F.2d 506, 513 (1st Cir. 1991).
184. *United States v. Boyle*, 675 F.2d 430, 432 (1st Cir. 1982).
185. *Napolitano v. United States*, 340 F.2d 313, 314 (1st Cir. 1965).
186. *Smith v. United States*, 324 F.2d 879, 882 (D.C. Cir. 1963).
187. *United States v. Krapf*, 285 F.2d 647 (3d Cir. 1961).
188. *People v. Sallow*, 165 N.Y.S. 915 (Ct. Gen. Sess. of Peace 1917).
189. *Bartletta v. McFeeley*, 152 A. 17 (N.J. 1930).
190. *Davis v. Mississippi*, 394 U.S. 721, 727 (1969).
191. *United States v. Laub Baking Co.*, 283 F. Supp. 217, 223 (N.D. Ohio 1968).
192. J. Neff, *Unfinished Murder: The Capture of Serial Rapist*, Pocket Books, New York, 1995.
193. Mass. Gen. ch. 22E, §10(d)(4).
194. A. Lippman, *Am. J. Pub. Health* **86**, 1030 (1996).
195. R.C. Scherer, *Geo. Law J.* **85**, 2007, 2016 (1997).

See other entries GENETIC INFORMATION, ETHICS, ETHICAL ISSUES IN TISSUE BANKING AND HUMAN SUBJECT RESEARCH IN STORED TISSUES; GENETIC INFORMATION, ETHICS, PRIVACY AND CONFIDENTIALITY: OVERVIEW; GENETIC INFORMATION, LEGAL, ERISA PREEMPTION, AND HIPAA PROTECTION; GENETIC INFORMATION, LEGAL, FDA REGULATION OF GENETIC TESTING; GENETIC INFORMATION, LEGAL, GENETIC PRIVACY LAWS.

GENETIC INFORMATION, LEGAL, ERISA PREEMPTION, AND HIPAA PROTECTION

MARY ANNE BOBINSKI
University of Houston Law Center
Houston, Texas

OUTLINE

INTRODUCTION

Researchers are constantly discovering new linkages between human traits, diseases, or conditions and human genes. The information obtained through the Human Genome Project is expected to be used for human benefit. But should the use of this information be restricted in any way, and if so, by whom? People might be reluctant to gather useful information about their genetic heritage if there is a chance this information will be used to their detriment by others. "Genetic discrimination" has a long history in the United States and is defined as "discrimination against an individual or a member of an individual's family solely on the basis of that individual's genotype" (1). Fears about genetic discrimination could limit the utility of new genetic discoveries.

A number of states have attempted to protect individuals by enacting legislation restricting the use of genetic information by employers, insurance companies, and others (Tables 1 and 2, Appendix). At the same time the federal government has enacted statutes that affect the necessity for, and utility of, state intervention. One federal statute in particular, the Employee Retirement Income Security Act of 1974 (ERISA), has created difficulties for states that wish to regulate the use of genetic information (2). The federal government recognized the difficulties faced by the states and enacted the Health Insurance Portability and Accountability Act (HIPAA) (3). This Act directly regulates the permissible use of genetic information by employment-based plans and authorizes states to engage in additional regulatory action in this area. This article is devoted to examining the roles of ERISA, HIPAA, and state statutes in regulating the use of genetic information. The discussion concludes that federal legislation has provided an important supplement to state regulation of the use of genetic information.

GENETIC INFORMATION

Although the technology for genetic testing has only recently become accessible, genetic information has been available for a long time. People have long understood the general concept of heredity and have been aware that some conditions, such as hemophilia, tend to be inherited within families. The general concept of hereditable conditions gained a new scientific foundation with Watson and Crick's description of the structure of DNA in the 1950s. Since that time scientists have been able to make an ever-increasing number of connections between the structure of an individual's DNA and that individual's traits or the expression of a variety of genetic conditions. What this means, first of all, is that the problems presented by the availability of genetic information are not new. Some genetic information has been available for use, properly or improperly, and with good or bad intent, for much of this century. Individuals with a family history of a known genetic condition have used this information to make important decisions, such as decisions about whether to reproduce or whether to provide health or life insurance for a member of the family.

States have also used this information to sanction the sterilization of individuals based on their genotype. For example, in 1927 the United States Supreme Court upheld the constitutionality of a Virginia statute that allowed the involuntary surgical sterilization of Carrie Buck, a female resident in a Virginia institution for the "feebleminded." In this now infamous case, *Buck v. Bell*, Justice Holmes concluded that "it is better for all the world, if instead of waiting to execute degenerate offspring for crime, or to let them starve for their imbecility, society can prevent those who are manifestly unfit from continuing their kind. . . . Three generations of imbeciles are enough" (4,5). Although modern courts are not likely to uphold state-mandated sterilization programs, this decision vividly demonstrates past misunderstandings about the nature of the hereditability of various traits as well as the misuse of that information to infringe individual rights.

While the availability of genetic information has a long history, the technology used to obtain this information has undergone rapid change. Advances in genetic research have improved the accuracy of attempts to relate the content of one's genetic code with its real life effects. In addition research has created an exponential growth in the number of conditions with an identified genetic correlate. This has enabled scientists to test for the presence of the particular gene or gene combination. Within the next several decades we are likely to have access to over 100 different tests for genetic variations that predispose persons to common diseases such as cancer, cardiovascular diseases, autoimmune diseases, and so on (6).

Table 1. Selected Recent State Legislative Genetic Information Initiatives

State (year enacted)	Type of Genetic Information	Area of Protection	Type of Protection
Alabama (1997)	Genetic test for predisposition to cancer	Health benefit plans (insured, self-insured)	No genetic testing or use of result for coverage or rates
Alaska (1997)	HIPAA-related provisions; genetic information	Health insurers	No discrimination based on health status; health status included genetic information; no preexisting condition unless genetic condition diagnosed
Arizona (1997)	Genetic condition	Life or disability insurance contracts	No coverage or rate discrimination unfair unless actuarial guidelines met
Arkansas (1997)	HIPAA-related provisions; genetic information	Group health plans	Genetic information not a preexisting condition without diagnosis of condition related to the information; no discrimination based on health status; health status includes genetic information
California (1994, 1995, 1996, 1998)	Asymptomatic genetic condition or predisposition	Health benefit plans (insured, self-insured, and MEWA) Disability insurance for medical expenses also regulated	No nontherapeutic use of information; confidentiality protections
Colorado (1994)	Direct tests for presence or absence of alterations in genetic material associated with disease	Health, group disability, long-term care insurance Does not cover life insurance or individual disability insurance	Therapeutic use permitted; no underwriting use in covered insurance
Connecticut (1997)	Information about genes or inherited characteristics derived from individual or family member	Individual or group health insurance	No use for coverage or rate determinations Can refuse to cover or can apply preexisting clause to person with symptomatic genetic disease.
Delaware (1998)	Broad definition of genetic information "about inherited genes or chromosomes, and of alterations thereof, whether obtained from an individual or family member, that is scientifically or medically believed to predispose an individual to disease, disorder or syndrome or believed to be associated with a statistically significant increased risk of development of a disease, disorder or syndrome"	Informed consent and confidentiality Health Insurance	Restricts access to genetic information, except where person consents or otherwise permitted by law Insurers are permitted to have access under some circumstances Prohibits discrimination in access or rates of health insurance based on genetic characteristics
Florida (1997)	Information from genetic testing for asymptomatic persons DNA analyses	Health benefit insurers (insured, self-insured plans) No protections for coverage and rate determinations of life insurance, disability income policies, long-term care policies, or certain other insurance policies Public or private entities performing DNA analyses	Health insurers cannot use genetic information in coverage or rate determination unless diagnosis of disease Other than in statutorily limited situations, public or private entities performing DNA analysis must obtain informed consent of person and result is property of individual

Table 1. ***Continued***

State (year enacted)	Type of Genetic Information	Area of Protection	Type of Protection
Georgia (1995)	Genetic testing for asymptomatic genetic conditions.	Health and sickness plan payers	Therapeutic use only
		Does not cover self-insured plans subject only to ERISA	Confidentiality protection
		Statute does not protect against use in life insurance, disability income policies, long-term care, Medigap, and other policies	
Hawaii (1997)	Genetic information of individual or family member	Health Insurance. Does not apply to life insurance, disability income insurance, and long-term care insurance	May not use genetic info for coverage or rate determinations
	HIPAA provisions		Confidentiality protection
		Group and individual health insurance	Implements HIPAA
Idaho (1997)	Small employer reforms	Small employers	Limits on use of health status
Illinois (1997)	Genetic testing for abnormalities or deficiencies linked to current disorders or susceptibility. No confidentiality protection for determination that person suffers from disease, "whether or not currently symptomatic."	Accident, health insurance policies	Accident and health insurers may not seek or use for nontherapeutic purposes
		Employers	Confidentiality protection
			Individual can release favorable results to insurers for consideration
			Employers may use if consistent w/ADA
Indiana (1997)	Results of genetic tests	Insurers other than life insurers	Insurance companies not entitled to access to genetic test results unless individual gives specific written consent
	Direct genetic screening or testing of individual's genes for defects linked to disorder or susceptibility or damage	Health care services coverage	Non-life insurer may not use individual or family members genetic testing results to determine coverage or rates
	Does not cover detection of genetic disorder through its manifestation		Insurer may consider favorable results released by individual
			Confidentiality protection
Iowa (1992, 1998)	Genetic Testing in Employment	Employers, labor organizations, licensing authorities	No mandatory pre-employment genetic testing; employee can volunteer and give informed consent for genetic testing related to occupational risks
	Small group reforms	Small groups	
			Restrictions on use of health status, including genetic information
Kansas (1997)	Genetic screening or tests	Health insurers, HMOs	Covered entities should not request test or condition insurance on testing; covered entities should not establish rates based on results
	HIPAA -related	Group and individual policies must be renewed without consideration of health status, which includes genetic information	Life insurance, disability, and long term care coverage must set rates based on reasonable risks
			Life insurers not covered

(*continued*)

Table 1. ***Continued***

State (year enacted)	Type of Genetic Information	Area of Protection	Type of Protection
			Group and individual policies must be renewed without consideration of health status, which includes genetic information
Kentucky (1998)	Genetic testing or information	Group or individual health plan; group insurers; disability income insurers	"A group or individual health benefit plan or insurer offering health insurance in connection with a health benefit plan or an insurer offering a disability income plan may not request or require an applicant, participant, or beneficiary to disclose to the plan or insurer any genetic test about the participant, beneficiary, or applicant"
Louisiana (1997)	Genetic information is all information about genes, inherited characteristics, or family history/pedigree expressed in common language	Health insurers, including employee benefit plans Life disability income, and long-term care insurance policies excluded	Health insurers may not use genetic information of individual or family member in coverage or rate determinations Confidentiality protection
Maine (1997)	Genetic tests or results	Employers Health insurers Life, disability, long term care insurers	Employers prohibited from discriminating based on refusal to take genetic test or results of genetic test unless bona fide occupational qualification Health insurers cannot use genetic test results to discriminate Life, disability, long-term care insurers may discriminate only if reasonably related to expected claims experience
Maryland (1997)	Genetic test used to identify alterations in genetic material associated with disease or illness	Health insurance policies or contracts Section does not apply to life insurance policies, annuity contracts or disability insurance policies	May not use genetic test results for coverage or rate determinations Confidentiality protections
Minnesota (1995)	Presymptomatic test of genes associated with genetic conditions or predispositions	Health plan companies Life insurance policies	Health plan companies may not require or use genetic tests in coverage and rate determinations Life insurance companies may require use of genetic test but informed consent and confidentiality provisions apply
Missouri (1998)	Genetic testing and information	Insurers, employers	Insurers cannot require tests or consider results Excludes disability income insurance and long-term care insurance Employers can use genetic information when directly related to job responsibilities Confidentiality protections

Table 1. ***Continued***

State (year enacted)	Type of Genetic Information	Area of Protection	Type of Protection
Montana (1991, 1999)	Genetic condition Genetic information and genetic tests	Life and disability insurance Individual or group insurers	Coverage and rate discrimination permitted only where "substantial" differences in claims likely Insurers may not require genetic testing Insurers may not discriminate in coverage or rates Provisions do not apply to life insurance, disability insurance, or long-term care insurance
Nebraska (1997)	Genetic information	Individual and group insurers	Prohibits discrimination based on health status, which includes genetic information
Nevada (1997)	Genetic information is that obtained from a genetic test to determine abnormalities linked to disorder or susceptibility to disease	Health insurance Disability income and long-term care coverage excluded	Health insurers cannot require test for individual or family members and cannot use results in coverage and rate determination
Nebraska (1997, 1996)	Genetic information or a typical hereditary cellular or blood trait; genetic testing Genetic characteristics are inherited gene or chromosome or alteration thereof scientifically or medically believed to predispose individual to disorder or to predispose to disorder	Employers Health insurance; life insurance and annuities	Employers cannot use genetic information, etc., as basis for discrimination Confidentiality protections Genetic characteristics cannot be used in coverage and rate determinations for health insurance Discrimination in life insurance must be reasonably related to expected claims experience
Nevada (1997)	Genetic information and testing HIPAA	Group and individual health insurers HIPAA Implementation	Health insurers prohibited from requiring genetic tests or from using the results of such tests to discriminate in coverage or rates Long-term care or disability insurance not covered HIPAA Implementation
New Hampshire (1995)	Genetic Testing	Employers, labor organizations	Cannot require genetic testing or use to affect terms and conditions of employment Individual can consent to tests for susceptibility to workplace chemicals if employer takes no adverse actions based on results Genetic tests can be used to determine insurability for life, disability income, or long-term care insurance as part of employee benefit plan

(*continued*)

Table 1. ***Continued***

State (year enacted)	Type of Genetic Information	Area of Protection	Type of Protection
	Genetic information	Health insurers	Cannot require individual or family member to undergo testing and cannot use information in coverage or rate determinations
		Life, disability income, and long-term care insurance	Life, disability income, and long-term care insurers can use information
New Jersey (1996, 1997)	Genetic Information and Testing	Restrictions on testing apply to all "persons"	Informed consent required for most genetic testing
	Employers not permitted to require genetic tests	Special rules for life insurance or disability income insurance.	Commissioner on Banking and Insurance to establish regulations
	Genetic characteristics	Employers	Limits ability of employers to require testing or require access to test results
	HIPAA implementation	Group and individual insurers	
		Group and individual coverage	Prohibits discrimination in coverage or rates based on genetic characteristics
			HIPAA implementation
New Mexico (1998)	Genetic Information Privacy Act	Applies to most "persons"	Genetic analysis prohibited without informed consent
			Life, disability, long- term care entities are exempted
			Prohibits discrimination based on genetic analysis, genetic information, or genetic propensity, except that life, disability income, and long-term care insurers are permitted to make actuarily reasonable adjustments
New York (1996)	Genetic predisposition	Employers	Unlawful for employers to discriminate against individuals based on their disability, genetic predisposition, or carrier status
	Genetic tests, genetic predispositions	All persons; insurance companies	Informed consent and confidentiality required; special rules for consent and confidentiality for insurers
North Carolina (1997)	Genetic information: from individual or family member, about genes, gene products or inherited characteristics	Health benefit plans, including all those where regulation permitted by ERISA	Insurers may not make coverage or rate determinations based on genetic information
	HIPAA amendments	Excluded types of plans include disability income, long-term care, and Medigap coverage	No person corporations, etc., can engage in employment discrimination based on genetic information about the individual or family members
		Persons, corporations, etc.	HIPAA conforming amendments
		Group and individual plans	
North Dakota (1997)	Genetic information	Hospital and medical insurance	Genetic information is not a preexisting condition absent a diagnosis

Table 1. ***Continued***

State (year enacted)	Type of Genetic Information	Area of Protection	Type of Protection
Ohio (1997, repealed effective 2004; replacement provision)	Genetic screening or testing	Health insurers; government self-insurers	Health insurers cannot require testing or use results unless favorable results volunteered by applicant Provision repealed effective 2004; protections for genetic information obtained before 2004 remain
Oklahoma (1998)	Genetic Nondiscrimination Insurance Act; genetic information means result of a genetic test and does not include family history	Health and accident insurance, not disability income or long-term care	Prohibits discrimination based on genetic information "except to the extent and in the same fashion as an insurer limits coverage, or increases premiums for loss caused or contributed to by other medical conditions presenting an increased degree of risk" Insurers may discriminate based on manifestations of conditions Weak confidentiality protections
Oregon (1995, 1997)	Genetic characteristic: gene or chromosome or alteration thereof believed to cause or predispose to disease Genetic information can be about individual or family	Insurance providers include all those subject to state regulation; health providers	Informed consent and confidentiality protections Insurers cannot use favorable genetic tests as inducement to purchase insurance Genetic information cannot be used negatively in hospital and medical expense insurance
Pennsylvania (1996)	PKU and Insurance	Insurance companies	Insurance must cover PKU-related formula
Rhode Island (1997)	Genetic testing	Individual or group coverage, HMOs, nonprofit health corps and insurers	No use of genetic tests or results to affect coverage or rates for health coverage Disability income and long-term care policies not covered
South Carolina (1998)	HIPAA Implementation Genetic privacy act protecting genetic characteristics and genetic information	MEWAs Health coverage	Impermissible "health status" discrimination includes use of genetic information Informed consent and confidentiality provisions; prohibits discrimination in coverage or rates based on genetic information Excludes disability income, long-term care coverage, and other nongeneral health insurance types of policies
South Dakota (1997)	Genetic information and preexisting medical conditions	Individual, group, and small employer policies	Preexisting condition cannot include genetic information unless related condition has been diagnosed

(*continued*)

Table 1. ***Continued***

State (year enacted)	Type of Genetic Information	Area of Protection	Type of Protection
Tennessee (1997)	Genetic information; carrier status, genes that cause or predispose to disease	Insurance providers, but focused on medical or health insurance; excluding life insurance, disability income, long-term care policy and other types of insurance	No coverage or rate discrimination.
			Confidentiality protection
	Questions about family history excluded		Restrictions on ability of insurer or plan to consider health status; health status includes genetic information
	Insurers are permitted to ask questions about health of applicant and family	Small group market, individual and group coverage	
	HIPAA implementation		
Texas (1997)	Genetic information: that derived from genetic test for genes associated with predisposition to disorder	Employers, licensing authorities; group health benefit plans as permitted by ERISA	Discrimination based on genetic information or refusal of testing prohibited by employers, group health benefit plans, and licensing authorities
		Excluded insurance plans: specific disease plans; accidental death or dismemberment; Medigap, works compensation, long-term care, etc.	Confidentiality protection. Insurers may not coerce abortion in pregnant women carrying children with genetic conditions
Vermont (1997)	Results of genetic testing	Business of insurance	Genetic test results can be used where there is a reasonable relationship between the information and anticipated claims experience
	Genetic testing in employment, licensure, and insurance	Employers, licensing authorities, insurers	
			Employers or licensing authorities cannot require or use genetic test results
			Health insurers cannot require testing or use results in underwriting
			Disability or long-term care coverage excluded
Virginia (1996)	Genetic Information	Health insurers	No discrimination in coverage or rates based on genetic information
	HIPAA-related reforms on genetic information and health status	Insurer issuing group or individual coverage	Confidentiality protections
			Disability income insurance excluded
			Genetic information cannot constitute preexisting condition unless diagnosed condition
Washington (1988)	PKU-related provisions	Insurers, HMOs	Coverage of PKU-related formula
West Virginia (1997)	HIPAA implementing provisions	Various health plans	Limitation on use of health status; health status includes genetic information; preexisting medical condition only where genetic condition diagnosed
Wisconsin (1991, 1997)	Genetic test for disease or predisposition	Insurer or self-insured governmental programs	Cannot require testing or revelation of results and cannot use in coverage and rate determinations
	Genetic tests or information	Employment	
	HIPAA-related amendments	Various insurers or plans	Use in life insurance and income continuation insurance must be reasonably related to risks

Table 1. ***Continued***

State (year enacted)	Type of Genetic Information	Area of Protection	Type of Protection
			Employers may not require or use genetic tests; employees may request and provide informed consent for genetic tests related to occupational safety issues
			Genetic information not a preexisting condition without diagnosis of condition
Wyoming (1997, 1998)	HIPAA-related amendments	Group and small employer health plans	Insurers cannot discriminate based on health status; health status includes genetic information; no preexisting condition unless genetic condition diagnosed

Note: The table was first created using a Westlaw search of the state legislative database in Fall 1997. The table was updated with an additional Westlaw search in Fall 1999. The search results were compared with a separate survey published in Ref. (25).

There are at least two sources of genetic information: the general medical records of the individual or his or her family (which may reveal information about genetic disease), and the results of specific genetic tests performed on the individual or his or her family. Policy makers interested in protecting genetic information must regulate both sources. Advances in genetic testing have further complicated the issue by providing information about an individual's susceptibility to particular genetic conditions. An individual may be determined to be a "carrier" of a gene, that is, able to pass the trait on through reproduction but at no risk for expression of a genetic condition. For example, individuals can carry the gene associated with Tay-Sachs without personality experiencing the disease. Alternatively, genetic testing can reveal that a person has a genetic condition or disease ("diagnostic testing"), such as hemochromatosis. It can also reveal that a person will develop a genetic condition at some future point ("predictive testing"), such as Huntington's disease, or that the person is at greater-than-average risk for developing the condition or disease being tested, such as certain types of colon or breast cancer (7).

USE OF GENETIC INFORMATION BY EMPLOYERS, INSURERS, AND OTHERS

Buck v. Bell involved the use of genetic information by the state, which mandated sterilization as a method of protecting the public welfare. But private entities, such as insurance companies and employers, have also been interested in obtaining genetic information. These private parties have an interest in using the genetic information to make decisions about employment, insurance, and other issues. Employers might want to exclude potential employees who present higher health insurance costs or who might be susceptible to workplace injury or illness (8). In the 1970s, for example, employers began screening workers for sickle cell anemia, which led to stigmatization and discrimination against sickle cell anemia carriers in employment (1). Insurance companies selling health, disability, life, long-term care, or other types of policies might want to exclude applicants or raise premiums for those who have higher rates of illness, disability, or premature death (7,8).

Individuals who anticipate these forms of discrimination and social stigmatization will have an interest in restricting access to genetic information. There are some circumstances, however, where an individual might want to release genetic information to employers or others in order to gain more favorable treatment than they would otherwise receive. This discussion is concerned with the state and federal regulation of the use of genetic information by private parties in the context of health, life, and other types of insurance. The central issue in this complex web of regulation is control. Who will control the decision about whether to undergo genetic testing, the individual or some third party such as an employer or an insurer? Who will control access to and use of genetic information: the individual or a third party?

A number of different commissions, working groups, and other organizations have concluded that it is important to restrict access to genetic information. The view of the Task Force on Genetic Testing created by the National Institutes of Health (NIH) and Department of Energy (DOE) Working Group on Ethical, Legal, and Social Implications of Human Genome Research is typical:

> Protecting the confidentiality of information is essential for all uses of genetic tests.... Results should be released only to those individuals for whom the test recipient has given consent for information release.... Under no circumstances should results ... be provided to any outside parties, including employers, insurers, or government agencies, without the test recipient's written consent....No individual should be subjected to unfair discrimination by a third party on the basis of having a genetic test or receiving an abnormal test result. (9, pp. 14–15)

The common belief that the privacy and confidentiality of genetic information must be protected to encourage individuals to undergo testing does not resolve all questions.

Table 2. Legal Documents Used for Table 1

Alabama Stat. §§27-53-1 to 27-53-4 (Alabama 1999) (effective 1997)
Alaska Stat. §§21.54.100, 21.54.110 (Alaska & Mathew Bender 1999) (HIPAA-related provisions enacted 1997)
Arizona Revised Statues Annotated §20-448 et seq. (West 1999) (main genetic provisions enacted 1997). See also, 2000 Ariz. S.B. 1330 (amendments enacted during publication process).
Ark. Code Ann. 23-86-subch. 3 (Arkansas 1999) (HIPAA-related provisions enacted 1997)
Cal.Civ.Code Ann. §56.17 (West 1999) (main provisions enacted in 1995 and 1996)
Cal. Health & Safety Code 1374.7 (West 1999) (multiple enactments and amendments, including 1995, 1996, 1998)
Cal.Health & Safety Code Ann. §124975 et seq. (West 1999) (enacted 1995)
Cal. Ins. Code §742.405, 10123.3, 10140, 10143, 10146 et seq. (West 1999) (multiple amendments, including 1994, 1995, 1996, 1998)
Colorado Rev. Stat.Ann. §10-3-1104.7 (West 1999) (enacted 1994)
Connecticut General Statutes Annotated §§38a-476, 38a-476a (West 1999) (enacted 1996 for HIPAA conformity)
Delaware Code Annotated, Title 16, §1220-1227 (Del. 1998) (informed consent and confidentiality for genetic information) (effective 1998)
Delaware Code Annotated, Title 18, §2317 (Del. 1998) (prohibiting discrimination in health insurance) (effective 1998)
Florida Statutes Annotated §§627.4301, 760.40 (West 1999) (enacted 1997)
Official Code of Georgia Annotated §33-54-1 to -8 (1997) (enacted 1995)
Haw. Rev. Stat. Ann. §§431:10A-118 (individual health insurance); 432:1-507 (group health coverage); 432:1-607 (mutual benefit societies); 431:2-201.5 (HIPAA) (West 1999) (enacted 1997)
Id. Code §§41-4708 (Lexis 1999) (amendments including 1997)
Illinois Compiled Stats. Ann. 215, §5/356v (West 1999) (effective 1998)
Illinois Compiled Stats. Ann. 410, §§513/5, 15, 20, 30 (West 1999) (effective 1998)
Burns Indiana Code Ann. §16-39-5-2 (West 1999) (amended 1997)
Burns Indiana Code Ann. §27-8-26-1 to 11 (West 1999) (enacted 1997)
Iowa Code §§5132B.9A, 513B.10 (small group reforms, 1997); §729.6 (genetic testing in employment, 1992)
Kansas Statutes Annotated §40-2209, -2257, -2259 (Revisor, Kansas 1998) (sections amended or enacted 1997, 1998)
Kentucky Rev. Stat. Ann. §§304.12-085 (West 1999) (enacted 1998)
Louisiana Rev. Stat. Ann. §§22:213.7, 22:250.1 to 22:250.16, 22:1214 (West 1999) (provisions enacted 1997)
Maine Rev. Stat. Ann. 5:§19302, 24:§2159-C (West 1999) (enacted 1997)
Maryland Ins. Code Annotated §§27-208; 27-909 (1997)
Minnesota Stat. §72A.139 (West 1999) (enacted 1995)
Vernon's Missouri Statutes Ann., 24:375.1300 to .1312 (West 1999) (enacted 1998)
Montana Code Anno. §33-18-206 (enacted 1991), 33-18-901 to 904 (enacted 1999) (West 1999)
Neb.Rev.St. §§44-787, 44-6910, 44-6915,44-6916 (Neb. 1999) (enacted and amended, 1997, 1998, 1999)
Nevada Rev. Stat. §689A, §689B.420, §695B.069, §695.317 (enacted 1997) (West 1999)
New Hampshire Revised Statutes Annotated 141-H:1 to H:6 (N.H. and Lexis 1999) (enacted 1995)

Table 2. ***Continued***

New Jersey Statutes Ann. 10:5-12 (non-discrimination in employment); 10:5-43 to 10:5-49 (genetic privacy act); §17:48-6.18 (individual or group health policies); 17:48A-6.11 (medical service corporation contracts); 17:48E-15.2 (health service contracts); 17B:26-3.2 (individual health policies); 17B:27-36.2 (group health insurance); 17B:27-54 (HIPAA implementation); 17B:30-12 (insurance trade practices) (West 1999) (most provisions enacted in 1996, HIPAA implementation in 1997)
New Mexico Statutes, Ch. 24, Art. 21 (N.M. 1999) (enacted 1998)
McKinney's New York Civ. R.Law §79-1 (consent and confidentiality); McKinney's New York Ins. Law §2612, 3221 (insurance consent and confidentiality); McKinney's New York Exec. Law §292,296 (employers; discrimination) (West 1999) (many provisions enacted in 1996)
North Carolina Gen. Stat. §58-3-215 (genetic information in health insurance); §95-28.1A (prohibiting discrimination in employment); Ch. 58, Art. 68 (HIPAA amendments) (Lexis 1999) (most enactments 1997)
North Dakota Century Code 26.1-36.4-03.1 (Lexis 1999) (enacted 1997)
Ohio Rev. Code Annotated §1751.64-65; 3901.49 to 3901.50 (West 1999) (major portions enacted 1996)
Oklahoma Stat., Title 36, §§3614.1 to 3614.4 (West 1999) (enacted 1998)
Oregon Revised Statutes §659.036 (employers use of genetic information limited to bona fide occupational qualifications); §§659.700-.720 (genetic privacy), §746.135 (discrimination in health insurance) (1995, 1996)
Pa. Cons. Stat. §3902 (West 1999) (enacted 1996) (insurance must cover PKU-related formula)
R.I. Gen. Laws §27-18-52 (genetic testing and insurance), §§27-20-39 & 27-19-44 (non-profit hospital services/non-profit insurance), §27-41-53 (HMOs) (R.I. and Lexis 1998) (enacted 1998)
S.C. Code Ann. §38-41-45 (MEWAs), §38-71-670 (HIPAA and individual coverage), §38-71-840 (HIPAA and group coverage), §38-93-10 to -60 (S.C. 1999) (HIPAA implementation 1997, Genetic Privacy enacted 1998)
S.D. Codified Laws §§58-17-84, 58-18-45, 58-18B-27 (S.D. 1999) (pre-existing condition cannot include genetic information unless related condition has been diagnosed) (enacted 1997)
Tennessee Code Ann. §56-7-2701 to 2708 (genetic information non-discrimination in health insurance act), §56-7-pt 28 (HIPAA implementation) (Tenn. 1999) (enacted 1997)
Texas Labor Code Ann. §21.401 to 21.405 (West 1999) (discriminatory use of genetic information prohibited enacted 1997); Texas Insurance Code Ann. §21.73 (West 1999) (non-discrimination in insurance, enacted 1997); Texas Civ. St. Ann. Art. 9031 (limitation on use by licensing authority) (West 1999) (enacted 1997)
Vermont Stat. Ann. §§8:4724, 8:9331-9335, 18:9332-9333 (Vt. 1999) (enacted 1997)
Va. Code Ann. §§38.2-508.4 (genetic information privacy), 38.2-613 (confidentiality, enacted 1996) (Lexis 1999); §38.2-3431, 38.2-3432.3 (HIPAA-related amendments in 1997, 1998, 1999) (Lexis 1999)
Washington Rev. Code Ann. §§48.20.520, 48.21.300, 48.44.440, 48.46.510 (West 1999) (PKU provisions, enacted 1988)
W.Va. §§33-15-2a, 33-16-1a, 33-16-3k (Lexis 1999) (enacted 1997)
Wisc. Stats. Ann. §§111.372, 111.39 (employment, 1991), §631.89 (genetic testing and insurance, 1991), §§632.746 (HIPAA implementation, 1997) (West 1999)
Wyo. Stat. Ann. §§26-19-107, 26-19-306 (Wyo. 1999) (HIPAA amendments enacted 1997, 1998)

Which level of government should become involved in regulating genetic information, the federal government or the states? What uses of genetic information are permissible? When is the use of genetic information "unfair"?

OVERVIEW OF STATE AND FEDERAL REGULATION OF GENETIC INFORMATION

It is important and useful to understand the distinctions between the regulatory authority of the federal and state governments regarding the control of genetic information. The federal government has power over interstate commerce which theoretically could include the regulation of insurance companies and the provision of health and other benefits by employers. Nevertheless, the federal government for many years failed to enact much substantive regulation touching on the problems created by genetic information. Indeed, under the McCarran-Ferguson Act, the federal government has long ceded primary authority over the regulation of the business of insurance to the states (10). Where federal law existed, it tended not to impose substantive standards of conduct regarding the use of genetic information.

Under the Employee Retirement Income Security Act of 1974 (ERISA), the federal government established various requirements for the maintenance and operation of employee benefit plans (2). These federally regulated benefit plans include those in which the employer undertakes to provide medical or health coverage for its employees. But ERISA did not originally regulate the substantive content of those employee benefit plans in any way that impinged on the ability of the employee benefit plan to seek out and make use of genetic information. Other federal statutes, such as the Rehabilitation Act of 1973 or the Americans with Disabilities Act (ADA), prohibited certain types of discrimination by employers and others, but these statutes were not drafted to make clear that they included protections for all types of genetic information or discrimination (11,12).

Thus, for several years, states were alone in regulating access to and use of genetic information in insurance or employment-related areas. Many states enacted specific statutory protections for genetic information as a matter of self-initiated state policy (see Table 1). State legislation in this area was sparked by the knowledge that advances in genetic technology were creating conflicts between individuals, insurers, and employers. Individuals often wanted to know relevant genetic information but then realistically or unrealistically feared discrimination by others (11,13).

The federal ERISA statute governing employee benefit plans created some significant barriers to state regulation (2). ERISA contains a vigorous "preemption" clause. The clause invalidates state attempts to regulate many types of employee benefits, including state efforts to regulate the way in which employers use genetic information in certain employee benefit plans such as in health benefits coverage. Congress appeared to recognize the regulatory vacuum created by the ERISA preemption clause when it enacted the Health Insurance Portability and Accountability Act (HIPAA) (3). This federal statute established substantive guidelines for the use of genetic information in health plans and also explicitly authorized consistent state regulation. HIPAA sparked a new wave of state legislative activity so that now nearly all states have enacted measures governing the use of genetic information in at least some contexts (Table 1).

STATE REGULATION AND ERISA PREEMPTION BEFORE THE HIPAA AMENDMENTS

State Regulation Before HIPAA

States began to be concerned with the use and misuse of genetic information over 20 years ago because of the problems associated with sickle cell anemia testing. These concerns were augmented over time as geneticists began to identify a host of conditions thought to be related to underlying genetic traits. States sought to encourage individuals to use genetic testing services by protecting the voluntary and confidential nature of genetic testing. On the other hand, states recognized that providing access to genetic information for individuals but not insurers could create problems. Is it "fair," for example, to permit an individual who knows she is susceptible to early disability and death, but does not share this information with her insurer, to purchase large amounts of disability and life insurance coverage (7,14,15)? In an effort to deal with some of these problems, states attempted to establish when insurers, or others, should be permitted to gain access to an individual's genetic information and to establish certain domains within which discrimination based on genetic characteristics would be permissible.

As Table 1 indicates, a number of states responded to these conflicts by enacting strong privacy protections for genetic information in the early to mid-1990s. State genetic privacy acts, such as those enacted in California, Georgia, Nebraska, and New York, generally provided strict protections for the confidentiality of genetic information. Some states even provided that a person's genetic information would be the "property" of the individual. Typically an individual could not be tested without giving specific informed consent, and genetic test results could not be released without the individual's consent.

Some states prohibited employment discrimination based on genetic factors. New Hampshire, Nebraska, and New York, for example, prohibit employers from using genetic test results to affect the terms and conditions of employment. These state employment discrimination statutes supplement the sometimes weak protections offered by existing federal statutes.

Some states also sought to regulate the use of genetic information by insurers, particularly health insurers before HIPAA's enactment. California, Colorado, Georgia, Minnesota, Montana, Oregon, and Wisconsin adopted restrictions on the use of genetic testing or genetic information by insurers. Each of these states focused on health insurance, leaving life insurers and other types of insurers relatively free from state regulation.

The early state efforts to regulate the use of genetic information by insurers appeared to be driven by three principles: (*1*) genetic information is a "good"

which individuals should be encouraged to obtain, (*2*) access to medical and health insurance should not be restricted based on genetic information, and (*3*) genetic discrimination in other types of insurance, such as disability income or life insurance, should not be prohibited. The fact that so many states attempted to regulate in this area suggests a high level of state concern about the impact of genetic information on the health insurance market.

State efforts to regulate health insurance were complicated by a complex federal and state regulatory structure. Congress largely delegated the regulation of the business of insurance to the states in the McCarran-Ferguson Act (10). Yet two important federal statutes affected the ability of states to regulate in this area. The first, and most important, impediment to state regulation of the insurance market was ERISA (2). The second relevant federal statute was HIPAA, which amended, ERISA and prohibital certain types of genetic discrimination (3).

Employee Retirement Income Security Act of 1974

ERISA regulates employee benefit plans, including health coverage, disability, and life insurance provided as a benefit of employment (2). The benefit plans covered by ERISA include those in which an employer enters into a contract with another entity such as an insurance company. ERISA also applies where an employer "self-insures," or bears the risk of paying the benefits directly. (Governmental plans, church plans, and a few others are excluded from ERISA regulation.) (2).

About 70 percent of the people who have private health insurance coverage obtain ERISA benefit plans through employment (19). However, most employees who have health insurance through employment are covered by plans that are "self-insured," that is, their employer bears the risk of medical expenses (19).

Until recently ERISA was silent about the employer's ability to take genetic conditions into account in the terms and conditions of his or her employee benefit plans. Nevertheless, the ERISA preemption clause presented a substantial barrier to effective state regulation of genetic information.

Federal Preemption of State Regulation of Genetic Information

Unlike many federal laws, ERISA was drafted with an explicit and broad preemption clause that prevents states from regulating employee benefit plans. The statute states that ERISA "shall supersede any and all state laws insofar as they may now or hereafter relate to any employee benefit plan" (2). The major exception, for our purposes, is the "savings clause," which explains that "nothing in this subchapter shall be construed to exempt or relieve any person from any law of any state which regulates insurance" (2). This provision "saves" state laws regulating insurance from ERISA preemption. However, ERISA makes clear, that in regard to employer self-insured plans, "an employee benefit plan shall [not] be deemed an insurance company ... or to be engaged in the business of insurance ... for purposes of any law of any state purporting to regulate insurance companies" (2). Self-insured employee benefit plans are thus completely protected from state regulation.

The Supreme Court has often been called upon to interpret this complex and confusing provision. In order to decide whether ERISA preempts a state law, the Court must start with the presumption that Congress does not intend to supplant state law (20,21). In evaluating whether the normal presumption against preemption has been overcome, the Court considers whether the state law "relates to" an employee benefit plan. The Supreme Court has said that a state law "relates to" an employee benefit plan if it (*1*) has a "connection with" or (*2*) "reference to" such a plan (22). State laws that "relate to" an employee benefit plan in this fashion are preempted and given no effect because of the ERISA preemption clause, unless an exception to preemption can be found.

Once the Court rules that a state law "relates to" an employee benefit plan and thereby falls under ERISA preemption, it must determine if the law "regulates insurance" and thus escapes preemption under the "savings clause." To determine if a law regulates the "business of insurance," the Court must first consider whether, from a commonsense view, the contested prescription regulates insurance (23). The Court must also consider three factors to determine whether the regulation fits within the "business of insurance," the phrase used in both the McCarran-Ferguson Act and the ERISA: (*1*) whether the practice has the effect of transferring or spreading a policyholder's risk, (*2*) whether the practice is an integral part of the policy relationship between the insurer and the insured, and (*3*) whether the practice is limited to entities within the insurance industry (23).

Under the "deemer" clause, a state law that relates to an employee benefit plan but is "saved" from preemption because it constitutes the state regulation of insurance can still not be applied to self-insured employee benefit plans. A "self-insured" plan — one in which the employer bears the risk of his or her employee's health care costs directly — cannot be treated as an insurance plan and thereby be subjected to state regulation of insurance (20).

ERISA's broad preemption clause has presented a significant barrier to state regulation of the use of genetic information. A large number of states have sought to prevent entities from using genetic information to determine whether to provide health benefits for individuals. California law provides, for example, that genetic information should not be used in making coverage or rate determinations for medical benefits (see Table 1). Since most persons who have private insurance obtain it as a benefit of employment, California's statute is only effective if it is applied to employee benefit plans. Under the ERISA preemption clause, California's statute may be considered state "regulation of insurance" and may be "saved" from preemption (2). This means that the state statute can be applied to insurance companies that sell health insurance to employers for the benefit of their employees. The statute cannot be applied, however, to employers who self-insure their health benefits plans. Employers who self-insure are immune from state regulation under the "deemer" clause of ERISA. It is for

this reason, to avoid state regulation, that many employers provide employee benefits, such as health care coverage, through self-insurance.

The same principles apply to state attempts to regulate the use of genetic information in other areas, such as disability income and life insurance. Under ERISA, states can regulate insurance companies selling these types of benefits and may prohibit or permit the use of genetic information in determining the terms and conditions of the insurance plan. However, states are barred from attempting to regulate the use of genetic information by employers who self-insure a disability or death benefit plan. Consequently most states permit the use of genetic information as an underwriting consideration in life and disability income insurance (24). Indeed, ERISA preemption has little impact on employers who self-insure because state law often permits use of genetic information in coverage or rate determinations for disability and life insurance.

EFFECT OF THE HEALTH INSURANCE PORTABILITY AND ACCOUNTABILITY ACT OF 1996

HIPAA Amendments to ERISA

Although by 1996 many states had attempted to regulate the use of genetic information, the ERISA preemption clause limited the effectiveness of state regulation, since it shielded self-insured employee benefit plans, such as those providing health coverage, from state regulation. In addition, more than half the states had not enacted *any* legislation governing the use of genetic information. Critics were quick to note that state regulation had not been an effective method of restricting improper use of genetic information in health care coverage determinations.

Congress responded to these critics with amendments to ERISA that govern the use of genetic information in many types of health insurance coverage. While recognizing that state regulation is still important, particularly in areas outside employer provision of health coverage, in 1996 Congress amended ERISA to provide protection from genetic discrimination in some contexts.

HIPAA amends ERISA to directly regulate the use of genetic information in several important ways (3). The most serious restrictions are imposed on group health plans. HIPAA prohibits the use of "genetic information" in making eligibility determinations for group health plans, whether the health plan is insured or self-insured. Group plans are also prohibited from charging individual members higher premiums because of genetic information about themselves or their dependents. HIPAA provides that these plans may not consider an individual's genetic characteristics to be a preexisting condition unless the genetic characteristic has given rise to a diagnosis of an actual condition related to the genetic information (3).

Other sections of HIPAA, while not specifically focused on the issue of genetic information, provide important benefits for individuals who have a genetic condition. The statute establishes a "credit" for prior insurance coverage that may be applied to the preexisting limitations period of a new policy. This means that an individual with a diagnosed genetic condition will actually have greater job mobility, since he or she will be able to achieve continuous coverage for medical treatments even after switching jobs (3). The statute also contains some protections that apply to health insurer issuers (including HMOs) and individual health plans. The statute establishes standards for guaranteed availability and renewability, for example, that could prove helpful to individuals who have genetic conditions (3).

Finally, HIPAA does not amend the general ERISA preemption clause but does provide that states may continue to regulate certain aspects of how health insurance companies use genetic information (3). States will monitor the implementation of the federal rules governing insurers and HMOs, for example. States can also enact alternate protections, so long as their efforts do not prevent the application of federal law. Thus states may enact certain measures that provide greater protection to individual enrollees than that under federal law. Most states already have passed state legislation implementing HIPAA's protections in the state insurance market.

Even with HIPAA, there are still big gaps in protection for those concerned about loss of confidentiality and/or genetic discrimination (17). First, HIPAA regulates the "use" of genetic information but does not otherwise protect the creation or privacy of the information. Second, HIPAA focuses federal regulation on group plans; individual health insurance policies are offered lesser protection. Third, even for group plans, HIPAA does not prohibit an insurer from charging higher overall rates for a plan based on genetic information about plan participants.

Current Status of State Regulation

As Table 1 shows, most of the states that now regulate the use of genetic information enacted their legislation during or after 1996, the year of HIPAA's enactment. Most state regulation now focuses on insurance issues. A large number of states attempt to control the use of genetic information by various types of insurers. There are three basic variables that characterize state legislation: the definition of "genetic information," the types of insurers or entities regulated, and the substance of the insurance regulation (25).

The first important variable is the type of genetic information covered under the state statute. States have taken a variety of approaches to defining the types of genetic information subject to statutory protection. Most states focus on the use of "genetic testing" and protect test "results." Several take into consideration the need to protect both the test results of the "individual" and his or her "family members." Louisiana recognizes the need to protect other types of genetic information, such as information about genetic conditions that might be found in the family history or medical records of an individual.

A significant number of jurisdictions, such as California, protect only genetic information about "predispositions" or "susceptibility," while permitting disclosure of information about actual genetic conditions or diseases. Some states, such as North Carolina, seem to protect the full range of genetic information, up to and including a determination that an individual in fact is currently exhibiting the effects of a genetic condition.

The second variable in state regulation is the identity of the entities subject to regulation. As shown in Table 1, states typically have sought to regulate a wide range of insurers and insurance-like entities. Thus state laws focus on health benefit plans, whether insured or self-insured (e.g., see Alabama, California, and Florida). The statutes also attempt to restrict the use of genetic information by health maintenance organizations, preferred provider organizations, and other health insurer-hybrids. State regulatory schemes mention other types of insurance, such as life insurance, disability income insurance, Medigap coverage, and other types of policies.

The third variation in state regulation is the differential treatment of genetic information based on the insurance coverage type. Each state regulating the use of genetic information by insurers has enacted strong prohibitions against use by health or medical insurance entities. A large number of states, however, permit discrimination — or at least "actuarially sound" discrimination — by life, disability income, and other types of insurers. Most of states regulating the use of genetic information in insurance thus have established that discrimination in health or medical benefits is "unfair" while discrimination in other types of insurance is often "fair" and should be permitted (5).

Some of the legislation would doubtlessly have been enacted even without HIPAA. But many of the post-1996 statutes have clear objectives on the implementation of HIPAA-type provisions within the state. Post-HIPAA, state legislation that relates to employee benefit plans will consequently not be preempted so long as it does not prevent the application of federal law (3). Indeed, states can and have considered provisions that would establish standards more protective than those offered under HIPAA.

CONCLUSION

Genetic information is becoming an increasingly important area for state and federal regulation. Most states have already specifically regulated the confidentiality and use of genetic information. There are significant barriers, however, to effective state regulation, including federal preemption under ERISA.

HIPAA's enactment and the subsequent flurry of state legislative activity have been helpful to those concerned about the confidentiality of genetic information and the risk of genetic discrimination. However, commentators have identified the gaps that still remain. It is clear that further federal action must address at least two of these gaps: (*1*) the risk of discrimination in group premiums or rates based on genetic information, and (*2*) the risk of the loss of privacy for genetic information (17). Before the Congress are versions of the Patients Bill of Rights that would protect individuals in group plans from discrimination in the provision of services based on genetic information (26,27). Congress is also considering bills specifically designed to protect the confidentiality of genetic information and to restrict the imposition of genetic testing (28).

BIBLIOGRAPHY

1. M.B. Kaufmann, *Loyola Univ. Chicago Law J.* **30**, 393–438, 400–401 (1999).
2. Employee Retirement Income Security Act of 1974, 29 U.S.C.A. §§1001–1169 (West 1997).
3. Health Insurance Portability and Accountability Act, Public Law 104–191.
4. *Buck v. Bell*, 274 U.S. 200 (1927).
5. P.A. Lombardo, *N.Y.U. Law Rev.* **60**, 30–62 (1985).
6. L. Hood and L. Rowen, in Rothstein, ed., *Genetic Privacy*, 1997.
7. J. Gaulding, *Cornell Law Rev.* **80**, 1646–1694 (1995).
8. K. Rothenberg et al., *Science* **275**, 1755–1757 (1997).
9. Task Force on Genetic Testing, NIH-DOE Working Group on Ethical, Legal, and Social Implications of Human Genome Research, *Final Report: Promoting Safe and Effective Genetic Testing in the United States*, N.A. Holtzman and M.S. Watson, eds., U.S. Government Printing Office, Washington, DC, 1997.
10. McCarran-Ferguson Act, 15 U.S.C.A. §1011–1014 (West 1999).
11. Rehabilitation Act of 1973, 29 U.S.C.A. §§701–796 (West 1999).
12. Americans with Disabilities Act, 42 U.S.C.A. §12101–12213 (West 1999).
13. *Norman-Bloodsaw v. Lawrence Berkeley Lab.*, 135 F.3d 1260 (9th Cir. 1998).
14. E.M. Holmes, *Kentucky Law J.* **85**, 503–664 (1996, 1997).
15. R.A. Bornestein, *J. Law Policy* **4**, 551–610 (1996).
16. L.F. Rothstein, Disabilities and the Law, §5.02 (Shepard's/McGraw-Hill, 1992 and 1997 Supplement).
17. M. Rothstein and S. Hoffman, *Wake Forest Law Rev.* **34**, 849–888 (1999).
18. Civil Rights Act of 1964 (Title VII), 42 U.S.C.A. §2000e-2 (West 1999).
19. M.A. Rothstein, *J. Law Health* **9**, 109, 113 (1994, 1995).
20. M.A. Bobinski, *University of California at Davis Law Rev.* **24**, 277–278 (1990).
21. *De Buono v. NYSA-ILA Med. and Clinical Serv. Fund*, 520 U.S. 806 (1997).
22. *California Div. of Labor Standards Enforcement v. Dillingham Const.*, 519 U.S. 316 (1997).
23. *UNUM Life Ins. Co. v. Ward,* 119 S. Ct. 1380 (1999).
24. C.M. Keefer, *Indiana Law J.* **74**, 1375–1395 (1999).
25. W.F. Mulholland II and A.S. Jaeger, *Jurimetrics* **39**, 317–320 (1999).
26. Patients Bill of Rights Act of 1999, 1999 U.S. Congr. Senate Bill 1344 [introduced by Senator Lott (R)].
27. Patients Bill of Rights Act of 1999, 1999 U.S. Congr. Senate Bill 1256 [introduced by Senator Daschle (D)].
28. Genetic Information Disclosures, 1999 U.S. Congress House Bill 2555, 1st Sess. 106th Congr.

See other entries GENETIC INFORMATION, ETHICS, INFORMED CONSENT TO TESTING AND SCREENING; GENETIC INFORMATION, LAW, LEGAL ISSUES IN LAW ENFORCEMENT DNA DATABANKS; GENETIC INFORMATION, LEGAL, FDA REGULATION OF GENETIC TESTING; GENETIC INFORMATION, LEGAL, GENETIC PRIVACY LAWS.

GENETIC INFORMATION, LEGAL, FDA REGULATION OF GENETIC TESTING

ANNY HUANG
Wachtell, Lipton, Rosen, and Katz
New York, New York

OUTLINE

INTRODUCTION

Genetic tests are the means by which individuals may unlock the wealth of information contained in their genetic constitution, and as the understanding and significance of genetic information advances, the scope of genetic testing expands accordingly. However, growth in genetic testing has opened new avenues of commercial exploitation that raise problems with ensuring adequate evaluation of the scientific legitimacy of genetic tests being offered to the public. The Food and Drug Administration (FDA) subjects genetic tests sold as kits to full regulatory review but leaves genetic tests provided as services entirely free from FDA scrutiny. The federal government does not review genetic testing services for the soundness of the scientific claims made for the tests, and there is very little substantive review of genetic testing services at the state level. The issue of whether the current state of regulation and the distinction between test kits and testing services is reasonable becomes more pressing as the genetic testing industry gains momentum.

Like many achievements in biotechnology, developments in genetic testing have been shadowed by concerns about their use and implications: scientists worry how their research will be used, physicians wonder whether the benefits outweigh the harms, and ordinary people face philosophically uneasy choices about whether knowledge leads one to the garden or away from it. The nuanced and weighty problems surrounding the use of genetic technology have been struggling for years toward resolutions that might suggest a basis for practical action. While numerous states have passed laws restricting the use of genetic information (1,2) and a consensus appears to be growing regarding the privacy of genetic information, basic questions about regulating genetic tests themselves remain mired in complex theoretical and scientific controversies. For many years the Task Force on Genetic Testing, an entity organized by the National Institutes of Health (NIH) and the U.S. Department of Energy, provided leadership in the development of national policy on regulating genetic tests. The Task Force produced its pivotal Final Report in 1997 (3) which contained extensive analysis of the regulatory and scientific issues related to genetic testing and provided groundbreaking detail on the economics and practice of the genetic testing industry (3). However, the role of FDA, a natural candidate to lead regulation of the safety and efficacy of genetic tests, was left surprisingly murky in the Report's recommendations. In fact the Final Report did not specify an appropriate candidate to monitor the scientific validity of genetic testing services. This result is remarkable in light of the fact that the Task Force's creation derived in part from an Institute of Medicine report criticizing the disparity between test kits regulated by the FDA and unregulated testing services (4,5). Panel members were clearly attentive to the issue and had questioned publicly the business strategy of marketing tests as services rather than kits specifically for the purpose of evading FDA oversight. Dr. Neil A. Holtzman, chairman of the Task Force, observed, "Companies don't create kits so they can circumvent the FDA regulatory process" (6). Yet the Final Report refrained from proposing an extension of FDA oversight to genetic testing services.

Such reserve, however, may have derived more from the Task Force's consensus style of decision making than from a fundamental objection to FDA involvement in genetic testing regulation. Part of what made the Task Force's conclusions so forceful and well-informed was the diversity of interests represented on the panel, including from the biotechnology industry (3). The Final Report of the Task Force had the distinction of being unanimously approval by its members with no abstentions. However, several Task Force members had openly doubted whether the FDA was the right entity to regulate genetic testing services (3). In fact, as early as April 1997, the Task Force on Genetic Testing "tabled any efforts to come up with a recommendation for the role that the FDA should play in regulating genetic testing" (3). In addition FDA itself showed considerable reluctance in assuming greater responsibility. In her testimony before the House Subcommittee on Technology, the FDA Deputy Commissioner Mary K. Pendergast stated that, "[t]o date, the FDA has minimal involvement with genetic testing," and cautioned that careful weighing on the relevance of further FDA efforts would be necessary:

> If the FDA is to do any of the additional regulating, we would have to evaluate how these concerns fit with other concerns facing the Agency, e.g., product approval and regulation, infectious disease transmission through foods, blood, and tissues, and examine how the harms from inaccurate genetic testing stack up against those other priorities. If additional oversight is mandated, there is the question of resources and how to pay for the oversight (7).

The agency has repeatedly taken the position that it will not exercise jurisdiction over tests marketed as services (8). Another full-length study, which supported the imposition by "federal regulators" (9) of minimum standards for the positive predictive value (PPV) of genetic testing services, came to the following conclusion: "In light of the public and political pressures on the FDA, such regulation might best be introduced through the [Centers for Disease Control], [Federal Trade Commission], [Health Care Financing Administration], or [Department of Health and Human Services] by, for example, modifying [the Clinical Laboratories Improvement Act]" (9, p. 1299).

While the new Advisory Committee on Genetic Testing in the Department of Health and Human Services (DHHS), formed at the recommendation of the Task Force, closely involves FDA in its activities, it is proceeding with the collection and analysis of data on the analytical validity, clinical validity and clinical utility of genetic tests leaving aside, as premature, the issue of the appropriate source and level of oversight (10). The question that remains unanswered is why the distinction between commercial services, which FDA believes it has authority to regulate, and kits makes sense. A Task Force member affiliated with OncorMed expressed the view that FDA did not have the appropriate level of expertise (6), but the criticism applies equally well to the agency's competence in regulating genetic testing kits, whose ranks can be expected to swell in the next few years. According to an FDA official: "At present we estimate that there are, or soon will be, dozens of companies or laboratories offering hundreds of different genetic tests to the public and it is projected that this number will grow substantially" (7). Currently 5,000 genes associated with genetic disorders have been cataloged for the Human Genome Initiative by a group at Johns Hopkins University (11). It appears that amid herculean efforts to collect information, resolve grave social issues, and achieve consensus among the various constituencies in genetic testing, the proposals refrained from charging any particular regulatory body with assuring consumers that the genetic tests being performed as a service, which represents the vast majority of emerging tests, are scientifically valid, despite agreement as to the necessity of such measures (9,12). That is not to say that detailed and fact sensitive proposals were not made to secure the validity and analytical sensitivity of these tests through improvements at the provider or clinical level. With regard to some of the principles, however, it cannot be enough to presume the clinics will regulate themselves, and, perhaps due to political constraints, the Final Report did not specify which regulatory body should enforce the recommendations.

The regulatory standards were introduced presumably to ensure that genetic testing is made available in clinical laboratories whose clinical validity has been established, that is, their positive predictive value (PPV), unless it is collecting data on clinical validity under either an IRB-approved protocol or conditional premarket approval agreement with the FDA.

The task of developing an understanding of how FDA *should* and, from a legal perspective, *could* fit into a coherent and flexible regulatory plan remains open. The efforts of the Task Force and other institutions and professionals have laid the difficult ground work for the development of standards and protocols in the genetic testing community, and a slight shift in perspective could suggest a path to greater clarity in regulations and liability. By focusing on the problem from a test recipient's perspective, the issue becomes remarkably simple. Despite some empirical differences in delivery, the information dynamics from the patient's perspective produce as great a need for protection in the area of commercial genetic testing services as for a test packaged as a kit. Indeed, apparent factual distinctions that may seem important from a regulatory standpoint turn out to be irrelevant when considered from the view of a genetic test recipient. Emphasizing these considerations is quite appropriate in determining the applicability of FDA oversight, furthermore, because the national protection of consumers is the animating purpose behind food and drug legislation even when it encroaches upon the prerogatives of medical practice. Subject to the boundary issues involved in legal jurisdiction (14), FDA is the most logical and efficient choice as the regulatory actor. Nothing in the Final Report or in available information defeats this view or suggests a better alternative. While it is important to recognize the political and institutional constraints that may dissuade the agency from undertaking regulation of genetic testing services, it is equally important to be aware that these external considerations may be charting a course in the wrong direction.

Ensuring public health safety in an age of new biological technologies will undoubtedly subject FDA to evolutionary pressures, but the novelty and complexity of genetic technology should not deflect the agency from its traditional goals and areas of oversight. FDA already has begun to review the early products of genetic technology, and it will doubtless face many more in this millennium (16–18). The use of innovative forms to deliver essentially commercial products should not confound regulation. It is important, however, to make a distinction between commercial testing services that have the character of products and traditional laboratory analysis. A bright-line rule not only serves regulatory goals and eases compliance, but is advisable to stay well within the bounds of federal authority. The dangers of genetic information are profound and warrant the public delegating the resources and mandate to the FDA to ensure that the troubling issues and agonizing choices occasioned by genetic testing are not compounded by poorly developed or even misleading information.

BACKGROUND ON GENETIC TESTING

The potential scope of genetic testing as a commercial enterprise has become increasingly clear over the last decade (19). In 1986 the conductors of the survey found only 118 companies likely to be offering or researching genetic tests, and of those 85 companies responded to the survey. Only 22 were performing or developing tests. In contrast, the 1996 survey found a target audience of 594 biotechnology companies of which 461 responded. About a third of respondents were engaged in genetic testing

activity (3). The once largely academic activity has moved into the marketplace, where more and more companies are realizing that it can be big business. A survey conducted in 1986 found 22 biotechnology companies offering or developing genetic tests, whereas a similar survey conducted ten years later found 147 companies. Tests have been developed for hundreds of conditions (20), including Alzheimer's, colorectal cancer, and melanoma with others such as asthma and even deafness on the horizon (21–24) and according to NIH, more than 450 research programs are working to develop more (21). It should be noted that the press and literature on genetic tests quote a large variety of figures for how many genetic tests have been developed deriving largely from differences in definition. Because thousands of genes linked to disorders have been identified, it would be possible to say that thousands of tests have been developed. In many of these cases, however, the disorder is extremely rare or the linkages have not been established with sufficient generality to justify use in a clinical setting. Nonetheless, the commercial potential of viable tests is considerable. The total DNA diagnostic market is estimated to exceed $6 billion by 2005 (25). Depending on the complexity involved and the number of genes to be screened, a single test may cost anywhere from a few hundred to several thousand dollars (21,26,27) and might be used by millions every year. For example, the most commonly used genetic screening test for phenylketonuria is used on millions of newborns annually (20). The proceeds from even one test could prove very lucrative. Consider, for example, the osteoporosis test being developed by Medical Science Systems, Inc. (28). The company estimated that if 2 percent of all affected Americans, approximately 500,000 people, are tested at $200 each, the market for this one test alone is worth $100,000,000 in revenues (28). MSSI focuses on the national market in estimating the commercial potential of a test. (Also companies such as Myriad, Salt Lake City, Utah, and Genetics and I.V.F. Institute, Fairfax, Virginia, comply with the regulations of jurisdictions as far away as New York in their quest for a national audience.)

Like the rest of the biotechnology industry, however, these endeavors are new and are dependent on recent breakthroughs in science (3). It is a young and dynamic industry (29). According to the Task Force Report, "The companies engaged in testing activities operate in an extremely dynamic environment, frequently undergoing restructuring, forming new partnerships, embarking on new initiatives, or dropping projects" (3, app. III). The various obstacles involved in entering the genetic testing market at this time — regulatory uncertainties, ethical issues, rapid technological evolution, the inherent limit on the utility of a test to once per patient — serve to dampen rapid maturation in this sector (3). Also most single-gene disorders are quite rare and have less market potential, so to earn high profits the tests must involve more common disorders, which often turn out to be more genetically complex (3,30). Investing in research for more common multiple-gene disorders such as cancers and heart ailments is expensive and risky (3), but promises the biggest payoffs in terms of patents and first to market profits. Consider, for example, the unseemly race to patent the BRCA1 breast cancer gene test, which generated a considerable amount of controversy in 1996 (31). Alternatively, a company could commercialize technology licensed from an academic or research institution (3). In either situation an attractive way to reap the full benefits of developing the product and to leverage research and development and production costs is promotion to a national market (28,32). For example, the fee for a single breast cancer genetic test is $2000, for the two genes involved in here ditary nonpolyposis colorectal cancer, $870. The international competition to identify the first breast cancer gene was probably the most publicized scientific "race" of the 1990s. And when the winners — Myriad Genetics, a Utah-based biotechnology company, working with the University of Utah — promptly sought a very broad patent over BRCA1 in 1994, there was considerable disquiet. Opposition came not only from "genetic interest groups," arguing that genes are natural human blueprints which should not be patented, but also from other scientists who had co-operated with the Utah group during earlier stages of the research (31).

This commercial model for test development, however, is of fairly recent origin (26). Most genetic testing still takes place in research and academic settings (33). In the survey of genetic testing discussed above, twice as many nonprofit organizations were engaged in genetic testing activities as biotechnology companies (3). Perhaps more significantly, most of the basic research underlying the science of testing continues to be generated in public and nonprofit institutions and universities through sequencing under the Human Genome Project (HGP) and investigational studies of particular disorders (34–37). A comparison of recent headlines is instructive. Several research centers organized by the National Institute for Alcohol Abuse have found links between specific chromosomal areas and a propensity for alcohol abuse (34). The development of private sector initiatives is part of a general burgeoning of genetic testing research and practice rather than representative of a transition from public to private leadership. Symbolic of the growth of both research and commercial endeavors, a new publication known as the *Journal of Genetic Testing* was recently introduced (38).

What is a Genetic Test?

The landscape for the policy debate on genetic testing takes place against a complex and rapidly evolving scientific background of genetic testing technology. The details and methods used in genetics are complicated, but the general concept of a genetic test is fairly simple. By now the image of DNA as a double helix should be quite familiar (20). DNA stands for deoxyribose nucleic acid. The deoxyribose, in conjunction with a phosphate group, alternately link together to form the backbone of the helix similar to the sides of a ladder. The nucleic acid, of which there are four kinds, projects off the deoxyribose molecule to form the step of the ladder known as the base. Each step in the winding ladder of the helix is called a base-pair, and a person's genetic code contains billions of base-pairs (20). Imagine if the two sides of the ladder were pulled apart, splitting all the steps in half (20). (More specifically 23 human chromosomes contain

about 3 billion base-pairs. Each side of the ladder is a reverse of the other because of the base-pairing.) Each half of a base-pair is called a *nucleotide*. The sequence of nucleotides provides specific directions for the production of all the parts of a living organism. Originally a *gene* was defined as a set of nucleotides that codes for a protein (20), although the term now commonly is used to identify sets of nucleotides that are associated with more general traits. Although variation in human genes is as natural as variation in fingerprints, there are specific variations in some genes that cause or are linked to a genetic disease (20). For example, the common inherited disorder of cystic fibrosis results from mutations or deletions in a transmembrane protein known as cystic fibrosis transmembrane conductance regulator. Thus, locating and cloning a disease gene can be important in determining the molecular pathology of inherited disorders. A genetic test determines whether an individual has a harmful genetic variation or any other specified genetic sequence. At present directly sequencing a gene (20) is a method reserved primarily for finding the gene originally. It is uncommon for a genetic test to simply sequence each person's gene at the relevant location because sequencing has until recently been a time-consuming and expensive process. Increasingly powerful engines for the sequencing of DNA, however, make the prospects of such an approach not only possible, but likely in the future. For example, the Perkin-Elmer's 3700 DNA analyzer, which gained so much notoriety as the engine behind a bid to take HGP private, can analyze thousands of base-pairs a day and sells for $300,000 (39–42).

Causation in Genetic Diseases. There are several relatively high incidence single locus gene disorders such as sickle cell anemia, thalassemia, and cystic fibrosis where genetic diagnosis is now common (4). Of course, identifying a disease-related genetic variation is not always conclusive, nor is failing to identify one. There may be an extremely large number of disease-inducing variations amid many benign ones, requiring extensive clinical follow-up and research. Expression of a gene or genetic predisposition is subject to complex laws of inheritance and the influence of environmental factors (3,29). In addition common disorders such as heart and Alzheimer's diseases are usually multifactorial, which greatly increases the complexity of analysis (3,20,29). Tracing the relative significance of variable genetic and environmental factors is extremely troublesome from a research perspective and impedes conclusive determinations of risk and probability. The complexity of the underlying genetics requires flexibility and sensitivity to situation-specific information for effective genetic testing generally. Indeed, the pace of innovation subjects testing practices to continual change. These conditions, and the technical difficulty level involved, encourages testing to be performed as a service rather than packaged as kits at least in the near and medium term (19).

Future Technology. In the relatively near future it may become possible to screen a larger number of patients more routinely for a larger number of genetic diseases, and to do so for diseases with more allelic variation. For example, the "gene chip," a device with space for up to 409,000 probes, has been developed (44). This new technology is in the final stages of gaining approval for its first commercial use detecting mutations in the P53 gene, which may lead to development of certain cancers. This mutation usually is acquired during the course of a person's life, however, and there currently is no chip-based test commercially available for inherited diseases (44). There is no reason, however, that this technology could not be used to test more cheaply and rapidly for a large number of genetic diseases at once. Additionally the DNA chip may make it more feasible to test for diseases with a large number of different disease alleles such as cystic fibrosis (a three base-pair deletion detectable by a single probe accounts for 68 percent of cystic fibrosis mutations, but the other 30 percent arise from over 60 different kinds of mutations (20,45). Although there are no commercial genetic tests for inherited disorders based on this technology, the chip may be a view into the future of such testing. Given the chip's probable status as a device, it will be a future likely overseen by FDA.

Potential Harms in Genetic Testing

The technical aspects behind genetic testing indicate that use of these technologies may involve considerable uncertainty. Tests for multifactorial inheritance disorders such as Alzheimer's and breast cancer have generated the most concern because the clinical utility of a positive test in a healthy individual is often still unclear at the time the tests become commercially available. For example, a panel of experts has recently concluded that testing for ApoE4, the Alzheimer's "susceptibility" gene, does not provide sufficient predictive value to justify its use outside of research labs (46). While the test improves the clinical diagnosis of symptomatic patients, it is not appropriate for testing at large. These ambiguities and uncertainties, however, probably are not significantly greater than in other medically complex areas of diagnosis. What distinguishes genetic diagnosis from other kinds of complex or controversial forms of testing? In truth, there are many similarities, especially from a medical perspective. Misdiagnosis could lead to unnecessary, painful, and risky treatments (4), and the failure to diagnose might discourage diligence in taking care of the body and seeking medical attention. Imperfect tests in general create difficult choices, suffering, and uncertainty. In fact, when used to diagnose a symptomatic individual, there are probably few differences between genetic tests and other methods of assessing disease. Predictive genetic testing, however, is performed on healthy individuals and even on fetuses in utero. The tests frequently offer only estimates and probabilities of risk, and may become available before the validity and reliability of its predictive value have been established. The genetic diseases for which tests are being developed are generally more dangerous and severe within the range of ailments (2,30), and unfortunately there are rarely effective treatments, much less cures or means of prevention (26). The fundamental and basic nature of genetic abnormality changes the entire dynamic of disease. One's genetic

heritage, unlike a virus, bacterium, or a mutant cell, simply cannot be destroyed or abated (39,47). Genetic conditions, in a sense, are inseparable from the existence and identity of a person, and they can confer potential disease status upon someone with perfect health and no symptoms. The medical realities of genetic diseases form a background for a much broader, and sometimes more troubling, range of social and psychological concerns. The basic character of genetic information brings it to bear on fundamental life choices and understandings in a way that raises unique possibilities of harm to the individual, and perhaps, even to society.

Inherited Breast Cancer. The paradigmatic case and original source of much of the controversy and concern over the commercialization of genetic testing was the discovery and marketing of the BRCA breast cancer gene tests (31). The case provides a good illustration of the issues involved in the commercialization of genetic testing. Normal or wild-type BRCA genes apparently inhibit the growth of tumor cells in breast and ovarian tissue (48). Mutations in the genes may confer a lifetime risk of 80 percent for developing breast cancer and 50 percent for developing ovarian cancer (48). There are many deletions and variations, however, that may appear in a woman's BRCA gene of "unknown significance" (49), and in fact the high correlations to risk cited above are based on specific mutations found in women that were previously determined to be high risk (48,50). The possibility of selection bias compromises the probabilities (51–53). It has been speculated that the risk posed by the BRCA mutation for women with no family history of breast cancer might only be 40 percent or lower (54). Amid this background of uncertainty, and with the drastic measure of a prophylactic mastectomy as the only potentially preventative option, several companies moved to provide BRCA tests commercially (55). The Genetics and IVF. In vitro Fertilization Institute in Fairfax, Virginia, began offering the test to the general public on the theory that women had a right to know about their genetic status (56). Myriad Genetics, in its promotional literature to physicians, suggested that almost everyone could benefit from taking the test (57). Since then a University of Washington study has concluded that women without a strong family history of breast cancer should not worry about getting tested (58). Another study at Leiden University in the Netherlands found that the polymerase chain reaction (PCR) process used in the gene tests was missing about a third of the disease-causing mutations because they were too large (59).

At the center of the confusion and controversy, women were making private, painful decisions about whether to undergo testing despite few medical remedies except for surgically removing both breasts. Parents had to decide for daughters and weigh the benefits of alertness in life to symptoms of the disease against the possibility of ruining her healthy years with depression and apprehension. A woman told that she has a significantly increased risk would have to decide how to react to news reports that having no children or having them late increases the risk of breast cancer (60) and that smoking seems to reduce the risk (61). She would have to contemplate the possibility of passing on the gene to her children and the guilt that might entail. Some of the healthy women who tested positive chose the difficult course of prophylactic double mastectomies and oophorectomies (removal of the ovaries), although breast cancer could still occur in remaining tissue or even in the abdomen or colon (62). Questions remain whether the estimated 2.5 million other American women with BRCA mutations should undergo similar procedures (62). Today new options such as tamoxifen offer women alternative means of prevention and hold promise for future developments (51–53,63). An at-risk woman, however, would still have the burden of figuring out how reducing the risks posed by her particular mutation compares to side effects like increasing the risk of uterine cancer and life-threatening blood clots (64). The result is a biological version of Russian roulette that many women may simply avoid by not playing at all (65).

Dangerous Information. The discussion of breast cancer gives some brief indication of the hard choices and painful consequences that genetic testing can create. Only a superficial sense of these issues can be conveyed here, but it is important to realize the scope of the effects of genetic information. Perhaps the most obvious effect, as well as the most difficult to describe, is that testing positive may be "psychologically devastating" (26). It can cause deep depression or anxiety, and disrupt family relations. Even cases of suicide of otherwise healthy individuals have been reported (4,66). The problems seem most acute with respect to incurable late on-set disorders, such as Huntington's disease, where the utility of testing in light of the lack of treatment has been the subject of considerable controversy (67). Genetic information also substantially affects important life decisions such as whether to have children and even whom to marry (68). Such information may even define personal identity and status within one's community (69).

The broader social implications of genetic testing also are profound. Genetic testing already has begun to enable genetic discrimination by employers and insurers (70,71). The government has started creating databases for the DNA of employees, all military personnel, and criminals (70,71). Prenatal diagnosis provides a means of extending subtle social discriminations to selective termination. In California, the only state keeping track of prenatal genetic testing, 70 percent of women receive prenatal diagnosis (72). It is estimated that at least half the pregnant women in the country do so, and that about half the women who receive a positive test for a serious illness choose selective termination (72). There is little room to be sanguine in these matters. Consider, for example, the recent research indicating the genetic basis of Lou Gehrig's disease (73), and imagine that parents began terminating births of afflicted fetuses. An advanced society might think that the private activity of parents has resulted in a net benefit, but in fact, could the world have done without the likes of Stephen Hawking, or for that matter, Lou Gehrig (74)? An Institute of Medicine study commented:

> The development and widespread use of genetic tests ... raises issues about discrimination and privacy ... that

> people found to possess certain genetic characteristics will lose opportunities for employment, insurance and education. . .. [G]enetic testing raises worries about inequities and intolerance . . . that not everyone will share equitably in the benefits of genetic testing, that some will be stigmatized, and that the beauty of human diversity will be denigrated due to a narrowed definition of what is acceptable (4, p. 30).

In Germany genetic testing already is being used in immigration policies (75), and in the United States suits for "wrongful birth" have been brought against doctors who botched genetic tests indicating disability (76).

The stakes in genetic testing are not only the traumatic personal ones discussed above, but include far-ranging social ones involving democracy, privacy, and delicate notions of human worth. As one commentator has noted, "Although their proponents invariably proclaim that new technologies will bring unprecedented prosperity and freedom, they can also threaten our civic values. What Thomas Jefferson called 'cherished liberty' is not determined by our genes. It is determined by our eternal vigilance" (70). The difficult private and public decisions that genetic information requires, about the power of genetic selection, eugenics, and so on, seem impossible to render. Before individuals and society are asked to go where angels fear to tread (77–80); however, it seems reasonable to ensure at least that the momentous consequences provoked will flow from information sufficient and reliable enough to make a basic scientific determination. It is not easy to advise a parent whether to terminate selectively a fetus that has an 80 percent (or perhaps only 40 percent) chance of developing breast cancer (72) or for society to decide whether to intervene. In fact the only easy thing to say about these troubling and complex issues is that people should not be asked to resolve them based on information with no established clinical validity. It is perhaps the only easy decision to make in minimizing the harms that can result from genetic information. The question remains, however, over who should be responsible for doing so.

HISTORY AND JURISDICTION OF THE FOOD AND DRUG ADMINISTRATION

Though FDA regulates genetic testing kits, it has historically had a policy of not regulating clinical services (82). Before turning to the original reasons behind adoption of this policy, it might be appropriate to explain why the option of restraint might have particular currency with respect to genetic testing. The political climate in which FDA now operates greatly heightens the risks and barriers to undertaking new forms of regulatory oversight. Since 1994 there has been a remarkable shift in public, and especially congressional, attitudes toward FDA. While critical campaigns in the past traditionally centered on the agency's failure to protect sufficiently the public health or to meet fully its statutory obligations with respect to imposing regulations and review (83,84) recent criticisms have focused on excessive bureaucracy and oversight that prevented medical benefits from reaching the public with adequate speed (85,86). Comparisons were made to the availability of new products in European markets (87), and, particularly in the case of HIV treatments, the social cost of standards and requirements for testing and proof of efficacy became the basis for public outrage and demonstrations (88). A newly elected Republican Congress, sensing an opening, concentrated much of its energy on FDA in its campaign to deregulate American business (89–92). The fierce cost-cutting budgetary environment put further restrictive pressure on the scope of the agency's regulatory mandate.

The FDA also has not had positive experiences taking strong regulatory positions in the testing area. Its last few attempts to protect the public from too much self-knowledge proved to be an institutional bellyache. In the late 1980s FDA adopted a complete ban on HIV home testing kits, which came under considerable derision (93–97). Accusations of paternalism were leveled and public health experts argued that without encouraging testing of populations not apparently at-risk, efforts to control spread of the disease would not be effective (98). FDA retreated from its position in 1995 (99). Similar efforts to block the distribution of drug-testing kits also suffered considerable unpopularity (100).

Finally, the agency's lack of enthusiasm may stem from the intuition that regulation of genetic testing is a thankless task. The extraordinary volume of scholarly writings on genetic testing exploring the social and moral implications often include critiques of potential government policies, which writers prophesize will lack nuance and sensitivity to the exploding issues related to genetic testing (101–105). Almost any guidelines that can be produced by the limited resources available will prompt passionate objections and painstaking critiques. While physicians, scientists, consumer advocates and other thoughtful individuals repeatedly declare that some effort to establish validity should be undertaken (2,8,11,12), the general public may still question attempts to limit their access to information, however uncertain the information may be. People might prefer to judge for themselves.

In addition to these present considerations, the policies and possibilities of FDA regulation of genetic testing are shaped by its institutional and jurisdictional history. Its range of administration is formulated according to a legislatively constructed and evolving societal role that is defined by specific purposes and by its position in relation to other guardians of the public health.

Historical Account of the FDA's Mandate

FDA's creation and evolution have resulted largely from successive periods of public crisis (106). Its charter began at the turn of the century when industrialization created an urban workforce dependent on produce and packaged goods transported from rural areas (107). The market proved to be a poor regulator of the quality and content of these goods, and the horrors of mislabeled and adulterated foods became the subject of public outcries and criticism (108,109). For example, food labeled "potted chicken" or "potted turkey" in North Dakota was found by state government officials to contain no discernible amounts of chicken or turkey (108). The first food and drug law was passed in 1906 and was aimed solely at punishing transportation of food and drugs that were adulterated

or misbranded (110). There was no provision for direct regulation of safety and efficacy.

Congress began to entertain the possibility of more expansive regulation during the New Deal, but could not come to an agreement on the issue for many years (107). From the outset there was concern about administrative overreaching into the discretion and business decisions of the industry and fear that regulation would hinder growth (107). The trade-off between preservation of the public health and economic and technological development was immediately apparent. While Congress debated, in 1937 the elixir sulfanilamide disaster struck (111). The product, which had been tested for appearance and taste but not for safety, caused the deaths of almost 100 people. The only federal statutory violation was labeling the product as an elixir, which technically can apply only to alcohol solutions (111). The product was actually a sulfa compound dissolved in diethylene glycol with unfortunately fatal side effects. The manufacturer was not required to test for safety, or even to disclose the fatal ingredients. It was a tragic introduction to the inadequacies of existing legislation, and within a year the basis of modern food and drug law, the Federal Food, Drug, and Cosmetic Act of 1938 (FDCA), was passed (107,112).

In 1962, following the development of powerful new pharmaceutical products, regulation for effectiveness was added to the goals of FDA (107). The amendment was an important step in blending the work of the medical profession with the obligations of federal regulators. While the thrust of the laws continued to be the protection of individual consumers of mass-produced and marketed products, comprehensive regulation of all drugs (not merely for whether people could be harmed but for whether they would be helped) represented a major inroad on medical discretion. Extending FDA guardianship over both the safety and effectiveness of medical devices through the Medical Device Amendments of 1976 continued the assault (108). Congress was aware of the overlap, and specifically provided that effectiveness need only be reasonably substantiated (113). The statute made clear that no interference in medical practice, particularly with respect to the physician–patient relationship, was intended (114). The reluctance to regulate clinical services derives from this historical deference to medical practice.

With each successive wave of technological advancement in various health-related fields, the agency's purview has expanded and adapted to encompass novel issues and risks. Despite political attacks and criticisms of FDA, it will rarely be heard that the agency should be abolished. FDA's necessity has been demonstrated by experience, and despite complaints about the bureaucracy there is a certain efficiency to its existence. By having one national gatekeeper to monitor for potentially harmful drugs or devices, the independent review necessary to secure the public health takes place only once and is rendered in a methodical and dependable manner rather than duplicatively in offices and hospitals with inconsistent results depending on variations in time, access to information, and interest. The public benefits from consistency, standardization of review, avoidance of redundancy, and accumulation of institutional expertise generated by the existence of a central monitoring agency.

Boundaries of FDA Jurisdiction

While a fairly uncontroversial argument can be made that the agency has some authority to regulate the activities of services under the Medical Device Amendments to the FDCA (116), it seems equally clear that its discretion to do so is not unlimited. The question of where to draw the line is a matter of some novelty and ambiguity. There are in fact two separate but related inquiries involved: first, whether the regulation of an apparently intrastate service falls within the constitutional boundaries of federal power, and, if so, whether such power has been delegated by Congress to FDA. The extent to which a historical deference to the medical profession may bear upon proposed efforts to intervene in auxiliaries to clinical practice also must be considered.

Interstate Commerce Clause. FDCA derives its authority ultimately from the interstate commerce clause of the U.S. Constitution (117), which grants Congress the power, "[t]o regulate Commerce with foreign Nations, and among the several States, and with the Indian Tribes" (118). The significance of the clause has undergone considerable evolution as the legitimacy of national power has grown, the most memorable point of departure taking place during the New Deal. During the 1930s President Franklin D. Roosevelt oversaw enactment of a series of laws designed to help the country out of the Depression including the National Industrial Recovery Act. The Supreme Court invalidated the Act and other efforts as unconstitutional in several famous cases including *ALA. Shechter Poultry Corp. v. United States*, 295 U.S. 495 (1935) and *Carter v. Carter Coal Co.*, 298 U.S. 238 (1936). Roosevelt's response was a proposal to change the structure of the court to include one additional justice for each justice over the age of 70 who did not retire. In 1937 there were six such justices. Several key opinions on the Court changed in what has come to be known as "the switch in time that saved Nine." After that time the New Deal vision of national power was vindicated completely, and almost no judicial restraints on the federal commerce power seemed to remain:

> By 1945 the supreme court had come to the position that the primary and perhaps exclusive federalism-based constraints on Congress were imposed by the political process. Although Congress's regulation of the national economy has continued to grow, nearly all of its work falls well within the boundaries set by cases such as *Wickard* and *Darby* (119).

Present jurisprudence interprets the commerce power to permit Congress (*1*) to regulate the use of the channels of interstate commerce, (*2*) to regulate and protect the instrumentalities of interstate commerce, or persons and things in interstate commerce, even though the threat may come only from intrastate activities, and (*3*) to regulate those activities having a substantial relation to interstate commerce (120). FDCA already has been sustained as a constitutional exercise of the interstate commerce power (121–123). In determining the permissible extent of federal regulation of genetic testing, however, the constitutional question warrants preliminary consideration, because under the doctrine of "constitutional doubt"

statutes are construed, if fairly possible, so as to avoid raising a serious question as to its constitutionality (124).

The substantial relation to commerce theory is probably the best basis for regulating genetic testing services. A classic case illustrative of the generous boundaries of that test is *Wickard v. Filburn* (125), the authority of which recently was reiterated by the Supreme Court (120). *Wickard* sustained an implementation of the Agricultural Adjustment Act of 1938 that imposed a wheat quota on a farmer whose wheat was intended entirely for consumption on the farm (125). Not only was the wheat not intended for marketing or distribution out of the state, it was not intended for commercial sale at all. Determining that even the home consumption of wheat bore on the demand for it in the national market, the Court concluded, "This record leaves us in no doubt that Congress may properly have considered that wheat consumed on the farm where grown if wholly outside the scheme of regulation would have a substantial effect in defeating and obstructing its purpose to stimulate trade therein at increased prices" (125). The Court's holding, however, was premised on a specific finding that home consumption substantially affected the economics of the wheat industry. In *United States v. Lopez* the Supreme Court rejected the claim that bearing weapons near schools substantially affected interstate commerce" (120). There clearly are limits to the effects that will support jurisdiction. With respect to commercial genetic testing where samples or patients cross state lines to obtain service for fees or where the company itself has clinics in several states, regulation is clearly well within the bounds of Congress' constitutional power. The activity clearly bears on commerce that is interstate in nature. Even if a test is "home brewed" entirely in state and performed on local residents, it would probably fall within the rule enunciated in *Wickard* and subsequent cases that intrastate activity affecting interstate commerce may be regulated. The Supreme Court repeatedly has sustained regulation of intrastate activity that bears a substantial relation to interstate commerce (120). A more troublesome question arises with respect to programs offered by academic and research centers that are nonprofit. The findings of the Task Force suggest that despite the volume of testing in which such institutions engage, they do not compete significantly with commercial producers and providers (2). If the programs attract patients from out of state then regulation might be supported under the second prong of the commerce power, namely protecting persons in interstate commerce (120). Regulation of solely regional test centers and nonprofit organizations therefore present increasingly marginal constitutional cases for regulation.

Federal Food, Drug, and Cosmetic Act. The more substantial inquiry focuses on the breadth of Congress' delegation of regulatory power to the FDA as it bears on genetic testing. The relevant operating provision in the FDCA is codified at 21 U.S.C. §331, which prohibits:

> (a) The introduction or delivery for introduction into interstate commerce of any food, drug, device, or cosmetic that is adulterated or misbranded.
>
> (b) The adulteration or misbranding of any food, drug, device, or cosmetic in interstate commerce.
>
> (c) The receipt in interstate commerce of any food, drug, device, or cosmetic that is adulterated or misbranded, and the delivery or proffered delivery thereof for pay or otherwise. . . .
>
> (k) The alteration, mutilation, destruction, obliteration, or removal of the whole or any part of the labeling of, or the doing of any other act with respect to, a food, drug, device, or cosmetic, if such act is done while such article is held for sale (whether or not the first sale) after shipment in interstate commerce and results in such article being adulterated or misbranded.

Pursuant to its authority to regulate devices in this section and under other MDA provisions, FDA has asserted repeatedly that it has the power to regulate genetic testing services (126). To support its authority, FDA must identify both a device and the presence of interstate commerce. The language of FDCA is therefore more limited than the constitutional boundaries, regulating activity in interstate commerce rather than activity merely bearing a substantial relation to interstate commerce. The agency has several alternatives to establish jurisdiction: It may premise jurisdiction either upon the materials shipped in interstate commerce from which the genetic test is assembled, on the genetic test itself, or some blend of the two.

Materials Received in Interstate Commerce. FDA's regulation of analyte-specific reagents is an example of regulating ingredients or materials received in interstate commerce (116). Extending regulation to the entire genetic test assembled after the components are shipped in interstate commerce, however, adds a twist. Section 331(k) of title 21 of the *U.S. Code* prohibits adulteration or misbranding of a "device" after shipment in interstate commerce, and the statutory definition of "device" includes "any component, part, or accessory . . . intended for use in the diagnosis of disease or other conditions" (127) clearly encompassing the materials used to assemble a test. A logical and persuasive argument can be made that the components previously shipped in interstate commerce can be "adulterated" within the meaning of Section 331(k) by assembling them into a genetic test that does not conform to regulations prescribed by FDA. A similar reasoning was adopted by the U.S. Court of Appeals for the Ninth Circuit in *Baker v. United States* with respect to the sale of misbranded drugs (128). The court held that Section 331(k) applied where the ingredients were shipped in interstate commerce even though the final manufacture and sale of the drug took place in California (128). The U.S. Court of Appeals for the Eighth Circuit adopted a similar approach to extend regulation to animal biologics that were only distributed intrastate (129). In its argument for jurisdiction under this theory the FDA would be assisted by the long-standing principle in food and drug law jurisprudence that the "high purpose of the Act to protect consumers who under present conditions are largely unable to protect themselves . . . [should not] be easily defeated" (130). Ensuring the public health is a prime example of national power and prerogative, especially where the circumstances are buttressed by other

commercial and interstate elements. The weakness in the approach is that a testing service could theoretically evade regulation by assembling the test from materials which never crossed state lines, even if the service openly and aggressively engaged in national consumer-oriented marketing and served residents from outside the state.

Commercial Testing as Interstate Commerce. Alternatively, jurisdiction may be premised on the genetic test itself if it can be deemed to be an object in interstate commerce. There is some question whether a product technically manufactured and consumed locally may nevertheless, in light of the surrounding circumstances, be an object in interstate commerce. In the past courts have found interstate commerce where the consumer traveled from another state (131) or where information crossed state lines (132), both of which often are involved in providing commercial testing services by test producers. There appear to be no cases, however, where jurisdiction was premised on a device that did not itself at some point cross state lines (133). By providing testing as a service, companies inadvertently have found a means for circumventing the regulation of diagnostics. It is an exception not limited to genetic testing, and generates some disturbing dynamics. Beyond the lack of FDA regulation, it also makes manufacturers no longer dependent on the approval of the medical establishment and allows them to advertise and sell directly to the consumer. Consider, for example, the recent controversy over an Alzheimer's test marketed as a service by Nymox Pharmaceutical Corporation (134). The company ran an advertisement in *TV Guide* implying a conclusive test despite considerable doubt as to the scientific validity of the $400 test (134). A nationally advertised single-test "service," such as the Nymox Alzheimer's test, bears a rather suspicious resemblance to a nationally distributed product, the very object of historical FDA regulation. Adopting the theory that interstate commerce could be created by out-of-state advertising and servicing would require some revision of existing interpretations (135). Courts might be willing to view the test as taking place in interstate commerce under Section 331(a) or (b) if the regulation was restricted specifically to commercial providers who marketed their tests outside the state. Regional services that advertised only to local medical practitioners would provide little basis for regulation. The more interstate elements that are added in defining the scope of those regulated, the more compelling the case for applying Section 331(a) or (b). Using Section 331(a) or (b) seems to better reflect the substance of the activities' interstate nature. However, premising jurisdiction on the test components shipped in interstate commerce and proceeding under Section 331(k) seems a more certain approach under present case law.

Unwarranted Interference with the Practice of Medicine. Although there are dependable grounds for asserting statutory jurisdiction, the FDA must be wary of encroaching on medical practice. While many of the FDA's activities constrain the practice of medicine, such as regulating drugs and screening medical devices (136), there are limits to the permissibility of interference, perhaps even of a constitutional nature. In a very early opinion the Supreme Court stated, "Obviously, direct control of medical practice in the States is beyond the power of the federal government. . ." (139). Although it should be noted that this statement was made prior to the changes in the federal commerce power wrought by the New Deal (119), there is probably still some active residue of the principle remaining today. In *United States v. Evers*, the FDA tried to regulate a physician's use of a chelation drug in treating arteriosclerosis (138). The district court concluded that it was an unwarranted agency intrusion, stating:

> The courts have rather uniformly recognized the patients' rights to receive medical care in accordance with their licensed physician's best judgment and the physician's rights to administer it as it may be derived therefrom (138).

On appeal the U.S. Court of Appeals for the Fifth Circuit declined to endorse this reasoning and affirmed on entirely different grounds (139). The situation involved in *Evers* is quite different than the setting of genetic testing involving commercial labs and regulation of substantive changes in the device rather than merely regulating its medical use. In addition *Evers* reflects somewhat dated, paternalistic notions of the privileges of the medical profession that may have little currency in the modern era of patient's rights. Despite doubts about its authority, however, the reasoning of *Evers* strongly suggests that some care should be taken to avoid encroaching on the practice of medicine in fashioning regulation for genetic testing (140).

THE CASE FOR FDA REGULATION

FDCA arose partly out of a recognition that drug companies and medical products manufacturers, although they assist doctors in saving lives and relieving suffering, are not parties to the Hippocratic Oath. Their motive is profit. Regulation represents a determination that the dangers posed to public health are not sufficiently accounted for in the risk calculus of companies when they introduce products. As revealed in the historical discussion of FDA, the nation learned that lesson through painful experience and since that time has attempted to anticipate potential problems by delegating broad authority that can evolve to cover novel technologies (141,142). There seems to be little disagreement that some authority should determine that genetic tests being offered as services have positive predictive value for the disorder being tested (9,12). In tension with this simple purpose is a broad and varied range of settings in which genetic testing takes place. Any DNA analysis performed at the request of a physician or researcher, whether in a chemical or hospital laboratory including individualized linkage analyses used for diagnosis, can constitute a genetic test. Over 500 laboratories in universities, hospitals, public health departments, and commercial centers perform analysis utilizing innumerable target genetic variations and rapidly evolving test strategies (143). If that is the objective of regulation, it is quite apparent that FDA is not exceptionally suited to such a task. It seems, however, that the dangers generating unique concerns do not arise

in all settings where genetic testing services are provided. Some testing takes place in a research environment or as an extension of medical practice similar to other forms of laboratory analytical services where there has been no general call for regulation. Describing these settings is a helpful prelude to distinguishing precisely the objective of regulating genetic tests. When the goal is carefully identified, the strengths and expertise necessary to its implementation will also become apparent.

Information Dynamics in Genetic Testing

Genetic testing providers may engage in a wide range of activities: researching genetic linkages (144), developing new diagnostic products (28), performing genetic analysis, and serving people directly through clinics (145). To clarify the information dynamics and make the realities of the industry's structure simpler and more immediate, the discussion proceeds by tracing the progress of a hypothetical consumer of genetic testing services through the process. An individual may obtain a genetic test either as part of a clinical research program, through a physician, or directly from a commercial testing service. The fourth option of purchasing a home test kit is not currently available in the United States, although an over-the-counter genetic test for cystic fibrosis was recently introduced in the United Kingdom (146–148).

Patient Participating in Research. Individuals asked to participate in research programs often already have developed the condition or are related to someone who has the condition (4,9). The linkage studies used to identify disease genes are most productive using a familial database and comparing the genes of those who have the disorder with those who do not have it (144). The researchers are necessarily genetic specialists and academics who must comply with human subject requirements and other investigative protocols (9). The potential harms of poor information in this scenario are minimal since the individuals tested are at-risk to begin with, the information is often accompanied by education and counseling (26), and the results are delivered by professionals with the specialized knowledge necessary to explain the significance and experimental nature of the information and who have motive other than commercial profit. Sometimes the results are not even given to the test recipient (4). In addition to taking place in a "protected" setting (26), testing for research purposes yields important social benefits in increased scientific understanding of the human body and produces precisely the kind of information necessary to make medical determinations more reliable.

Requesting Testing through a Physician. There are basically three purposes for which genetic testing might be ordered through a physician. First, the physician might be using it to further the diagnosis and treatment of a symptomatic individual, which in the case of genetic disorders is usually a child (a related case involves presymptomatic testing of a close relative). Second, an expectant mother might seek prenatal diagnosis, or her obstetrician might recommend it based on a family history of previous birth defects. Finally, a physician may recommend it to an individual with a general family history of a condition, or the individual might raise the possibility on his or her own after having heard of the test on the news or even through advertisements.

Diagnosis of a Symptomatic Individual. As discussed above, when genetic tests are used to make or improve diagnosis of a symptomatic individual, the dangers implicated are not significantly different from those occasioned in other types of testing. Because the test recipient presently is endangered by illness, any medically useful information may facilitate treatment and prevent imminent physical harm. The ordering physician, furthermore, is usually a specialist in the particular disorder being tested for and has a special capacity to determine the presence of a valid indication for testing (149) and evaluate the medical implications of a positive result. The case will involve a patient under continuing care whose medical and family history will have been explored in conjunction with the disorder. For family disorders, close relatives who are tested will be subject to similar conditions. Although a high rate of misinterpretation may persist even in these circumstances because of the complexity of genetics (149), these tests generally offer a greater degree of medical certainty as a starting point (149) and physicians who work with genetic disorders are more likely to be able to understand and evaluate genetic information and research (2).

Prenatal Diagnosis. Prenatal diagnosis involves somewhat special considerations, and also blends aspects of the other two cases. In a survey of physicians, obstetricians earned some of the highest marks for genetic knowledge (2). Obstetricians' work exposes them to more genetics, and they are one of the main users of clinical genetic services, including genetic counseling (29). Furthermore, until recently, most prenatal diagnosis utilized the same tests employed by pediatricians and specialists to diagnose symptomatic individuals. To have the test performed, the obstetrician may send a sample to one of the national commercial laboratory services, a regional service or the laboratory at a local hospital, perhaps one with which she is affiliated. Because of the risks involved in prenatal diagnosis (48), the tests remain conservative and focus largely on severe, predominantly single-gene disorders with well-established genetic tests. Commercialization of tests like the one for breast cancer susceptibility, however, may not leave prenatal diagnosis immune. An article noted that "some couples already are asking for [testing] the genes that confer a fifty to eighty five percent breast cancer risk" (72). The Genetics and IVF Institute has responded that they might even do it "after counseling and careful consideration" (72).

General Physician Referral. Primary care providers are likely to be on the frontlines of both demands for genetic tests by patients (150) and advertising by biotechnology companies and national laboratories who develop tests. As one journalist noted, "It's almost a daily occurrence to pick up the morning newspaper and read that scientists have found another gene that causes a medical malady" (26). Furthermore, "Once we find the gene, the stories imply, a genetic cure may be just around the corner"

(26). Most Americans are confident in the ability of their primary care physician to tell them if they are at risk and believe that the family practitioner can correctly interpret a genetic test (151). Unless their doctor has had some previous experience with genetic testing, however, their reliance may be misplaced (150). A 1993 study found that primary care physicians earned low scores on average in a test of genetic knowledge, although there was marked improvement for more recent years of graduation (152). Unfamiliarity with the essentials of genetics makes the primary care provider a somewhat unlikely candidate for ensuring independent review of the clinical utility of a new genetic test. In fact, the media and even advertising may take the decision out of the hands of primary care providers as consumers are persuaded to pursue the test based on these extrinsic sources, and physicians feel pressured not to interfere in the patient's decision making (149).

At present there are only a handful of companies developing new, advanced tests (2). Most are young, aggressive companies "hoping to take the lead in a potentially lucrative field, unconstrained by the complex FDA regulations that beset drug approvals (28,153,154). The rest of the industry is comprised of major diagnostic companies, such as Boehringer Mannheim, Ciba-Geigy, and Johnson & Johnson, and national behemoths in genetic testing services, such as Genzyme, Lab-Corp., Abbot Laboratories, and SmithKline-Beecham, which have been slower to commit resources (153). The existence of the small companies revolves around their venture capital structure, which requires quick turnarounds on profit and creates intense pressure to get a product quickly to market, perhaps even before the test's safety and effectiveness can be established (153). The smaller regional commercial testing centers, such as Clinical Diagnostic Services serving the New York metro area and Laboratories for Genetic Services in Texas, traditionally have offered a more conservative array of genetic analysis but seem to have been expanding in recent years (155). Their advertising is almost exclusively local and is directed at physicians rather than the broader public because of expense considerations. Laboratories affiliated with universities and hospitals generally offer a smaller range of testing, specialize in particular kinds of tests, and rarely advertise their services.

Getting a Test on One's Own. The last consumer is one who has not been invited to be part of a research study and has not been referred for a test by a physician. Perhaps, the individual does not wish to divulge family disorders for fear of insurance discrimination, or he or she is motivated more by curiosity and apprehension than medical necessity. The laboratories at universities, hospitals, and regional commercial services are largely unavailable without a physician referral (155). A few testing services that cater to a nationwide audience, however, have given consumers direct access to genetic testing. A prominent and controversial example is the Genetics and IVF Institute, which permits walk-in testing for BRCA mutations (145). Third-party screening of the test's validity in this scenario obviously would be nonexistent.

Need for Regulation. The need for objective information on the soundness of tests being purveyed is greatest with respect to commercial services marketing tests directly to the public. In these instances, a genetic testing service bears the greatest resemblance to a conventional commercial product. Prior to commercialization of genetic testing, there were established, informal networks for obtaining a test through specialists and other experienced providers. If a test was well-proven, knowledge about the test filtered through the professional establishment to the appropriate physicians. Now, producers promote novel tests directly to all physicians and the public (153). The consuming physician and patient need not have any independent information to obtain access. The problems of misinformation and defective information seem to be most critical where the elements of commercial interest, puffery, ready access by the public, and lack of independent review converge. In fact, the essence of these elements lies in the flow of information, that is, who has it and who depends on it. A negative change in the source of information away from the supervising physician to a financially self-interested company or service may be summarized in one word: marketing. Relying for information on the very parties who have the most to benefit from praising the test, however, obviously leaves a great deal to be desired.

Current State of Regulation

However, the existing framework of regulation focuses not on where the greatest harm may exist, but on whether a test is packaged or performed on site. Providers and developers of genetic testing are currently subject to a patchwork of federal and state regulation based on the form of the activity engaged in rather than the substance. Biotechnology companies who package the test as a kit to be distributed to laboratories and physicians are subject to FDA regulation. Companies and laboratories that deliver the testing as a service, sometimes by accepting samples through the mail, are not subject to any regulations to ensure safety and effectiveness. Instead, only the technical competence of a laboratory's testing is regulated under the Clinical Laboratory Improvement Act. Laboratories also are licensed by the state and voluntarily may undertake limited FDA compliance by setting up Institutional Review Boards (IRBs) (8). Laboratories associated with academic and research institutions are regulated similarly, and usually have IRBs and prescribed human subjects protocols (8).

FDA Regulation of Kits. FDCA as amended by MDA (107) and the Safe Medical Devices Act of 1990 (156) established a comprehensive system to regulate the safety and effectiveness of medical devices. The level of review depends on a device's classification (159). Class I devices are devices that pose no unreasonable risk to health and whose safety and effectiveness can be ensured by general controls (158). Class II devices require specialized controls to ensure their safety and effectiveness (158). Most genetic tests, however, would probably fall into Class III because of their complexity (160,161). Class III devices require submission of a premarket approval application for FDA review unless "substantial equivalence" can be established with a previously approved device through Section 510(k) notification (162). When an application is

received, FDA assigns the product to an examiner who possesses the relevant scientific background to review the application. Although no more than a handful of genetic testing products have been submitted for review (163), the agency already has designated an examiner to specialize in genetic tests (164). Oncor's Inform Her-2/Neu breast cancer genetic diagnostic was one of the first genetic tests to pass FDA review (165). When Oncor first submitted a version of the gene diagnostic kit, FDA rejected the application (166). It was not until after the company "expanded reproducibility studies, set forth manufacturing criteria, and ... develop[ed] a training program for the physicians and technicians using the test" that it was approved for use in predicting the likelihood of recurrence in previously symptomatic patients (166). FDA also requires compliance with good manufacturing practice (GMP) standards, monitors advertising, and expedites clearance by utilizing postmarket control (83). The agency further might condition approval on the company developing a system to ensure counseling or physician consultation (167).

Clinical Laboratory Improvement Amendments. The Clinical Laboratory Improvement Amendments of 1988 (CLIA) were passed to establish minimum quality standards for laboratory testing (168). The Health Care Financing Administration (HCFA), which implements CLIA, reviews some 158,000 laboratories across the country pursuant to the act (169). CLIA sets forth quality standards for proficiency testing, patient test management, quality control, personnel qualifications and quality assurance (170). There is no effort under CLIA's regulatory administration to establish the utility of or the validity of using any particular test. Under CLIA, the laboratory need only demonstrate that its procedures result in accurate and reliable identification of the target, namely, technical competence in test performance (2). Furthermore, much of the review process is conducted through third-party accreditation organizations and programs (171). The scope of review is limited to laboratory procedures and its purpose, to certify labs for payment under federal health plans, is largely administrative.

State Regulation. The states of Washington, Oregon, and New York have laboratory review regulations rigorous enough to qualify them for exemption from CLIA regulation by the HCFA (172). However, only New York undertakes any significant assessment of genetic testing. If a company or laboratory located in New York or providing testing to New York residents wishes to offer a genetic test, it must become certified for the particular test, a process that involves consideration of the quality and effectiveness of the test (28,173).

Choosing a Regulatory Actor

There is thus a sharp regulatory divide between genetic testing kits, which receive the full intensity of FDA review of safety and effectiveness, and widely marketed genetic tests offered as a service, which generally receive no review of safety and effectiveness. It is a regulatory divide not warranted by any significant factual distinctions. A genetic testing recipient requires as much protection from the harms of a proprietary genetic test provided as a service, as a genetic test sold as a kit. If a company promotes a genetic test in interstate marketing directly to the public, the test, even if performed as a service, loses the characteristics of medical practice and carries the hallmarks of commercialization. It is a product being purveyed to the public, and as such, should be subject to the scrutiny of the nation's gatekeeper of medical devices, the FDA. In fact FDA is the only agency whose existing legislative authorization might cover regulation of genetic testing services.

The lack of convincing alternatives emphasizes the appropriateness of FDA oversight. Besides FDA, there is a fairly short list of other serious potential regulators of genetic testing: the Centers for Disease Control (CDC), the HCFA, or the Department of Health and Human Services (DHHS) (8). Both CDC and HCFA are agencies within DHHS (as is FDA). CDC "is responsible for promoting health and quality of life by preventing and controlling disease, injury, and disability" (174). CDC is the nation's prevention center; its efforts are directed primarily at monitoring, researching, and preventing diseases, and cover a broad range of social activities including collaborative projects with either HCFA or FDA (175,176). HCFA administers the Medicare and Medicaid programs and CLIA (177). HCFA's duties in certifying laboratories and reviewing testing proficiency are related to its administration of Medicare and Medicaid and are intended to ensure that laboratories are qualified for the work for which they are being reimbursed, which is similar to its authority to prescribe nursing home standards (178). Finally, DHHS is mentioned not so much to suggest that the department itself undertake regulation of genetic testing, but that a new body within DHHS be created to do so. As an initial matter, it is not apparent that these alternative regulatory bodies have any existing statutory authorization to regulate the clinical validity of genetic testing, which creates a substantial hurdle to effective action and suggests a lack of traditional jurisdiction. Under CLIA, the CDC determines the categorization for tests performed by laboratories, and HCFA reviews their competence. As discussed, however, CLIA does not include the authority to review the quality or scientific validity of a genetic test and focuses instead exclusively on proficiency. However, the public policy question remains as to whether the CDC or HCFA have more expertise and a more appropriate structure than FDA to substantively regulate genetic tests, or whether all the agencies are so poorly suited to the task that a new body should be invented.

HCFA does not appear to have developed particular genetics expertise in its regulation of clinical laboratories (179,180). While it may have a better infrastructure to undertake widespread regulation of many different sites, the clinical validity of a type of genetic test only needs to be established once. Nor does HCFA seem to have familiarity in conducting a substantive scientific review. CDC could be a more promising candidate because of its scientific expertise and prominent historical role in disease control. Setting up a new process for conducting safety

and effectiveness review for genetic testing services under CDC, however, seems redundant because FDA already is implementing similar measures to regulate genetic testing kits. The science and protocols for evaluating a genetic test for clinical utility should not change based on how the test is delivered. Furthermore the issues raised in regulating genetic testing conform to those routinely raised for FDA such as identifying a professional standard, expediting clearance through postmarket controls and tracking, monitoring advertising, and determining equivalence with a previously approved produce.

In fact, the only benefit of choosing another agency or charting a new one comes not from any advantage in scientific qualifications or institutional expertise, but in avoidance of perceived excessive bureaucracy at FDA. This concern, however, may be misjudging the matter. Government review and approval of a diagnostic is inherently bureaucratic. Creating a new organization to undertake the task or assigning it to one with no experience is nothing more than reinventing the wheel and hoping it turns out better this time. Not only is such a course somewhat doubtful in rationale, it is duplicative of existing FDA regulation of genetic testing kits and therefore wasteful of public resources. From the perspective of expertise and experience, it is clear that FDA may most readily provide the regulation called for in genetic testing services.

CONCLUSION

Perhaps the most disturbing aspect of the issues discussed is the immediate need for oversight. FDA is the best positioned agency from the perspective of jurisdiction and institutional expertise to undertake the task, and what it requires is not so much legislative authorization as a social and political mandate. Promptness in action is not only essential in light of the particular circumstances in genetic testing but foreshadows the proficiency of society to address these problems in the future. Commercialization of genetic testing is one of the first challenges presented by the wondrous and rapid developments in genetic technology, and it only can be hoped that the government will prove itself capable of managing the dangers with the energy and foresight necessary to ensure that the public safely receives the benefits of a new era in genetics.

ACKNOWLEDGMENT

Originally published as A. Huang, *Food Drug Law J.* **53**, 555–591 (1998). Reprinted with the permission of The Food and Drug Law Institute.

BIBLIOGRAPHY

1. President W.J. Clinton, *U.S. Newswire*, February 8 (2000).
2. S.N. Hurd, *Law Polic Rep.*, October, 161 (1997).
3. Task Force on Genetic Testing, NIH–DOE, *Final Report: Promoting Safe and Effective Genetic Testing in the United States*, U.S. Government Printing Office, Washington, DC, September 1997.
4. L. Andrews et al., *Institute of Medicine Assessing Genetic Risks: Implications for Health and Social Policy*, 1993.
5. E. Lane, *Newsday*, November 5, 17 (1993).
6. L. Seachrist, *Bioworld Today*, April 7, 1 (1997).
7. M.K. Pendergast, FDA Deputy Comm'r. & Special Advisor to the Comm'r. Hearing before House Comm. on Science, 104th Cong. 2nd Sess. (1996).
8. B. Burlington, Director, Center for Devices and Radiological Health, to FDA, Task Force on Genetic Testing, April 3 (1997). Contents available at: FDA Freedom of Information Office, 5600 Fisher's Lane #12A16, Rockville, MD 20857.
9. M.J. Malinowski and R.J.R. Blatt, *Tul. Law Rev.* **71**, 1211, 1296 (1997).
10. Department of Health and Human Services, Advisory Committee on Genetic Testing, *The Gray Sheet*, November 1 (1999).
11. Human Genome Initiative *Genome Database* (visited July 21, 1998), Available at: *<http://www.gdb.org/>*
12. *Nat. Biotech.*, December, 1627 (1996).
13. L. Wasowicz, *Assoc. Press, BC Cycle*, November 23, 1996, *Available in:* Westlaw, Allnews Database.
14. Medical Device Amendments (MDA), Pub. L. No. 94-295, 90 Stat. 539 (codified at 15 U.S.C. §55 (1994), scattered sections of 21 U.S.C.).
15. J.W. Hulse et al., *Food Drug Law J.* **48**, 285–287 (1993).
16. L.M. Fisher, *N.Y. Times*, January 1, A1 (1998).
17. *Bus. Wire*, May 25, 1998, *Available in:* Westlaw, Allnews Database.
18. *PR Newswire*, August 22, 1989, *Available in:* Westlaw, Allnews Database.
19. R. Weiss, *Washington Post*, May 26, A1 (1996).
20. J.D. Watson et al., *Recombinant DNA*, 2nd ed., 539, 1992.
21. M. Guttman, *USA Today (Weekend)*, February 9 (1997).
22. S. Seshadri et al., *Archi. Neurol.* 1074 (1995).
23. T.M. Maugh II, *L.A. Times, reprinted in The Record* (N. NJ), March 16, 1998, at H04.
24. *Pulse*, February 28, 32 (1998).
25. *Chem. Bus. NewsBase*, February 16 (1999).
26. N.J. Nelson, *J. Natl. Cancer Inst.*, January 17, 70 (1996).
27. W.R. Wineke, *Wis. St. J.*, March 9, 1C (1998).
28. R. Osborne, *Bioworld Today*, February 24, 1 (1998).
29. *Standard & Poor's*, August 28, 2 (1997).
30. H.M. Kingston, *ABC Clin. Genet.* 31 (1997).
31. C. Cookson, *Fin. Times (Tech.)*, June 27, 12 (1996). *Available in:* Westlaw 1996 WL 10598276.
32. Genetic Testing Quality Assurance Office, New York State Department of Public Health, New York State Genetic Testing Laboratories F12, F13 (1998).
33. Congressional testimony of Francis S. Collins, 1998 WL 12760444 (May 21, 1988).
34. S. Okie, *Washington Post*, May 26, Z11 (1998).
35. *Irish Times*, May 21, 15 (1998).
36. *Chicago. Trib.*, May 6, 7 (1998).
37. *Gene Therapy Wkly*, April 27 (1998).
38. P.S. Harper, in P.S. Harper and A.J. Clarke, eds., *Genetics, Society and Clinical Practice*, vol. 7, 1997.
39. *Marketletter*, May 18 (1998), *Available in* 1998 WL 11622549.
40. Z. Moukheiber, *Forbes*, June 1, 119 (1998).
41. R. Kotulak, *Chicago. Trib.*, May 14, 1 (1998).

42. J. Siegel-Itzkovich, *Jerusalem Post*, November 23, 10 (1997).
43. P. Recer, *Assoc'd. Press*, May 17 (1998).
44. D. Brown, *Washington Post*, February 20, E1 (1998).
45. D.L. Wheeler, *Chron. Higher Educ.*, November 29, A24 (1999).
46. M. Elias, *Red Flag Raised on Gene Test, USA Today*, October 27, 1D (1997).
47. N. Peterson, *Breast Cancer Action* (Breast Cancer Action, San Francisco, CA), June 1996.
48. P.A. Hoffee, *Med. Mol. Genet.* 279–280 (1998).
49. D. Shattuck-Eidens et al., *J. Am. Med. Assoc.* **278**, 1284 (1997).
50. P. Kahn, *Science* **274**, 496–498 (1996).
51. J. Gillis, *Washington Post*, April 7, C2 (1998).
52. Stanford University Program in Genomics, Ethics, and Society, Report of the Working Group on Genetic Testing for Breast Cancer Susceptibility (November 23, 1996), Available at: *<http://www-leland.stanford.edu/dept/scbe/brcaexec.htm>*
53. B.A. Koenig et al., *J. Women's Health* **7**, 531–545 (1998).
54. A. Marshall, *Nat. Biotech.* **14**, 1642 (1995).
55. B.A. Brenner, *Breast Cancer Action* (Breast Cancer Action, San Francisco, CA), June 1996.
56. G. Kolata, *N.Y. Times*, April 1, A1 (1996).
57. Myriad Genetics Laboratories, Inc., BRCA1 Genetic Susceptibility for Breast and Ovarian Cancer: A Reference for Healthcare Professionals in Anticipation of BRCA1 Genetic Susceptibility Testing (January 1996).
58. B. Newman et al., *J. Am. Med. Assoc.* **279**(12), 915–921 (1998).
59. A. Petrij-Bosch et al., *Nat. Genet.* **17**(3), 341 (1997).
60. L. Jennings, *Comm. Appeal*, April 26, F1 (1998).
61. T.H. Maugh II, *L.A. Times*, May 25, S2 (1998).
62. P. Lynden, *Am. Health for Women*, June 1, 29 (1997).
63. C. Wallis et al., *Time*, May 18, 48 (1998).
64. L. Seachrist, *Bioworld Today*, April 13 (1998).
65. B. Kuska, *J. Nat. Cancer Inst.* **87**, 1578 (1995).
66. L. Sowers, *Houston Chron.*, August 17, 1 (1997).
67. J. Brandt et al., *J. Am. Med. Assoc.* **261**, 3108 (1989).
68. N.M. Resnick, *Pittsburgh Post-Gazette*, March 17, G1 (1998).
69. D. Grady, *N.Y. Times*, January 7, C3 (1997).
70. P. Bereano and R. Sclove, *Washington Post*, March 22, C5 (1998).
71. *Smith v. Olin Chem.* Corp., 555 F.2d 1283 (5th Cir. 1977).
72. K. Painter, *USA Today*, August 15, 1A, 2A (1997).
73. J. Talan, *Newsday*, March 21 (1998), at 1998 WL 2663370.
74. Mayo Clinic, *Lou Gehrig and Stephen Hawking: Stars Despite ALS* (visited September 4, 1998), Available at: *<http://www.mayohealth.org/mayo/9604/htm/als_sb.htm>*
75. *Wkly J.*, January 30, 12 (1998).
76. C. Francke, *Baltimore Sun*, October 30, 1A (1998).
77. J. Price, *Washington Times*, September 2, A1 (1997).
78. C. Russell, *Washington Post*, February 24, Z08 (1998).
79. R. Kotulak, *Chicago. Trib.*, October 17 (1997).
80. M. Downey, *Atlanta J. & Atlanta Const.*, October 19, 1 (1997).
81. S. Roan, *L.A. Times*, April 29, 13 (1997).
82. *Fed. Regist.* **61**, 10,484 (1996) (to be codified at 21 CFR §809 & 864).
83. R.A. Merrill, *Virginia A. Law Rev.* **82**, 1753 (1996).
84. Regulation of New Drug R&D by the Food and Drug Administration, 1974: Joint Hearings before the Subcomm. on Health of the Senate Comm. on Labor and Public Welfare and the Subcomm. on Administrative Practice and Procedure of the Senate Comm. on the Judiciary, 93d Cong., 2d Sess. 207 (1974) (statement of Dr. Alexander M. Schmidt, the FDA Comm'r).
85. H. Burkholz, *The FDA Follies* 105-41 (1994).
    J. Schwartz and R. Marcus, *Washington Post*, July 13, A4 (1996).
86. J. Schwartz, *Washington Post*, July 15, A17 (1996).
87. M. Kraus, *Cal. W. Law Rev.* **33**, 101 (1996).
88. P. Duggan, *Washington Post*, October 12, B1 (1988).
89. J. Schwartz, *Conservative Foes of Government Regulation Focus on the FDA, Washington Post*, January 21, A7 (1995). 142 CONG. REC. H5631-38 (daily ed. May 29, 1996) (statements of Reps. Greenwood, Barton, Klug, Burr, and Fox).
90. M. Simons, *L.A. Times*, January 2, A1 (1995).
91. *World News Tonight* (ABC television broadcast, February 21, 1996).
92. *Med. Health* **49**, (Feb. 20), 1 (1995).
93. W.E. Leary, *N.Y. Times*, June 23, A18 (1994).
94. J. Mann, *Washington Post*, April 22, E3 (1994).
95. E.J. Millenson, *Washington Post*, August 27, C5 (1989).
96. *Buffalo News*, June 24, C2 (1994).
97. H.H. Harris, *Washington Post*, April 9, G1, G5 (1994).
98. C.T. Fang et al., *Patient Care* **23**, October 30, 19 (1989).
99. L. Garrett, *Newsday*, June 23, A7 (1994).
100. D.J. Murphy, *Investors Bus. Daily*, October 1, A4 (1994).
101. J.C. Fletcher and D.C. Wertz, *Emory Law J.* **39**, 747 (1990).
102. T. Duster, *Backdoor to Eugenics*, 1990.
103. D.B. Paul, *Soc. Res.* **59**, 663–683 (1992).
104. D.J. Kevles, *In the Name of Eugenics: Genetics and the Uses of Human Heredity*, 1985.
105. P. Reilly, *Genetics, Law, and Social Policy*, 1977.
106. J.L.J. Nuzzo, *Food Drug Law J.* **53**, 35, 40 (1998).
107. J.E. Hoffman, *Food and Drug Law*, 1–2 (1991).
108. Regier, *Law Contemp. Probs.* **1** (1933).
109. I. Scott Bass, *Food Drug Law* **61** (1991).
110. Pure Food and Drugs Act, Pub. L. No. 59–384, 34 Stat. 768 (1906) (codified at 21 U.S.C. §§1–15 (1934)) (repealed in 1938 by 21 U.S.C. §392(a)).
111. *Report of the Secretary of Agriculture on Deaths Due to Elixir Sulfanilamide-Massengill*, S. Doc. No. 124, 75th Cong., 1st Sess. (1937).
112. Pub. L. No. 75-717, 52 Stat. 1040 (codified at 21 U.S.C. §§301 et seq.).
113. 21 U.S.C. §355(d) (Supp. 1998).
114. S. Rep. No. 1744, 87th Cong., 2d Sess. 16 (1962).
115. *Fed. Regist.* **37**, 16,503 (August 15, 1972).
116. *Fed. Regist.* **62**, 62,243 (November 21, 1997).
117. *United States v. Apex Distributing Co.*, 148 F. Supp. 365 (D. R.I. 1957).
118. U.S. Const. art. I, §8, cl. 3.
119. G.R. Stone et al., *Constitutional Law* 194 (1991).
120. *United States v. Lopez*, 514 U.S. 549, 558 (1995).

121. *United States v. Sullivan*, 332 U.S. 689 (1948).
122. *United States v. Funk*, 412 F. 2d 452 (8th Cir. 1969).
123. *Dean Rubber Mfg. Co. v. United States*, 356 F. 2d 161 (8th Cir. 1966).
124. *Almendarez-Torres v. United States*, 118 S. Ct. 1219, 1227 (1998); *United States ex rel. Attorney General v. Delaware & Hudson Co.*, 213 U.S. 366, 408 (1909).
125. 317 U.S. 111 (1942).
126. N.A. Holtzman et al., *Science*, October 24, 602 (1997).
127. 21 U.S.C. §331(k) (FDCA §301(a)). 21 U.S.C. §321 (h) (FDCA §201(h))
128. 932 F.2d 813 (9th Cir. 1991).
129. *Grand Labs. v. Harris*, 660 F.2d 1288 (8th Cir. 1981) *cert. denied sub nom. Grand Labs. v. Schweiker*, 456 U.S. 927, 102 S. Ct. 1972, 72 L. Ed. 2d 442 (1982).
130. *Kordel v. United States*, 335 U.S. 345, 349 (1948).
131. *Drown v. United States*, 198 F.2d 999 (9th Cir. 1952).
132. *Western Union Tel. Co. v. Foster*, 247 U.S. 105 (1918).
133. *Barnes v. United States*, 142 F.2d 648 (9th Cir. 1944).
134. J. Weber, *Bus. Week*, March 16, 108 (1998).
135. *United States v. An Undetermined Number of Unlabeled Cases*, 21 F.3d 1026 (10th Cir. 1994).
136. C.J. Walsh and A. Pyrich, *Rutgers Law Rev.* **48**, 883, 890 (1996).
137. *Linder v. United States*, 268 U.S. 5, 18, 45 S. Ct. 446, 449, 69 L. Ed. 819 (1925).
138. 453 F. Supp. 1141 (M.D. Ala. 1978).
139. *United States v. Evers*, 643 F.2d 1043 (5th Cir. 1981).
140. Food and Drug Administration Modernization Act of 1997 (the FDAMA), Pub. L. No. 105–115, 111 Stat. 2296 (1997).
141. FDCA §201(h), (codified it 21 U.S.C. §321).
142. J.M. Zitter, *A.L.R. Fed.* **129**, 343 (1996).
143. University of California, San Diego, Biochemical Genetics, *Biochemical Genetics Test List* (UCSDW3BG) (visited July 22, 1998), Available at: *<http:www.biochemgen.ucsd.edu>*
144. S. Gottlieb, *Times Union*, March 12, A11 (1998).
145. Telephone Interview with Genetics and IVF Institute customer service representative (Apr. 8, 1998) (tape on file with author).
146. C. Cookson and C. Adams, *Fin. Times*, September 24, 12 (1997).
147. E. Chang, *Washington Post* (Magazine), September 7, W19 (1997).
148. A. Heller, *Drug Store News*, October 6, 21 (1997).
149. F.M. Giardiello et al., *N. Engl. J. Med.* **12**, 823, 825 (1997).
150. L. Seachrist, *Bio World Today*, March 12, 1 (1998).
151. *U.S. Newswire*, March 11 (1998).
152. K.J. Hoffman et al., *Acad. Med.* **68**, 625 (1993).
153. P.H. Silverman, *Hastings Cent. Rep.*, May 1, S16 (1997).
154. M. Guidera, *Baltimore Sun*, April 2, 1D (1998).
155. Telephone Interview with Joe Robinson, Coordinator, New England Regional Genetics Group (May 22, 1998).
156. Pub. L. No. 101–629, 104 Stat. 4511 (codified in scattered sections of 21 U.S.C. §§301 note-383note, 42 U.S.C. §263b-263n).
157. 21 U.S.C. §360c (FDCA §513).
158. 21 U.S.C. §360c(a)(1)(A) (FDCA §513(a)(1)(A)).
159. 21 U.S.C. §360c(a)(1)(B) (FDCA §513(a)(1)(B)).
160. M.J. Malinowski and M.A. O'Rourke, *Yale J. Reg.* **13**, 163, 206 (1996).
161. 21 U.S.C. §360c(a)(1)(C) (FDCA §513(a)(1)(C)).
162. Food and Drug Administration Modernization Act of 1997, Pub. L. No. 105–115, 111 Stat. at 2296.
163. Telephone Interview with Dr. Steven Gutman, Dir., Div. of Clinical Laboratory Devices, FDA (May 1, 1998).
164. Food and Drug Administration, Center for Devices and Radiological Health, Monitoring the Human Genome Project for Impact on Developments in Medical Devices [contents available at the FDA Freedom of Information Office, 5600 Fisher's Lane #12A 16, Rockville, MD 20857].
165. B. Berselli, *Washington Post*, January 26, F5 (1998).
166. *Daily Record*, June 11, 2 (1997).
167. *Chicago Trib.*, May 15, 3 (1996).
168. Pub. L. No. 100–578, 102 Stat. 2903 (codified at 42 U.S.C. 201 note, 263a, 263a note).
169. Health Care Financing Administration, *Clinical Laboratory Improvement Amendments* (visited July 28, 1998), Available at: *<http://www.hcfa.gov/medicare/hsqb/clia1.htm>*
170. 42 CFR 493 *et seq.*
171. Regulatory notices HCFA-2246-N (published 4/9/98) (approving the Joint Commission on Accreditation of Healthcare Organizations, the American Association of Blood Banks, and the American Osteopathic Association) and HSQ-242-N (published 5/19/97) (approving the Commission on Office Laboratory Accreditation).
172. Regulatory notices HSQ-243-N (published 4/9/98) (exempting Washington), HSQ-231-N (published 6/31/96) (exempting Oregon) and HSQ-230-N (published 8/28/95) (exempting New York).
173. Telephone Interview with Nanette P. Healy, Assistant Administrator, New York Genetic Testing Quality Assurance Office (January 20, 1998).
174. Centers for Disease Control and Prevention, *CDC FY 1999 President's Budget Request, General Statement from the FY99 Congressional Justification* (visited Aug. 11, 1998), Available at: *<http://www.cdc.gov/fy99bdgt/99genstm.htm>*
175. CLIA Program; Cytology Proficiency Testing, *Fed. Regist.* **60**, 61,509 (November 30, 1995) (to be codified at 42 C.F.R. pt. 493), 1995 WL 701433.
176. *Med. Malpractice Law Strategy* **14**, 1 (January 1997).
177. Office of the Inspector General, Health Care Financing Administration Projects, 995 PLI/Corp. 585, 639 (June 1997).
178. Medicaid Program; Standards for Intermediate Care Facilities for the Mentally Retarded, Proposed Rules, DHHS, HCFA, *Fed. Regist.* **51**, 7520 (March 4, 1986) (to be codified at 42 C.F.R. pts. 435, 442) 1986 WL 90426.
179. *CLIA & Genetic Testing*, F-D-C REP. ("The Blue Sheet"), (February 11, 1998), Available in Westlaw, 1998 WL 9283076.
180. Summaries of board meetings are available at HCFA, Office of Information Services, Division of Freedom of Information and Privacy, 7500 Security Boulevard, Room C2-01-11, Baltimore, MD 21244-1850.

See other entries FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE; see also GENETIC INFORMATION, ETHICS and GENETIC INFORMATION, LEGAL entries.

GENETIC INFORMATION, LEGAL, GENETIC PRIVACY LAWS

MARY R. ANDERLIK
University of Houston
Houston, Texas

REBECCA D. PENTZ
M.D. Anderson Cancer Center
Houston, Texas

OUTLINE

INTRODUCTION

Privacy is a broad concept. In the legal context it includes at least four categories of concern: (*1*) access to persons and personal spaces; (*2*) access to information by third parties, and also any subsequent disclosure of this information by third parties (the category of concern best captured by the term "confidentiality"); (*3*) third-party interference with personal choices, especially in intimate spheres such as procreation; and (*4*) ownership of materials and information derived from persons (1). Typically statutes characterized as "genetic privacy acts" address a number of these concerns, regulating genetic testing and other means of generating genetic information, limiting access to genetic information, and ruling out certain uses of genetic information by third parties. However, some states have laws that are much narrower. Further, even laws that are broad in terms of the range of privacy concerns addressed may be narrow in scope, targeting a particular industry or type of information keeper. For example, many genetic privacy laws were passed in response to abuses, actual or perceived or potential, in the insurance industry and by employers. Some laws are tailored to problems that arise in the areas of research or law enforcement or family relations.

For organizational purposes, this article is divided into four sections. Each section covers major developments at the state and federal levels, and common (or judge-made) law as well as statutory law, to the extent relevant. Given the range and complexity of the issues, this entry is intended as an overview. For more in-depth treatment, the reader may wish to consult the sources cited in the references.

LAWS REGULATING THE GENERATION, DISCLOSURE, AND USE OF GENETIC INFORMATION

State Laws

Genetic Privacy Laws. At present, the only laws that comprehensively address genetic privacy are at the state level. At least 35 states have some statutory law relating to genetic privacy (excluding laws authorizing creation of DNA banks and databases for law enforcement purposes, which may include privacy protections). See Table 1. The most comprehensive of these laws include general provisions covering genetic testing and the handling of genetic information, accompanied by more focused provisions addressing special concerns that arise in connection with insurance and employment, and in Texas and Wisconsin, occupational licensing.

Most genetic privacy laws cover at least two aspects of privacy, access to persons (for testing) and access to information. See Table 2. About half prohibit genetic testing of persons or samples without prior informed consent, subject to certain exceptions, for example, law enforcement, paternity determination, court order, and anonymous research. In many states the elements of the consent form are specified; standard elements include a description of the test, a statement of the purposes of testing, and the names of the persons or entities to whom results may be released. It is common for laws to contain a statement that genetic information is confidential, or "confidential and privileged," meaning that it is protected from subpoena in a civil proceeding, although production can still be compelled by a specific court order. Disclosure of genetic information to a third party without written authorization or consent is generally prohibited unless an exception applies. The standard list of exceptions parallels the list of exceptions to the consent requirement. Some states also permit disclosure of genetic information for the benefit of blood relatives if the subject is dead (2). Usually the authorization for disclosure must be specific rather than general. Many state laws provide that these legal protections only apply to genetic material or information that can be identified as belonging to an individual or family.

In the area of insurance, a major issue is breadth of application of genetic privacy laws. Many states limit special privacy protections for genetic testing and information to *health* insurance, leaving consumers with few or no safeguards in their dealings with life, disability

Table 1. State Genetic Privacy Laws (as of November 30, 1999)

State	Statute
Alabama	Ala. Code §27-53-1 et seq.
Arizona	Ariz. Rev. Stat. Ann. §20-448.02
California	Cal. Civ. Code §56.17; Cal. Health & Safety Code §1374.7; Cal. Ins. Code §§742.405, 742.407, 10123.3, 10123.35, 10140, 10140.1, 10146 et seq.
Colorado	Colo. Rev. Stat. §10-3-1104.7
Connecticut	Conn. Gen. Stat. Ann. §46a-60
Delaware	Del. Code Ann. tit. 16, §1220 et seq.; Del. Code Ann. tit. 19, §711
Florida	Fla. Stat. Ann. §§627.4301, 636.0201, 760.40
Georgia	Ga. Code Ann. §33-54-1 et seq.
Hawaii	Haw. Rev. Stat. §§431 : 10A-118, 432 : 1-607, 432D-26
Illinois	Ill. Comp. Stat. Ann. §513/1 et seq.
Indiana	Ind. Code Ann. §27-8-26-1 et seq.
Kansas	Kan. Stat. Ann. §44-1002 et seq.
Kentucky	Ky. Rev. Stat. Ann. §304.12-085
Louisiana	La. Rev. Stat. Ann. §22 : 213 : 7
Maine	Me. Rev. Stat. Ann. tit. 5, §19301-02; Me. Rev. Stat. Ann. tit. 24-A, §2159-C
Maryland	Md. Ins. Code Ann. §27-909
Minnesota	Minn. Stat. Ann. §72A.139
Missouri	Mo. Ann. Stat. §375.1300 et seq. (Vernon's)
Montana	Mont. Code Ann. §33-18-901 et seq.
Nevada	Nev. Rev. Stat. Ann. §§629.111 et seq., 689A.417, 689B.069, 689C.198, 695C.207, 695B.317
New Hampshire	N.H. Rev. Stat. Ann. §141-H : 1 et seq.
New Jersey	N.J. Stat. Ann. §§10 : 5-12, 10 : 5-43 et seq., 17B : 30-12
New Mexico	N.M. Stat. Ann. §24-21-1 et seq.
New York	N.Y. Civ. Rights Law §79.1; N.Y. Ins. Law §2612 (McKinney)
North Carolina	N.C. Gen. Stat. §58-3-215
Ohio	Ohio Rev. Code Ann. §§1751.64, 1751.65, 3729.46, 3901.49, 3901.491, 3901.50, 3901.501 (Baldwin)
Oklahoma	Okla. Stat. Ann. §tit. 36, §3614.1 et seq.
Oregon	Or. Rev. Stat. §§659.036, 659.700 et seq.
Rhode Island	R.I. Gen. Laws §§27-18-52, 27-19-44, 27-20-39, 27-41-53, 28-6.7-1 et seq.
South Carolina	S.C. Code Ann. §38-93-10 et seq.
Tennessee	Tenn. Code Ann. §56-7-2701 et seq.
Texas	Tex. Labor Code Ann. §21.401 et seq.; Tex. Ins. Code Ann. §21.73; Tex. Rev. Civ. Stat. Ann. §9031 (Vernon's)
Vermont	Vt. Stat. Ann. tit. 18, §9331 et seq.
Virginia	Va. Code Ann. §§38.2-508.4, 38.2-613
Wisconsin	Wis. Stat. Ann. §942.07

income, and long-term care insurers, among others. Texas is unusual in limiting protections affecting insurance to group health benefit plans (3). Legislation specifically directed at life insurers has been enacted by a number of states, but the protections afforded applicants for life insurance are minimal. Typically these laws require only that informed consent be obtained prior to performance of any genetic testing and/or that any use of genetic information in medical underwriting meet standards of actuarial fairness. Note that state protections are generally considered to be inapplicable to self-funded employer-sponsored benefit plans, including employer-sponsored health and life insurance, due to the operation of the federal Employee Retirement Income Security Act of 1974, commonly known as ERISA (4).

Many genetic privacy laws are silent on the issue of retention of samples (i.e., biological specimens obtained or retained for the purposes of genetic testing). A few states require destruction of samples upon specific request, or after the purpose for which the sample was obtained has been accomplished. The New York law requires that the sample be destroyed at the end of the testing process or not more than 60 days after the sample is taken, unless a longer period of retention is expressly authorized (5). Laws that require destruction of samples typically include exceptions related to research and law enforcement, areas discussed in more detail below.

States also vary in the sanctions imposed for violations of privacy protections. In most states, a violation is a misdemeanor punishable by fine or jail time or both. (A willful violation may be a felony.) Further, a number of states allow individuals to sue for equitable relief, such as an order to stop a violation, and damages, costs and attorney fees (6). The Louisiana law relating to health insurance allows recovery of the greater of actual damages or $50,000 ($100,000 in cases of willful violation) against persons who violate the law by negligently collecting or disclosing genetic information, and the law authorizes treble damages where a violation resulted in monetary gain (7).

Definitional Issues. Whether a law relates to insurance, employment, or has a more general orientation, the choices legislators make in two key areas have significant implications: (*1*) how to define the category or categories of protected material or information, and (*2*) whether and how to address the problem of compelled consent.

State genetic privacy laws typically focus on the generation, disclosure and use of "genetic information," a term that may be linked to a definition of "genetic test" and/or "genetic characteristic." Many states that have genetic privacy laws use one or more of these definitions to limit protections to persons who are presymptomatic or asymptomatic for disease. (The protections in an Alabama law are even more limited, only applying to presymptomatic testing for a predisposition to cancer [8].) From a policy perspective, this limitation may be defensible; it is unclear why persons with diseases that are genetic in origin should be favored vis-à-vis persons with diseases that arise in some other fashion. On the other hand, the "genetic revolution" has blurred many lines that formerly appeared sharp, even creating questions about what counts as a disease. For example, the Illinois Genetic Information Privacy Act provides that results of genetic testing that indicate the person is already afflicted with a disease, "whether or not currently symptomatic," are not subject to the confidentiality requirements of the act (9). This provision suggests that a disease may exist

Table 2. Summary of Key Provisions in State Genetic Privacy Laws (as of November 30, 1999)

Provision	State
General provisions	
Genetic testing generally prohibited without prior informed consent	AZ, DE, FL, GA, NV, NH, NJ, NM, NY, OR, SC, VT
Standard exceptions include law enforcement, paternity determination, court order, anonymous research	
Standard elements of consent form include description of test, statement of purpose(s), who receives results	
Person requesting genetic testing must advise of risk of discrimination	VT
Genetic information confidential and/or privileged	AZ, CA, CO, GA, IL, MO, NY, OK, OR, SC, TX
No release of genetic information without specific authorization	AZ, CA, CO, DE, FL, GA,
Standard exceptions include law enforcement, paternity determination, court order, anonymous research, benefit of relatives (subject deceased)	HI, IL, LA, MD,[a] MO, NV, NJ, NM, NY, OR, SC, TN,[a] TX,[b] VT, VA,[a] WI,[c]
Requests for genetic services expressly protected	DE, HI, LA, KY, MD, MN, MO, NV, NH, NM, RI, VT, VA
Destruction of (identifiable) sample generally required upon request and/or accomplishment of purpose	DE, LA, NV, NJ, NM, NY, OR, TX
Definitions	
Narrow definition of genetic test/characteristic/information:	
• Test (direct) for alterations in genes	CO, IN, KS, MO, OH
• Test for predisposition and/or environmental damage and/or carrier status	AZ, DE, IN, ME, NJ, NM, NY, TX, WI
• Person must be a- or presymptomatic	AL, CA, FL, GA, IL,[d] KY, ME, MD, MN, MT, OK, SC, TN, VT, VA
Genetic information expressly excludes family history	FL, MO, MT, OK, TN
Insurance and employment	
Health insurers prohibited from requiring genetic testing[e]	AL,[f] CA, FL, GA, HI, IL, IN, KY, LA, ME, MD, MN, MO, MT, NV, NH, NJ, OH, OK, RI, SC, TN, TX,[g]
Health insurers prohibited from requesting genetic testing and/or information (some states add "for nontherapeutic purposes")[e]	CA, CO, FL, GA, HI, IL,[h], IN,[h] KY, LA, MD, MN, MO, MT, NH, OH, OK, RI, TN
Employers prohibited from requiring genetic testing (some states add "unless job related" or similar language)	CT, DE, KS, ME, NH, NJ, OK, OR, RI, TX, VT
Employers prohibited from requesting genetic testing and/or information (some states and "unless job related" or similar language)	CT, DE, KS, NH, OK, OR, VT
Sanctions	
Individual right to sue	CA, CO, DE, GA, IL, LA, NV, NH, NJ, NM, RI, SC

[a] Insurance context.
[b] Employment, group health insurance, and occupational licensing contexts.
[c] Employment context.
[d] Provision in Act states that confidentiality protections do not apply if genetic testing indicates that the individual is at the time of the test afflicted with a disease, whether or not currently symptomatic.
[e] Self-funded plans covered by ERISA may be exempt.
[f] Genetic tests for predisposition to cancer.
[g] Group health insurers.
[h] Insurers may consider favorable test results voluntarily submitted by an individual.

purely on the basis of genotype, without any phenotypic manifestation. Another area of variation is whether "genetic information" includes biological materials such as tissue samples.

Some laws specify that only testing or information relating to inherited genes or genetic characteristics is covered, pointing up an ambiguity in the adjective "genetic." Many diseases, including all cancers, can be described as genetic in the sense that they are triggered by altered genes. However, only 5 to 10 percent of cancers are thought to be closely linked to a particular set of inherited genetic defects. Another area of variation is the stringency of the definition of genetic test. Some states limit the term to direct tests for alterations in genetic materials, while others include tests of proteins and other gene products. It is also common to exclude routine medical tests, such as cholesterol tests, human immuno-deficiency virus (HIV) tests, and drug tests from the definition of genetic test.

Genetic tests are not the only source of information about predisposition to disease or disease risk. Some of the first documented cases of genetic discrimination involved inferences from family history (10). Several

states use a broad definition of genetic information that appears to encompass information in a family history. For example, the Connecticut law defines genetic information as "information about genes, gene products or inherited characteristics that may derive from an individual or family member" (11). Louisiana expressly includes family history in its definition of genetic information, while Florida, Missouri, Montana, Oklahoma, and Tennessee expressly *exclude* family histories from protection (12). Finally, a number of states expressly extend privacy protections to requests for genetic services (13).

Compelled Consent. Genetic privacy laws commonly prohibit health insurers and employers from requiring genetic testing as a condition of insurance or employment and from using any genetic information acquired in a discriminatory fashion. Privacy advocates have long argued that these protections are fairly meaningless if insurers and employers can persuade or pressure unsuspecting individuals into submitting to genetic testing or sharing genetic information, or obtain genetic information from other sources. Once a third party has possession of information, it is very difficult to police its use (14).

To address these problems, some states prohibit covered insurers and/or employers from even requesting genetic testing or genetic information (15). Vermont, while not prohibiting requests, requires that the person requesting genetic testing advise the individual to be tested that the results may become part of the individual's permanent medical record and may be material to his or her ability to obtain insurance benefits (16). At least four states prohibit covered insurers from seeking genetic information for any nontherapeutic purpose (17). A Delaware law passed in 1998 makes it unlawful for an employer to "intentionally collect" genetic information unless it can be demonstrated that the information is job related and consistent with business necessity or is sought in connection with a bona fide employee welfare or benefit plan, and laws in Oklahoma and Oregon provide that an employer may not "seek to obtain" genetic information concerning an employee or prospective employee (18). It is unclear whether these laws will have any effect on employer access to health insurance claims data, a major area of concern for privacy advocates.

Generic Privacy Laws. In some states, either by default or design, privacy protections are "generic" rather than "genetic," that is, such protections as exist apply to the general category of medical record or health information rather than to genetic information per se. Generic approaches avoid many of the definitional or line-drawing problems that arise in statutes that seek to single out genetic tests and information for protection. They also avoid stigmatizing genetic conditions by treating them differently (19). However, general laws tend to be fairly weak. Laws providing for the confidentiality of physician–patient communications limit disclosure by providers of health care, but they typically permit blanket releases of information to insurance companies and other third parties. Oregon is one of the exceptions. The state-mandated form for medical record release authorizations requires that sensitive information (HIV/AIDs-related records, mental health information, genetic testing information, and drug/alcohol-related information) be initialed by the patient or legal surrogate in order to be included in the release (20).

The first-generation Insurance Information and Privacy Protection Model Act, released by the National Association of Insurance Commissioners in 1981 and adopted in 10 states, covers areas such as pretext interviews (i.e., attempts to gain information involving a misrepresentation of or refusal to provide identity), disclosure of information practices to applicants and policyholders, the content of authorization forms, access to recorded personal information including medical record information, opportunities for correction of recorded personal information, limitations and conditions affecting disclosure of personal information, and penalties and remedies for violations. It does not block exchanges of sensitive information or describe specific steps insurers must take to minimize the risk of unauthorized disclosure. It also bars any action for defamation, invasion of privacy, or negligence for disclosure of personal information in a manner permitted under the state insurance code, even if the information is false, absent malice or willful intent to injure, and a provision of this nature is common even in states that have not adopted the model law.

A Connecticut law passed in 1999 to prohibit the sale of medical record information and restrict disclosure for marketing purposes, and a similar Maine law passed in 1998, are examples of the next generation of generic privacy laws (21). The Connecticut law has a broad definition of "medical-record information," which expressly includes information obtained from a pharmacy or pharmacist. (However, the definition excludes information that lacks personal identifiers *or* has been encrypted or encoded.) The law requires insurance-related entities that regularly collect, use or disclose medical record information to develop and implement written policies, standards and procedures for its management, transfer and security. These must include limiting access to persons who need the information in order to do their jobs, employee training, institution of disciplinary measures, periodic monitoring of employee compliance, and an additional layer of protection of "sensitive health information," which includes information regarding genetic testing such as the fact that an individual has undergone a test.

The National Association of Insurance Commissioners' Health Information Privacy Model Act, released in 1998 and enacted in at least one state (Montana) in 1999, also protects health information generally rather than singling out genetic information. In the employment arena, an innovative Minnesota law imposes a "job-relatedness" condition on all medical evaluation by employers, and laws in a few other states offer similar protections (22). Massachusetts has a general Privacy Act that authorizes the award of damages as well as equitable remedies for "unreasonable, substantial, or serious" interventions with a person's privacy (23).

Constitutional and Common Law Protections. Genetic privacy may also be protected, generically, under state

constitutions. Several state constitutions recognize a right to privacy (24). The case of *Norman-Bloodsaw v. Lawrence Berkeley Laboratory* concerned unauthorized testing of clinical and administrative workers for conditions including sickle cell trait. Because the employer, a research laboratory, was operated by a California state agency and a federal agency, the workers had claims under the state and federal constitutions, as well as federal antidiscrimination laws. The Ninth Circuit Court of Appeals found that constitutionally protected privacy interests encompass medical information and extend beyond unauthorized disclosure to reach collection of information by illicit means. According to the court, "the most basic violation possible involves the performance of unauthorized tests — that is, the nonconsensual retrieval of previously unrevealed medical information" that may be unknown even to the test subjects (25). The court concluded that genetic conditions are among the conditions that should enjoy the greatest level of protection. Mere consent to a general medical examination and the taking of samples for routine testing would not constitute authorization for these highly sensitive tests.

Common law may be another source of protection and redress. Invasion of privacy is a blanket term for a number of separate claims under common law. The two most relevant to the genetic context are (*1*) public disclosure of private facts, and (*2*) intrusion upon seclusion. In *Doe v. High-Tech Institute*, a case involving unauthorized testing for HIV/AIDS, a Colorado court concluded that a person has a privacy interest in a blood sample and in the medical information obtained from it, and that an additional, unauthorized test can be sufficient to state a claim for relief for intrusion upon seclusion (26). The elements of the claim are intentional intrusion (which need not be physical) upon the solitude or seclusion of another or his private affairs or concerns, and evidence that the intrusion would be highly offensive to a reasonable person. The court in *Doe* commented that in a case of unauthorized testing the intrusion is the interference with autonomy, namely the right to control important health decisions such as whether to undergo testing for a particular disease, condition, or genetic trait. Offensiveness depends on the nature of the specific test and the circumstances under which the test was performed. A law prohibiting a particular kind of testing without consent can serve as evidence of a significant privacy interest.

Ownership of Genetic Information. A few states have statutes that at least obliquely address the question of ownership, the fourth category of concern subsumed under privacy. A section of the Oregon genetic privacy act states that "an individual's genetic information and DNA sample are the property of the individual except when the information or sample is used in anonymous research," but the law also states that this provision "does not apply to any law, contract or other arrangement that determines a person's rights to compensation relating to substances or information derived from a sample of an individual from which genetic information has been obtained" (27). (The Delaware and New Jersey laws contain similar disclaimers.) A 1999 amendment to the Oregon law, effective August 2, 1999, to January 1, 2002, declares any research conducted in accordance with federal regulations for the protection of human subjects "anonymous" (28). Several states have laws that contain general declarations that genetic information is the "unique" or "exclusive" property of the person tested (29). In a much-discussed case, *Moore v. Regents of University of California*, a court rejected a patient's argument that commercial development of a cell line derived from his excised cells amounted to theft, while recognizing that patients have a right to information concerning their physicians' economic and research interests as part of the consent process (30).

Federal Laws

Certain members of Congress have taken an interest in genetic privacy, and bills addressing genetic privacy have been circulating for several years. Most of the genetic privacy bills introduced recently are limited to the health insurance context, although one addresses health insurance and employment. As of May 15, 2000, none of these bills had been enacted. Hence state law remains the primary source of protection for genetic information, although a diverse array of federal laws that are broader in scope may have some bearing on conduct in this area. For example, laboratories that perform genetic testing may be subject to the Clinical Laboratory Improvement Act, and attention is currently being given to confidentiality standards. In the employment context, the Americans with Disabilities Act (ADA) of 1990 regulates the timing and scope of medical examinations conducted by employers. The ADA also requires that employers keep information obtained from medical examinations confidential.

A number of federal laws, as well as bills or regulations under consideration, address the privacy of medical records or health information generally. The Privacy Act of 1974 regulates the handling of health-related information by federal agencies. The Health Insurance Portability and Accountability Act (HIPAA) of 1996 provides protection against discrimination based on genetic information and specifies how genetic information is to be handled for purposes of preexisting condition clauses in group health insurance. HIPAA also mandates further attention to privacy concerns. On August 12, 1998, pursuant to a HIPAA requirement, the Department of Health and Human Services (HHS) published a proposed rule establishing security standards for all electronic transactions involving health information. A proposed rule intended to protect the privacy of individually identifiable health information was published on February 3, 1999. The rule does not single out genetic information for heightened protection. It does provide a framework for regulation of the use and disclosure of individually identifiable health information by health plans, health care clearinghouses, and health care providers that engage in electronic transactions. The rule would not preempt stronger privacy protections at the state level (31). If Congress does not take action, the rule will be issued in final form sometime in 2000 or 2001.

Several bills have been introduced in Congress that would introduce more comprehensive privacy regulation at the federal level, including the "Health Care Personal

Information Nondisclosure Act of 1999" (S. 578), the "Medical Information Protection Act of 1999" (S. 881), the "Medical Information Privacy and Security Act" (S. 573), the "Medical Information Protection and Research Enhancement Act of 1999" (H.R. 2470), and the "Health Information Act" (H.R. 1941). As of November 30, 1999, none had been enacted. Major points of contention include the scope of protected health information (the Medical Information Privacy and Security Act appears to be unique in explicitly including tissue samples) and the level of anonymity required for exclusion, the acceptability of blanket authorizations, rights of minors, preemption of state laws offering greater protections, and the creation of an individual right to sue for violations. Interestingly all provide that protections continue after death.

The U.S. Constitution is yet another source of privacy protection for health-related information. In *Whalen v. Roe*, a case involving a New York law requiring physicians to send copies of prescriptions for certain classes of drugs to the state health department, the U.S. Supreme Court ruled that the privacy interests secured under the U.S. Constitution extend to medical record information (32). However, the court found that the steps taken by the state to protect confidentiality were sufficient to meet constitutional requirements. In *Norman-Bloodsaw*, discussed above, a federal appeals court ruled that the right to privacy extends to genetic testing.

LAWS REGULATING RESEARCH

Basic Regulatory Framework

The Federal Policy for the Protection of Human Subjects (often referred to as the "Common Rule") and other regulations governing human subject research are codified at Title 45 Part 46 of the Code of Federal Regulations. The Office for Protection from Research Risks (OPRR) within the National Institutes of Health has primary responsibility for implementation. Significantly, these regulations only apply to research conducted, supported, or subject to regulation by a department or agency of the federal government. Further, since the term "human subject" is limited to living individuals, biological materials and data derived from deceased persons are not covered. Also, to be covered, research must involve (*1*) intervention or interaction with the individual or (*2*) *identifiable* private information, meaning that the identity of the subject is or may readily be ascertained by the researcher or associated with the information. If neither condition is met, research would not be considered human subject research. OPRR has taken the position that information is identifiable where codes can be broken with the cooperation of others (33). Some research, while covered, may be eligible for an exemption from regulatory requirements. The categories of exempt research include research involving the collection or study of existing data or specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified directly or indirectly through identifiers.

For research that falls within the scope of the regulations and does not qualify for an exemption, there are two basic requirements: approval must be given by an Institutional Review Board (IRB) and proper informed consent must be obtained from participants. One requirement for IRB approval is that "when appropriate" adequate provisions are made to protect the privacy of subjects and maintain the confidentiality of data. The consent requirement has aroused more controversy. Traditionally tissues, blood, and other biological materials used for research have been stored and studied without explicit consent based on a waiver provision in the Common Rule. Section 46.116(d) states that informed consent for research can be altered or waived if an IRB finds and documents that (*1*) the research involves no more than minimal risk to the subjects, (*2*) the waiver or alteration will not adversely affect the rights and welfare of the subjects, (*3*) the research could not practicably be carried out without the waiver or alteration, and (*4*) whenever appropriate, the subjects will be provided with additional pertinent information after participation. With the advent of genetic research, some have argued that the risk associated with the use of identified or identifiable data or specimens is no longer minimal and consent should be obtained, particularly where protocols propose sharing information with the source or a third party (34).

In August 1999 the National Bioethics Advisory Commission issued recommendations for the interpretation or modification of existing federal regulations relating to the use of human biological materials in research. The Commission found that use of specimens for purposes other than the purpose for which they were originally collected raises the strongest privacy concerns. Pursuant to the recommendations, even research limited to manipulation of existing specimens, if identified or coded (identifiable), would be judged greater than minimal risk and therefore ineligible for a consent waiver unless an IRB were to find that a particular study adequately protects the confidentiality of personal information and incorporates an appropriate plan for whether and how to reveal findings to the donors or their physicians. Consent forms for research would provide a menu of options; future research uses involving unidentified and unlinked samples would generally be accorded similar treatment, as would identified and coded samples (33).

Turning to the state level, states that have genetic privacy laws often exempt research activities, but only if there is some level of anonymity. Requirements vary, ranging from a requirement that the genetic information used for research be free of all identifiers, to a requirement that the only identifier be a code, to a requirement that the identities of donors not be released to researchers. States may parse the research exemption as an exemption from the mandate to obtain informed consent before performing any genetic testing or as an exemption to the mandate to obtain authorization for disclosure of genetic information to a third party. The Missouri law exemplifies the latter approach: genetic information is confidential and cannot be disclosed without written authorization except for health research conducted in accordance with the Common Rule or health research "using medical archives or databases in which the identity of individuals is protected from disclosure by coding or encryption, or by removing all

identities" (35). Florida's genetic privacy law contains no exemption for research (36).

Like much biological materials research, medical records research has traditionally been conducted without the knowledge of patients, let alone their informed consent. State medical record confidentiality laws frequently permit access to records for research purposes, so long as researchers maintain the confidentiality of identifiable information (37). Minnesota generated much controversy when it passed a law imposing unusually strict requirements on external researchers seeking access to medical records. An executive of a biotechnology company that conducts a large proportion of its clinical trials at the Mayo Clinic reports that this state law has dramatically reduced the medical records available for research, from 97 percent when the law was passed to 70 percent in 1999 (38).

The privacy bills currently under consideration by Congress would permit disclosure of protected health information for health research that satisfies certain requirements, typically compliance with the Common Rule (or the equivalent for the Food and Drug Administration, FDA) or, in the case of existing information, review by an IRB or IRB-like entity. The Medical Information Privacy and Security Act would make all health research subject to Part 46 of Title 45 of the Code of Federal Regulations, whether federally funded or not. The proposed rule on health information privacy, referred to above, would subject all medical records research to scrutiny by an IRB or privacy board, but only with respect to certain privacy criteria (31).

The federal and state laws and regulations that protect research subjects either by requiring consent or by requiring anonymity for health information and biological materials used in research do not address the potential for abuse of information about specific populations. Research involving biological materials that have been stripped of individual identifiers should not produce information that will result in direct harm to particular donors, but it may reveal sensitive information about the groups to which the donors belong.

DNA Banking for Research Purposes

Banking of biological materials for research, and storage of related information, is largely unregulated. DNA banks range from the collections of individual researchers to the extensive stores of the U.S. Centers for Disease Control and Prevention. State public health agencies may also have extensive collections of biological materials and health information (e.g., cancer registries and "Guthrie cards," specimens of infant's blood on filter paper), subject to varying levels of privacy protection. At least four states require that genetic information be destroyed at the completion of a research study or upon withdrawal of the person from the study, unless retention is explicitly authorized. In a controversial move, the Michigan Commission on Genetic Privacy has recommended that the state permanently preserve blood samples of newborns, which were originally collected to screen for rare congenital disorders, for use in research conducted in accordance with the Common Rule (unless parents elect to opt-out of research) (39).

Research and Commercial Development

The potential for commercial development based on genetic research has added another layer to research informed consent. While the court in *Moore* refused to recognize a patient's property rights in a cell line derived from his tissue, it did find that physicians must disclose personal interests unrelated to the patient's health that may affect their judgment and failure to do so may give rise to a cause of action for performing a medical procedure without informed consent or breach of fiduciary duty. In response to the *Moore* ruling, consent forms have been expanded to include explicit permission to retain cells for development with potential commercial value. It is not clear that current laws adequately address issues of commercialization. The Uniform Anatomical Gift Act, adopted in some form by the majority of states, gives the donor the authority to decide particular use of body tissue, while the National Organ Transplantation Act prohibits the sale of human tissue and organs for transplantation. Neither law was designed to address the commercialization of genetic material (40).

Law and Chilling Effects on Research

Some have argued that the new emphasis on informed consent for research use of biological materials and medical records will have a chilling effect on research (41,42). The major concern is that the ethical benefits of obtaining informed consent (e.g., demonstrating respect for persons and for privacy as personal choice) are far outweighed by the burdens to research. Consent requirements mean added time, expense, and paperwork, and selective participation may introduce significant bias into research, detracting from its value. A further concern is the lack of uniformity in state laws. Major research projects frequently cross state lines in order to obtain adequate accrual of subjects.

Research may also suffer due to uncertainty concerning confidentiality protections for data in the possession of researchers. A number of celebrated products liability cases from the 1980s showed the willingness of courts to compel production of research data. For example, in *Deitchman v. E.R. Squibb*, Dr. Arthur L. Herbst was required to produce research data linking adenocarcinoma of the genital tract and exposure in utero to diethylstilbestrol (DES); Herbst was not a party to the case nor was he even being called to testify (43). A provision of the Omnibus Consolidated and Emergency Supplemental Appropriations Act (P.L. 105-277) passed in October 1998 requires that federal agencies make data produced under awards to nonprofit organizations available to the public through Freedom of Information Act procedures. To be sure, individual identifiers are removed before data are released to litigants or requestors, but this does not allay the concerns of researchers and privacy advocates. With expanded capacities to manipulate data, redaction to eliminate obvious identifies may not adequately protect the privacy of individuals and groups. Under the

Public Health Service Act, Certificates of Confidentiality are available for certain types of sensitive research including genetic research. Certificates of Confidentiality, however, only protect against forced disclosure of identifying information about individuals, not forced disclosure of aggregate data.

Other Emerging Issues

The fecundity of biological materials raises two further concerns. Particular samples may have potential for usefulness in future research unanticipated at the time of collection. Should the individual be recontacted for permission to do further research on the sample or is a generic consent to use of the sample in future research sufficient? More pressing, with advancements in genetics, the amount of useful medical information that a sample can yield continues to expand. What is the responsibility of a researcher to inform the donor of the sample if medically significant information is discovered? Since most state laws do not require total anonymization of samples used in research, it is theoretically possible to identify the donor. With the possibility of identification may come a responsibility to inform, but the potential psychological harms and risk of discrimination that accompany unsolicited disclosure or "inflicted insight" must be carefully evaluated. Cases addressing the obligations of treating physicians to disclose information to potentially affected family members (discussed below) appear to be the closest analogue. Subjects might be queried concerning their desire for recontact for further studies and for contact with medically significant information, yet questions of interpretation will remain.

LAWS REGULATING DNA BANKS AND DNA DATABASES

Large collections of DNA samples, and records reflecting the results of analysis of those samples, are assembled primarily for one of three purposes: law enforcement, identification of human remains (chiefly in the military context), and medical treatment and research. This section addresses collections of DNA samples (or "banks") and DNA records (or "databases") related to the first two purposes. Interest in this area has been focused on the first two categories of privacy concern, bodily integrity and confidentiality. In particular, questions have arisen concerning the constitutionality of coercive extraction of DNA samples and subsequent testing, and the adequacy of measures to ensure confidentiality and protect samples and information from misuse by third parties. Although their primary goal is identification of perpetrators of crimes, DNA banks and databases maintained by law enforcement agencies may also contain DNA records of unidentified persons, and, in a few states, relatives of missing persons. Texas also authorizes inclusion of DNA records of "a person at risk of becoming lost" (44).

DNA Collection by Law Enforcement

The collection and analysis of DNA samples for law enforcement purposes is concentrated at the state level, with some coordination through the Federal Bureau of Investigations. All 50 states have passed authorizing legislation. States must meet certain federal guidelines in order to participate in data exchange through the Combined DNA Index System or CODIS, established under the DNA Identification Act of 1994 (45). Recently, controversy has centered on the scope of state collection efforts. State laws usually authorize collection of a sample (blood or a buccal swab) upon conviction of a crime and/or as a condition of parole or release from custody. Many state laws cover juvenile offenders (or children adjudicated delinquent) as well as adult offenders, and they contain no special provision for the expungement of DNA records or samples upon attainment of majority.

Initially the focus was on individuals convicted of sex crimes and other serious felonies, offenses associated with high rates of recidivism and biological evidence. However, at least 14 states have added burglary, a nonviolent property crime, to the list of covered offenses. Further, in 1999, several states considered legislation that would permit the creation of DNA records for felony suspects at the time of arrest, a practice contemplated under a Louisiana law that has been enacted but is not yet effective (46). (Significantly the DNA Identification Act limits entries into CODIS to DNA identification records of persons convicted of crimes, analysis of DNA samples recovered from crime scenes, and analysis of DNA samples from unidentified human remains.) Such an expansion of scope may implicate group as well as individual privacy interests. Given that certain minority groups are disproportionately represented in the prison population, and therefore in DNA databases, these technologies may be more effective in "fingering" members of those groups. This problem is likely to be compounded if all arrestees are sampled. The Attorney General has asked the National Commission on the Future of DNA Evidence to study the legality of taking samples upon arrest.

Inmates challenging state laws on Fourth Amendment unreasonable search and seizure grounds have lost with a fair degree of consistency. For example, in *Boling v. Romer*, the Tenth Circuit Court of Appeals found that obtaining and analyzing the DNA of an inmate convicted of a sex offense is a search and seizure raising Fourth Amendment concerns, but that the search and seizure was a reasonable one in light of an inmate's diminished privacy rights, the minimal intrusion of saliva or blood tests, and the legitimate government interest in the investigation and prosecution of unsolved and future crimes by the use of DNA in a manner not significantly different from the use of fingerprints, that is, for purposes of identification (47). First Amendment free exercise, Fifth Amendment self-incrimination and due process, Eighth Amendment cruel and unusual punishment, Fourteenth Amendment equal protection, and ex post facto clause challenges have also failed. The balancing tests may come out differently, however, when the subjects of collection and testing have yet to be convicted of any crime. It is significant that most state laws provide for expungement of samples and information when a conviction is overturned. The Rhode Island law also provides that all identifiable information and samples be destroyed upon official proof that the subject has been deceased for at least three years (48).

At present the biggest deterrent to expansion of DNA sampling by law enforcement agencies is a lack of capacity to analyze samples, rather than any fear of lawsuits. As of June 1999 the backlog of samples awaiting analysis exceeded 500,000 nationwide (49). In response, states are sending samples to private laboratories, a practice that heightens concerns about security and potential for misuse of data.

The DNA profiles entered in CODIS databases currently consist of 13 DNA loci or markers. These DNA markers were selected for their variability, rather than their intrinsic information value (e.g., links to disease or personality traits). Hence advocates for the use of DNA databases in law enforcement believe the potential for abuse is limited. In this regard it is important to distinguish between the samples stored in DNA banks, which can be used to generate other, more sensitive information so long as they are in existence, and the profiles stored in DNA databases. Given the interest in behavioral genetics, one can imagine the eagerness with which some researchers might pursue access to genetic information concerning this particular subset of the population.

State laws address confidentiality concerns in a number of ways. They typically contain a statement that DNA profiles are not public records. Some laws require custodians of samples or records to implement security measures. The Rhode Island law is one of the more detailed in this respect, requiring that an encryption code be used for access to records, that DNA samples be securely locked with access only by director of health department and head of laboratory, and that all identifiers be removed when samples are provided to third parties for creation of DNA records (50). Some state statutes require the relevant agency to develop privacy standards for laboratories under contract. The California law declares that the computer software and database structures, as well as the data, are confidential (51).

Most state laws mimic the language of the DNA Identification Act limiting disclosure of DNA analysis to criminal justice agencies for law enforcement purposes, in judicial proceedings, and, to defendants for criminal defense purposes. Where personally identifiable information is removed, disclosure is also permitted for population statistics databases, for identification research and protocol development purposes, and for quality control purposes. While the reference to population statistics databases might be interpreted broadly, in this context the term refers to analysis of the frequency of occurrence of genetic characteristics in local populations, a type of research closely related to the identification objective for which the DNA banks and databases have been established. Even so, a few states seem to take a fairly permissive view. A number allow dissemination of statistical or research information if there are no identifiers, and the Alabama law permits the state's DNA population statistical database to be used to provide data relative to the causation, detection and prevention of disease or disability and "to assist in other humanitarian endeavors including, but not limited to, educational research or medical research or development" (52). The Massachusetts law allows release of DNA records for the purpose of assisting in the identification of human remains from mass disasters, assisting the identification and recovery of missing persons, and "advancing other humanitarian purposes" (53). At the other end of the spectrum, Indiana, Rhode Island, and Wyoming expressly prohibit the use of the state collections for the purpose of obtaining information about human physical traits or predisposition for disease (54).

DNA Collection by the Military

The Department of Defense is compiling the world's largest employer-held DNA bank; all inductees and all active duty and reserve personnel are required to provide samples to the military's DNA Specimen Repository. The purpose of the repository is to permit identification of remains, with DNA extracted from samples only as needed. However, concerns have been expressed about possible diversion of the repository to other uses, such as criminal investigation (55). In 1995 a federal district court upheld the sampling program against a challenge on constitutional and other grounds, although the judgment was vacated on mootness grounds after the plaintiffs had been honorably discharged (56).

LAWS REGULATING FAMILY RELATIONS

Disclosure to Potentially Affected Family Members

In the clinical context, a major question has been whether health care providers have a legal and/or ethical duty to disclose genetic information to family members for whom the information may have medical or reproductive implications. There is little direct statutory guidance in this area, and where statutes exist, they tend to permit disclosure rather than mandate it.

The Uniform Health Care Information Act (National Conference of Commissioners on Uniform State Laws, 1985) recognizes several exceptions to the general obligation to keep information confidential. One exception permits, but does not require, a health care provider to disclose information from a patient's record where the health care provider has reason to believe that disclosure will avoid or minimize imminent danger to the health or safety of the patient or another individual. The exception developed out of common law precedents permitting or requiring disclosure in situations where a patient had an infectious disease and the facts were such that the patient could be expected to infect others absent disclosure, or where a patient threatened harm to self or others. The Act has been adopted in only two states, Montana and Washington, and Montana declined to include this provision (57). Other statutes that limit disclosure of health care information fail to address this issue directly, and it has fallen to the courts to define the scope of a legal duty to disclose genetic information.

Judges have grappled with this issue in two cases, *Pate v. Threlkel* and *Safer v. Estate of Pack*, with somewhat different results (58). In interpreting the results, it is important to recognize that both cases were decided at a stage in the proceedings where the court was required to accept as correct the plaintiffs' assertions concerning

the standard of care required of physicians during the relevant time period (and other factual matters). Also the decisions in these cases are binding only in the respective jurisdictions, Florida and New Jersey. Other courts will accord them weight based only on the persuasiveness of the reasoning relative to the law in their home states.

Pate was decided first, with the basic facts as follows: Heidi Pate discovered that she had medullary thyroid carcinoma, an autosomal dominant disorder, three years after her mother received treatment for the same disease. She sued her mother's physicians and their employers, arguing that the physicians had a duty to warn her mother of the risk of genetic transmission and to recommend testing of any children. The Florida Supreme Court ruled that if the standard of care was to warn a patient of the genetically transferable nature of a condition, as Pate alleged, then the intended beneficiaries of the standard would include the patient's children as well as the patient. In other words, the patient's children would be entitled to recover for a breach of the standard of care. However, in light of state laws protecting the confidentiality of medical information, the court found no requirement that a physician warn a patient's children. Rather, the court found that in any circumstances in which the physician has a duty to warn of a genetically transferable disease, that duty will be satisfied by warning the patient. The court believed a more expansive standard would be burdensome and unmanageable.

Safer involved similar facts. Donna Safer's father was treated over an extended period of time for colon cancer associated with adenomatous polyposis coli, an autosomal dominant disorder. Almost two decades after his death, Safer was diagnosed with metastatic colon cancer associated with adenomatous polyposis coli. Safer then sued the estate of George Pack, her father's physician, for Pack's failure to warn of the risk to her health. Two additional facts, taken as true for purposes of the court's decision, were significant. First, Safer's mother testified that on at least one occasion she asked Pack whether what he referred to as an "infection" would affect her children and was told not to worry. Second, Safer contended that careful monitoring of her condition would have provided an opportunity to "avoid the most baneful consequences of the condition," a consideration giving force to her argument that the standard of care in the circumstances was to warn any children at risk.

Examining the precedents, the New Jersey Appeals Court found no essential difference between this case with its "genetic threat" and traditional duty-to-warn cases involving the menace of infection or threat of physical harm. The court concluded that a duty to warn in the genetics context would be quite manageable, commenting that those at risk are easily identified. The court noted the potential to avert or minimize substantial future harm by a timely and effective warning to potentially affected persons. The court failed to state how the duty to warn might be discharged, especially in cases involving small children. The relation of this ruling to confidentiality protections was considered only in the court's concluding speculations about possible complications, such as the existence of an instruction from the patient not to disclose information to family members. The court offered no resolution of this dilemma.

Some states do have statutes that address the question of disclosure for the benefit of family members in very narrow circumstances. For example, a section of the California Welfare and Institutions Code concerning involuntary commitment provides that information pertaining to the existence of a genetically handicapping condition may be released to a qualified professional for purposes of genetic counseling for a blood relative upon the request of the blood relative (59). This can happen either with the consent of the patient or after reasonable attempts have been made over a two-week period to get a response from the patient.

It is important to remember that broad dissemination of genetic information is not an unalloyed benefit to potentially affected family members. In addition to possible psychological harms, family members may face discrimination if genetic information finds its way into their medical records or becomes part of their knowledge base and so must be disclosed on applications for insurance. This could happen without their cooperation where a health care provider shares information without first verifying that family members with to receive it, or where an insurer or other third party stores information concerning more than one family member and fails to prevent information flow. An awareness of this problem appears to explain a New York law that prohibits any person in possession of information derived from a genetic test from incorporating that information into the records of a nonconsenting individual who may be genetically related to the tested individual (60).

Testing of Children

The right of parents to order genetic testing of their children has also provoked some discussion. The issue is particularly vexing where the testing is for an adult-onset condition that cannot be prevented, ameliorated or cured by any action taken during childhood. In such cases it is hard to argue that testing confers any benefit on the child, or any benefit on the parents such as the ability to plan for burdens likely to affect the family unit (as might be the case with an incurable condition manifesting in childhood). The general rule is that parents control medical decision making for their children. The Delaware genetic privacy act expressly states that it "does not alter any right of parents or guardians to order medical and/or genetic tests of their children" (61). An Illinois law addresses a situation in which testing is requested by or ordered for a child and a question arises concerning the proper recipient of the results. The statute requires a health care provider who orders a genetic test for a minor to notify the minor's parent or legal guardian of the results of the test, but only if the provider determines that notification would be in the best interest of the minor and has first sought unsuccessfully to persuade the minor to give the notice (62).

Paternity Determination

The use of genetic testing in the determination of paternity is most analogous to the use of genetic testing in

law enforcement. As in the law enforcement context, testing aims at identification rather than production of information about health or behavior, and similar privacy interests are implicated, namely the interest in being free of unwanted bodily invasion, and the interest in preventing access to or disclosure of personal information. If an individual refuses to submit to paternity testing, the result may be a contempt citation or a presumption that the results would be adverse to the individual. Laws governing genetic testing to determine paternity typically provide that the results of testing are confidential and cannot be disclosed to third parties absent exceptional circumstances. Sample retention is not always addressed. Some laws do provide that if a man is found not to be the father of a child in a paternity proceeding, the man's genetic material must be destroyed (63).

Adoption

Adoption is yet another context in which genetics, privacy concerns, and the law come together. A critical issue here is the extent of adoptive parent and adoptee access to genetic information. State adoption laws typically require preparation of a report of the complete family medical and social history of a candidate for adoption (so long as parental identity is not disclosed), including identification of any known genetic disorders. Ten states require information concerning extended family if available (64). Generally, the adoptive parents, and the adopted child upon attainment of majority, are allowed access to this information. In a few states, where an adopted child has died, genetic information (and other medical information) is available to the spouse of the adopted child if he or she is the legal parent of the adopted child's progeny, and also to any progeny of the adopted child age 18 or older.

Some states create voluntary registries that allow nonidentifying genetic information to flow back *to* birth parents (e.g., notice that the child has been diagnosed with sickle cell anemia, meaning that the birth parents are carriers of sickle cell trait), as well as permitting transfer of additional information *from* birth parents as it becomes available. In Arizona, other biological children of the birth parent can also access nonidentifying information in adoption records upon request (65). In several states, a certified statement from a physician explaining why genetic or other critical medical information should be communicated to adoptive parents or an adopted child, or an adopted child's genetic parent or sibling, is the trigger for an effort by the registry to notify the affected person(s) that the nonidentifying information is available from the registry (66). North Dakota places information exchange within the discretion of the child-placing agency (67).

Opening of sealed adoption records that include identifying information is usually permitted only by court order upon a showing of "good cause." Proceedings to open records have been brought by adopted children seeking to contact genetic relatives as potential donors of tissue or bone marrow for life-saving transplants, or for other health-related purposes. In such cases courts typically apply a balancing test, weighing the interests of the adopted child, the adoptive parents, the biological parents, and society. In *Golan v. Louise Wise Services*, New York's highest court reversed a lower court order that would have permitted an adoptee, seeking genetic information for evaluation of his heart condition, access to the identities of his biological parents. The court found that consideration of the adoptee's well-being alone in cases involving medical problems with genetic implications would "swallow" the state's strong policy against disclosure. The court suggested correspondence through a guardian ad litem as an alternative to unsealing adoption records (68).

Other scenarios put forward by commentators, in which privacy interests may be at odds with other interests, are more speculative, and appear not to be directly addressed under current law. For example, there are concerns about excessive testing of children prior to adoption. Genetic testing or genetic information might also be demanded of prospective adoptive parents, as bearing on their suitability for parenthood. Concerns have also been expressed about the use of genetic information as ammunition in child custody disputes, as affecting the likelihood that a parent will "be there" for a child in the future (10,69).

Postmortem Testing for the Benefit of Family Members

Several states authorize sharing of genetic information with family members once a person is dead. The next logical step is genetic testing of the dead to obtain information desired by family members. Under normal circumstances the next-of-kin usually controls disposition of the body and has the right to consent to or refuse an autopsy (or the use of samples in research). An argument could be made that by extension family members have the right to order genetic testing, although laws prohibiting genetic testing without consent with only limited exceptions might be cited as evidence against such a right. In at least one case, reported in the media, a daughter was given access to a sample of her deceased father's blood for paternity testing. Her mother sued the hospital to try to block testing and lost (70). There appear to be few laws directly addressing this question. A notable exception is a New York law that expressly authorizes genetic testing on specimens from deceased persons if informed consent is provided by the next of kin (71). (Autopsies may also raise the unsolicited disclosure problem, where a pathologist obtains genetic information that may be material to family members in the course of an autopsy and must determine whether to disclose this information to family members.) The Native American Graves Protection and Repatriation Act addresses the issue by stipulating that Native American remains in the possession or control of federal agencies are the property of lineal descendants or tribes and must be repatriated upon request (72).

MISCELLANEOUS

Genetic information is increasingly sought in the context of civil litigation. Defendants in personal injury lawsuits may be eager to prove that injuries resulted from the plaintiffs' genetic defects rather than their own negligent

conduct. As noted above, state laws may declare that genetic information is privileged and hence protected from routine discovery in the investigational phase of a civil proceeding. However, a judge may order testing or disclosure of information if persuaded of its relevance. For example, a defendant in a lawsuit arising out of an automobile accident sought to compel genetic testing of the plaintiff for Huntington's disease, as a possible causal factor, and the court ordered the testing over the plaintiff's objections (73). Colorado law creates a barrier to recovery for injury arising from genetic counseling and screening, prenatal care, labor, delivery, or postnatal care where it can be established that the injury was the result of a genetic disease or disorder (74). A provision of the law expressly permits discovery of medical information concerning the plaintiff and makes this information admissable as evidence at trial. In addition the law allows discovery of medical information relating to genetic siblings, parents, and grandparents of the plaintiff, if the defendant cannot secure voluntary releases and persuades the court of the possible relevancy of the information.

BIBLIOGRAPHY

1. A.L. Allen, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 31–59.
2. E.g., Del. Code Ann. tit. 16, §1224; Nev. Rev. Stat. §629.171; N.J. Stat. Ann. §10 : 5-47; Or. Rev. Stat. §659.720; S.C. Code Ann. §38-93-30; Tex. Labor Code Ann. §21.403, Tex. Ins. Code art. 21.73, §4, and Tex. Rev. Civ. Stat. Ann. art. 9031, §3 (Vernon's).
3. Tex. Ins. Code, art. 21.73.
4. 29 U.S.C. §§1140 et. seq.
5. N.Y. Civ. Rights Law §79.1.
6. E.g., Colo. Rev. Stat. Ann. §10-3-1104.7(12); Nev. Rev. Stat. Ann. §629.201; N.M. Stat. Ann. §24-21-6.
7. La. Rev. Stat. Ann. tit. 22, §213.7(F).
8. Ala. Code §27-53-1 et seq.
9. Ill. Stat. Ann. ch. 410, para. 513/15.
10. L.N. Geller et al., *Sci. Eng. Ethics* **2**(1), 71–88 (1996).
11. Conn. Gen. Stat. Ann. §46a-60(a)(11) (West).
12. La. Rev. Stat. Ann. tit. 22, §213.7(A)(8); Fla. Stat. Ann. §627.4301; Mo. Ann. Stat. §375.1300 (Vernon's); Mont. Code Ann. §33-18-901; Okla. Stat. Ann. tit. 36 §3614.1; Tenn. Code Ann. §56-7-2706.
13. E.g., N.M. Stat. Ann. §24-21-2.
14. M.A. Rothstein, *Am. J. Law, Med. Ethics* **26**, 198–204 (1998).
15. E.g., Conn. Gen. Stat. Ann. §46a-60(a)(11); Fla. Stat. Ann. §627.4301(2)(b); Ga. Code Ann. §33-54-3; Haw. Rev. Stat. Ann. §431 : 10A-118; Ill. Stat. Ann. ch. 410 §513.20); Ind. Code Ann., §27-8-26-5; 1999 Kansas Sess. Laws S.B. 22; Ky. Rev. Stat. Ann. §304.12-085; 1999 Maryland Laws ch. 50 (S.B. 774); Minn. Stat. Ann. §72A.139; Mo. Ann. Stat. §375.1303; Ohio Rev. Code Ann. §1751.65; R.I. Gen. Laws §27-18-52; Tenn. Code Ann. §56-7-2704.
16. Vt. Stat. Ann. tit. 18, §9332.
17. E.g., Cal. Health & Safety Code §1374.7(b); Ga. Code Ann. §33-54-4; Mont. Code Ann. §33-18-904; Colo. Rev. Stat. Ann. §10-3-1104.7(3)(b).
18. Del. Code Ann. tit. 19, §711; Okla. Stat. Ann. tit. 36, §3614.2; Or. Rev. Stat. §659.036.
19. M.A. Rothstein, *Am. J. Law Med.* **24**, 399–416 (1998).
20. Or. Rev. Stat. §192.525.
21. 1999 Conn. Pub. Acts No. 99-284; Me. Rev. Stat. Ann. tit. 22, §1711-C.
22. Minn. Stat. Ann. §363.01-.20.
23. Mass. Gen. Laws Ann. ch. 214 §1B.
24. E.g., Constitution of the State of California, Art. I, §1; Florida Constitution, Art. I, §23; Constitution of the State of Illinois, Art. I, §§6, 12; Constitution of the State of Louisiana, Art. I, §5; Constitution of the State of Montana, Art. II, §10.
25. 135 F.3d 1260 (9th Cir. 1998).
26. 972 P.2d 1060 (Colo. Ct. App. 1998).
27. Or. Rev. Stat. §659.715; Del. Code Ann. §1225; N.J. Stat. Ann. §10 : 5-46.
28. 1999 Oregon Laws Ch. 921 (S.B. 937).
29. E.g., Colo. Rev. Stat. Ann. §10-3-1104.7(1)(a); Fla. Stat. Ann. §760.40; Ga. Code Ann. §33-54-1; La. Rev. Stat. Ann., tit. 22, §213.7(E).
30. 793 P.2d 479 (Cal. 1989), *cert. denied*, 499 U.S. 936 (1991).
31. Standards for Privacy of Individually Identifiable Health Information, Proposed Rule, 64 *Fed. Reg.* 59,918 (1999) (to be codified at 45 CFR Parts 160 through 164).
32. 429 U.S. 589 (1977).
33. Report and Recommendations of the National Bioethics Advisory Commission, *Research Involving Human Biological Materials: Ethical Issues and Policy Guidance*, vol. 1, National Bioethics Advisory Commission, Rockville, MD, 1999.
34. E.W. Clayton et al., *J. Am. Med. Assoc.* **274**, 1786–1792 (1995).
35. Mo. Ann. Stat. §375.1309 (Vernon's).
36. Fla. Stat. Ann. §760.40.
37. *The State of Health Privacy: An Uneven Terrain*, Health Privacy Project, Georgetown University Institute for Health Care Research and Policy, Washington, DC, 1999.
38. *Bluesheet* **42**, 18 (1999).
39. Michigan Commission on Genetic Privacy and Progress, *Final Report and Recommendations*, February 1999.
40. R.G. Harman, *J. Legal Med.* **14**, 463–477 (1993).
41. L.J. Melton III, *N. Engl. J. Med.* **337**, 1466–1470 (1997).
42. W. Grizzle et al., *Archiv. Pathol. Lab. Med.* **123**(4), 296–300 (1999).
43. 740 F.2d 556 (7th Cir. 1984).
44. Tex. Gov. Code §411.142.
45. 42 U.S.C. §13701 et seq.
46. *DNA Detection of Sexual and Violent Offenders Law*, La. Rev. Stat. Ann. §15 : 601 et seq., suspended by H.C.R. 40 (1999).
47. 101 F.3d 1336 (10th Cir. 1996).
48. R.I. Stat. Ann. §12-1.5-13.
49. R.S. Wilson, L. Forman, and C.H. Asplen, *Corrections Today* **61**(3), 20–22 (1999).
50. R.I. Gen. Laws §12-1.5-10.
51. Cal. Penal Code §299.5.
52. Ala. Code §36-18-31.
53. Mass. Gen. Laws Ann. ch. 22E, §10.
54. Ind. Code Ann. §10-1-9-18; R.I. Gen. Laws §12-1.5-10; Wyo. Stat. §7-19-404.
55. J.E. McEwen, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 231–251.

56. *Mayfield v. Dalton*, 901 F.Supp. 300 (D. Haw. 1995), *vacated*, 109 F.3d 1423 (9th Cir. 1997).
57. Mont. Code Ann. §50-16-529; Wash. Rev. Code Ann. §70.02.050.
58. 661 So.2d 278 (Fla. 1995); 291 N.J. Super. 619, 677 A.2d 1188 (N.J. Super. Ct. App. Div. 1996), *cert. denied*, 683 A.2d 1163 (N.J. 1996).
59. Cal. Welf. & Inst. Code §5328.
60. N.Y. Civ. Rights Law §79.1.
61. Del. Code Ann. §1226.
62. Ill. Stat. Ann. ch. 410, para. 513/30.
63. E.g., Mich. Stat. Ann. §722.716a.
64. L.B. Andrews, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 255–280.
65. Ariz. Rev. Stat. Ann. §8-129.
66. E.g., Minn. Stat. Ann. §259.83; Miss. Code Ann. §93-17-205; Va. Code Ann. §63.1-236.02; Vt. Stat. Ann. §tit. 15A, §6-104.
67. N.D. Cent. Code §14-15-16.
68. *Golan v. Louise Wise Services*, 69 N.Y.2d 343, 507 N.E.2d 275 (N.Y. Ct. of Appeals 1987).
69. C. Lerman et al., *J. Health Care Law Policy* **1**, 353–372 (1998).
70. D. Nelkin and L.B. Andrews, *Chron. Higher Educ.* **45**, B6 (1999).
71. N.Y. Civ. Rights Law §79.1. Or. Rev. Stat. §659.715.
72. Native American Graves Protection and Repatriation Act, 25 U.S.C. §3001–30013.
73. M.A. Rothstein, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 451–495.
74. Colo. Rev. Stat. Ann. §13-64-502.

See other GENETIC INFORMATION, ETHICS and GENETIC INFORMATION, LEGAL entries.

GENETIC INFORMATION, LEGAL, GENETICS AND THE AMERICANS WITH DISABILITIES ACT

MARY R. ANDERLIK
University of Houston
Houston, Texas

OUTLINE

INTRODUCTION

The Americans with Disabilities Act (ADA) is one of the landmark pieces of legislation of the twentieth century (1). Its purpose is to "provide a clear and comprehensive national mandate for the elimination of discrimination against individuals with disabilities" (2). Many people affected by genetic conditions fear discrimination, especially in the areas of employment and insurance. The ADA contains provisions that address disability discrimination affecting the conduct of employers and insurers. However, the protections of the ADA may be of little help to those who believe they have suffered discrimination owing to a genetic characteristic. This is so owing to the restrictive definition of disability in the ADA and other features such as a provision that shelters traditional insurance underwriting practices from challenge under the ADA.

Structural Reasons for Discrimination

It may be helpful to review some of the reasons for genetic discrimination, and the available evidence concerning its prevalence, before proceeding to a discussion of the statute itself. Discussions of genetic discrimination tend to focus on the employment and insurance contexts. This is so because given the current arrangements for financing health care, both employers and insurers have strong financial incentives to discriminate on the basis of genetic information (and other information concerning present or future health status) in order to control costs.

Employment

In the United States, employers frequently play a role in obtaining health insurance for employees. Where an employer assumes responsibility for paying some part of the premiums for experience-rated health insurance or chooses to self-insure, employee health problems have a direct effect on the employer's financial performance. Experience rating means that premiums are set based on the claims history of persons covered under the employer's group policy. For small employers, in particular, one employee with an illness that is expensive to treat can send premiums sky-rocketing, increasing costs to the employer and to other employees. Where an employer chooses to self-insure, the employer bears the burden of any

expenditures for covered medical care and captures the benefit from any reduction in such expenditures. In 1997, 13 percent of all employers had self-funded health plans, and 56 percent of firms with 500 or more employees had self-funded health plans (3). Strategies that self-insured employers might be tempted to pursue to hold down health care related costs include cutting salaries or increasing employee cost-sharing to offset any increases in health care costs, weeding out workers likely to have health problems through hiring and firing processes, identifying conditions that are likely to prove costly and excluding them from coverage (a strategy that is more effective if the health risks of actual employees are known), and dropping health coverage entirely (4).

In addition employers may seek to exclude individuals with above-average susceptibility to a toxin from certain jobs out of paternalistic concern or fear of liability. Another response would be to increase protections (e.g., reduce exposures) for all workers. In the employment context, it is important to distinguish between screening and monitoring. Screening refers to efforts to test applicants or employees for conditions that may render them more susceptible to harm from workplace substances than the average person. The target is the individual, and the goal is to improve productivity and lower workers' compensation and health insurance costs. Monitoring refers to the performance of periodic examinations of employees to identify and assess changes. The target is the active workforce and the goal is to identify workplace risks that can be reduced through implementation of prevention programs (5). Finally, employers may wish to exclude persons with genetic conditions from the workforce based on concerns about attendance and productivity.

Insurance

Insurers, such as issuers of health, disability, long-term care, and life insurance policies, discriminate among insureds and applicants for insurance based on health status due to bottom-line concerns and philosophical commitments. The bottom-line concern may be solvency or profitability. In addition many insurers have a commitment to "actuarial fairness," meaning that policies are priced to accurately reflect risk or expected losses. The forms of health status related discrimination operative in insurance are captured by the term "medical underwriting." Insurers obtain health information in order to decide as a threshold matter whether an applicant is an acceptable insurance risk, and also for use in setting premiums. Insurers can also design benefit packages strategically to limit potential losses. For example, health insurers can exclude certain conditions or procedures from coverage, or they can use "preexisting condition" clauses to eliminate coverage for any health care needs that relate to a condition diagnosed or treated prior to enrollment.

As a justification for these practices, insurers cite the problem of adverse selection. Adverse selection is the disproportionately heavy purchase of insurance by individuals at higher risk for claims than their insurers are aware. When individuals learn that genetic or other factors put them at high risk for disease, disability, and/or early death, they may load up on insurance. If insurers are ignorant of the information concerning risk, they cannot incorporate it into the process of underwriting. As the proportion of higher-risk individuals in an insurance pool increases, payouts will increase, and as payouts increase, premiums for all policyholders will go up. In a voluntary system of insurance, premium increases can be expected to drive out lower-risk individuals, resulting in a further increase in the proportion of higher-risk individuals, and so on. At least for health insurance, the problem of adverse selection could be addressed through universal coverage (6). The preferred solution for U.S. insurers has been to ensure that they have access to any health-related information that may be available to applicants for insurance.

Anecdotal and Survey Evidence of Genetic Discrimination

Although incentives to engage in genetic discrimination are present, the limited evidence available suggests that few employers or insurers systematically collect or use information derived from genetic testing. The expense involved in testing is surely a factor. A government-sponsored survey of Fortune 500 companies in 1989 found that of the 330 respondents, 12 reported conducting current biochemical genetic screening of employees. None reported conducting direct-DNA screening (5). Significantly more employers may solicit family histories containing genetic information.

In testimony before a congressional committee in 1998, a spokesperson for the Health Insurance Association of America stated that a survey of association members found none required genetic testing or solicited information regarding genetic testing as part of the application process (7). Authors of a recent study of genetic discrimination in health insurance concluded that a person with a serious genetic condition who is presymptomatic currently faces little or no difficulty in obtaining health insurance (8). Further over 75 percent of Americans obtain health insurance through their employers or the government and would not be subject to individual medical underwriting (9).

By way of contrast, almost 75 percent of life insurance policies are individual policies (10). As a result the majority of Americans with life insurance coverage are subject to medical underwriting. An individual may be denied life insurance if a family member suffers from a heritable disorder associated with premature death. For example, the practice among life insurers has been to decline to write individual policies for children of people with Huntington's disease until they are over 50 years old (11). Underwriting standards for individual policies of disability-income insurance may be even stricter. A number of researchers have surveyed currently healthy people with known genetic predispositions to disease or family members affected by genetic disorders. These researchers consistently find that a number of respondents report personal knowledge of instances of genetic discrimination in insurance and employment, although the magnitude of the problem is difficult to establish from these kinds of surveys (12).

Investigative reports and lawsuits are another source of information on genetic discrimination. A *New York Times* investigative report on genetic testing in the workplace

in 1980 found that DuPont de Nemours & Co. tested black job applicants for sickle cell trait. DuPont also tested job applicants for two enzyme deficiencies correlated with ethnicity. The purpose, according to the company's medical director, was to determine whether these tests would be of value in protecting the health of susceptible employees. (Another employer, Dow Chemical Company, engaged in genetic monitoring to detect any changes in chromosomal structure that might be attributable to workplace exposures.) Many years later, Lawrence Berkeley Laboratory, one of the national laboratories involved in the Human Genome Project, was sued for testing clerical and administrative employees for sickle cell trait, allegedly without employee consent. In each of these cases, it was unclear how the genetic information was used in decision making. DuPont apparently reassigned a small number of employees based on results of one of the enzyme tests (13).

Although the evidence concerning the actual extent of genetic discrimination is sparse, there is overwhelming evidence that people fear genetic discrimination. Nearly two-thirds of respondents in a 1997 survey reported that they would not undergo genetic testing if employers and health insurers would have access to the results. A 1995 survey found that over 85 percent of respondents were very or somewhat concerned about access to and use of genetic information by employers and insurers (12).

ADA: OVERVIEW

The ADA represents a blending of civil rights law and disability law. Key provisions were influenced by Title VII of the Civil Rights Act of 1964 and by the Rehabilitation Act of 1973 and its implementing regulations. From civil rights law the ADA derives its broad ambition of emancipating people with disabilities from a history of discrimination. From disability law it derives its demand for accommodation rather than simple equality, its mandate of community integration to combat isolation and segregation, and its emphasis on individualized or case-by-case decision making, taking account of the reasonableness of any accommodations requested and the burdens imposed on public and private actors. In addition the ADA reflects the belief that the denial of opportunity translates into dependency and nonproductivity, giving rise to substantial social costs.

The ADA was signed into law on July 26, 1990; most provisions became effective on or before July 26, 1992. Its structure is fairly neat. The law begins with a recital of findings and purposes and a number of general definitions, including a definition of the key term "disability." Each of the first three titles addresses a particular arena in which discrimination may occur. Title I governs employment. It applies to employers with 15 or more employees, excluding the United States and certain private clubs. Title II governs public services and applies to state and local governments and specified transportation agencies. Title III governs public accommodations and services and applies to all private entities with operations affecting interstate commerce. Each of these titles contains a prohibition of discrimination — with some variation in phrasing — and an enforcement provision. A section labeled "miscellaneous provisions" offers guidance on interpretation of the statute generally, and Title IV concerns telecommunications.

The U.S. government is subject to the antidiscrimination mandates of Section 501 of the Rehabilitation Act (14). The Rehabilitation Act also applies to federal government contractors (Section 503) and any entity receiving federal financial assistance (Section 504). Differences between the ADA and the Rehabilitation Act include the language in the prohibitions on discrimination. The Rehabilitation Act prohibits discrimination "solely" by reason of a person's disability (15). Titles I through III of ADA use "because of" or "on the basis of" disability, suggesting that the disability need not be the only factor motivating the adverse treatment (16). Enforcement is another area of difference. The Equal Employment Opportunity Commission (EEOC) is charged with issuing regulations under Title I of ADA and enforcing its provisions, and the Department of Justice is responsible for Titles II and III of ADA. The Rehabilitation Act does not concentrate authority in this manner. For example, the Department of Health and Human Services (DHHS) issues implementing regulations governing its grantees and enforces the law as applied to its own activities and those of its grantees through its Office for Civil Rights. The remedies available under the two statutes also differ. For example, the courts have established that money damages may be awarded under the Rehabilitation Act, whereas this remedy is sometimes unavailable under ADA.

The sections of ADA that are most relevant to the area of genetic discrimination are the definition of disability, the provisions in Titles I and III affecting employment and insurance, and the miscellaneous provisions that address the relationship of ADA to other laws and to insurance. While there are similarities, Titles I, II, and III are not perfectly symmetrical. In particular, a provision of Title I that regulates the collection of medical information by employers has no equivalent elsewhere in ADA. This important prophylactic provision will be discussed first, before consideration of whether and how the ADA might affect genetic discrimination in employment and other areas.

THRESHOLD QUESTION: IS A GENETIC CONDITION A "DISABILITY"?

Statutory Language

The definition of disability is important because the antidiscrimination provisions of ADA protect individuals with disabilities from discrimination based on disability. Under ADA, disability can be established in one of three ways. First, an individual can show that he or she has "a physical or mental impairment that substantially limits one or more of the major life activities of such individual." Second, an individual can show that he or she has "a record of such an impairment." Third and finally, an individual can show that he or she is "being regarded as having such an impairment" (17). Although other provisions of the statute categorically exclude certain conditions from the definition of disability (e.g., transvestitism, pedophilia,

compulsive gambling), the determination of whether a particular condition satisfies one of the three prongs of the definition of disability is generally made on a case-by-case basis.

Under the first prong of the definition, an affected individual might argue that a genetic condition is a physical or mental impairment that presently substantially limits the major life activity of reproduction, or will in the future substantially limit one or more major life activities of the individual. Under prong three, an affected individual might argue that he or she is being regarded as having a physical or mental impairment that presently substantially limits one or more major life activities, or that the anticipated future impairment, whether certain or merely more likely than for the average person, is being imputed to the present as evidenced by the discriminatory conduct of a third party. The strength of the argument will likely depend on the nature of the condition. The adjective "genetic" alone is fairly uninformative for ADA purposes; generally, ADA is concerned with function rather than causation. Commentators have identified at least seven categories of genetic conditions that may merit separate analysis: (*1*) already-expressed severe genetic conditions, such as symptomatic Huntington's disease, (*2*) already-expressed minor genetic conditions, such as polydactyly expressed in an extra finger or toe, (*3*) unexpressed late-onset genetic conditions, such as presymptomatic Huntington's disease detected through genetic testing, (*4*) genetic mutations associated with increased risk of disease (predispositions), such as BRCA1 or BRCA2 mutations detected through genetic testing, (*5*) unaffected carriers of recessive and X-linked disorders, such as carrier-status for cystic fibrosis, (*6*) genetic conditions that are cured or kept under control through treatment, such as phenylketonuria controlled through diet, and (*7*) conditions with a genetic basis that do not limit major life activities but are stigmatized or misunderstood, such as Down syndrome or Tourette syndrome (18). Category 1 conditions should satisfy the first prong of the definition of disability. Category 2 conditions would appear to fall outside the definition of disability (unless they are stigmatized or misunderstood). For the other categories, the outcome is uncertain, although the legislative history and the opinions of administrative agencies and the courts can be mined for insight.

Legislative History and Agency Interpretations

The legislative history of the ADA concerning genetic conditions is scanty at best. On the day the House of Representatives voted on the final conference report, three congressmen entered in the record statements that the law would protect "carriers of a disease-associated gene" from employment discrimination based on speculation about future illness or increased health care costs for carriers or their dependents (18). There is no record of debate on this point. To the extent the issue was thought of at all, then, the cases that came readily to mind for the few genetically minded legislators were autosomal recessive conditions such as sickle cell anemia. This is not surprising, since presymptomatic predispositional testing for diseases such as breast cancer and colon cancer has only recently become widely available.

Reliance on agency interpretations is complicated by the fact that the introductory sections of ADA, unlike Titles I through IV, contain no delegation of authority to a particular agency. EEOC, which has charge of Title I, has produced several documents that elaborate on the definition of disability. EEOC's Title I regulations attempt to clarify certain aspects of the definition of disability, although they do not specifically mention genetic conditions. According to the Title I regulations, a "physical or mental impairment" includes "any physiological disorder, or condition, cosmetic disfigurement, or anatomical loss" affecting at least one of the major body systems. Some examples of major life activities are given, such as caring for oneself, walking, and working. To "substantially limit" is to significantly restrict as to condition, manner, or duration of performance relative to the performance of the average person. Factors to be considered include the nature and severity of the impairment, its duration or expected duration, and "the permanent or long term impact, or the expected permanent or long term impact of or resulting from the impairment" (19). The Title I regulations also present several possible variants of "regarded as" disability; the unifying element is the focus on what the third party's treatment of the individual suggests, rather than on the physical or mental state of the individual. The Title III regulations issued by the Department of Justice are similar to the Title I regulations on these points (20).

Interpretive guidelines receive considerably less deference from the courts than regulations but may still carry some weight. EEOC's interpretive guidance for Title I, published as an appendix to the regulations, affirms the importance of case-by-case determinations. However, it contains some rather unnuanced assertions. It states that the definition of impairment does not include "characteristic predisposition to illness or disease" (21). This suggests that regardless of contextual factors, genetic predispositions to cancer or heart disease are not disabilities protected under ADA — unless the "regarded as" prong of the definition fits the case. On the other hand, EEOC makes HIV infection an example of an impairment that is inherently substantially limiting. Hence, to the extent that a genetic mutation can be analogized to HIV infection, this language suggests a strong case can be made for recognizing the mutation not only as an impairment but also as a disability under the first prong of the ADA definition.

EEOC has also issued a compliance manual for Title I. On March 15, 1995, an amended manual was released that included the following language concerning "regarded as" disability: "This part of the definition of 'disability' applies to individuals who are subjected to discrimination on the basis of genetic information relating to illness, disease, or other disorders. Covered entities that discriminate against individuals on the basis of such genetic information are regarding the individuals as having impairments that substantially limit a major life activity" (22). In the compliance manual, EEOC gives the example of an asymptomatic individual with a genetic

mutation conferring an increased risk of colon cancer; an employer discovers this information after making a conditional offer of employment and withdraws the offer due to concerns about attendance, productivity, and insurance costs. The example suggests that where action is taken on the basis of present fears, the definition of disability is satisfied, even if the fears are about future performance or future costs and even if the fears are not, strictly speaking, unfounded. Although the emphasis in EEOC's interpretation of "regarded as" disability is on myths, stereotypes, and misperceptions, EEOC states that the individual does not have to demonstrate that the employer's perception is wrong, for example, that health care costs will not increase if persons are hired who are at elevated risk of serious illness. (On the other hand, a finding of disability under ADA will not necessarily lead to a finding of liability. As discussed below, a prima facie case must include a showing that discrimination occurred because of or on the basis of the disability. Establishing a violation may be difficult, given that employers are unlikely to document that decisions are being made on the basis of genetic information.) The Department of Justice has not addressed genetic conditions in its interpretive guidance for Title III.

Judicial Opinions

No published judicial opinion addresses whether an individual with a genetic mutation associated with disease, but not yet expressed in symptoms, has a disability under ADA. A number of cases have raised somewhat similar issues. In *Bragdon v. Abbott*, the U.S. Supreme Court held that HIV infection, even in its early stages, is a disability under the first prong of the ADA definition (23). It is important to understand the court's reasoning, since an unexpressed genetic condition may or may not share the characteristics the court found significant in *Bragdon*. First, the court addressed whether HIV infection is an impairment. The court found that given the immediacy of the damage to the hemic and lymphatic systems, the predictable course of the disease, and its severity, an impairment exists from the moment of infection. Next the court concluded that as to the plaintiff reproduction was a major life activity. The court further concluded that the risk of infecting a partner and the risk of infecting a child could substantially limit this activity. Even assuming the risk of perinatal transmission could be lowered from 25 to 8 percent using antiretroviral therapies, the opinion stated that it is not possible to say as a matter of law "that an 8 percent risk of transmitting a dread and fatal disease to one's child does not represent a substantial limitation on reproduction."

Huntington's disease and other similar late-onset genetic conditions would appear to meet the *Bragdon* criteria of predictability and severity. Assuming some damage to a body system could be established prior to full expression, they would likely qualify as impairments from the moment of transmission. The case for genetic mutations associated with increased risk of disease is considerably weaker under the *Bragdon* criteria, since the element of predictability is missing; the same would be true for the category of unaffected carriers. (There is some irony here. Many people are struck by the unfairness where a third-party treats a possibility of disease as if it were a certainty as a risk avoidance measure, but the case for protection under ADA appears stronger where the question is not whether but when a disease will develop.) Genetic conditions would pose risks of transmission to a child analogous to the risk of infection associated with HIV, although the risk to a partner would be absent. For monogenic genetic disorders, the risk of transmission will generally be 25 or 50 percent. Factors affecting the reproductive options of particular plaintiffs, such as the availability of preimplantation genetic diagnosis, would also appear relevant under the Supreme Court's framework for analysis.

The concurring and dissenting opinions in *Bragdon* are of interest as well. Justice Ginsburg, concurring, stated that "[n]o rational legislator ... would require nondiscrimination once symptoms become visible but permit discrimination when the disease, though present, is not yet visible." This reasoning lends further support to a distinction, in the assessment of impairment, between those genetic mutations that inevitably give rise to disease and those that simply increase the risk of disease. Writing for the dissenters, Chief Justice Rehnquist stated that taken to its extreme the logic of *Bragdon* would "render every individual with a genetic marker for some debilitating disease 'disabled' here and now because of some possible future effects." Justice Rehnquist meant this as a warning of a peril to be avoided, but the language could be used by those who favor just such an extension.

In *Sutton v. United Air Lines*, the Supreme Court returned to the definition of disability, focusing on the substantial limitation requirement (24). The case involved two women who suffered from severe myopia but had 20/20 vision with corrective lenses. The court held that what matters for purposes of determining disability is an individual's present state, which includes mitigating measures such as corrective lenses; being "potentially or hypothetically" substantially limited does not suffice. (The court was influenced by a congressional finding, in the introductory provisions of ADA, that 43 million Americans had one or more physical or mental disabilities. The court found this number hard to reconcile with a broad interpretation of disability.) While the holding in *Sutton* would not change the result in *Bragdon*, or an analogous case in which the potential for transmission of a genetic mutation would substantially limit current reproductive options, it does suggest that any argument that a mutation qualifies as a disability under the first prong of the definition due to its anticipated effects will fail. Certainly those with genetic conditions that are cured or kept under control through treatment will have a hard time establishing disability, unless they can show that the side effects of treatment are themselves disabling, or satisfy one of the other prongs of the definition.

Unfortunately for plaintiffs, *Sutton* also puts up a barrier to establishing "regarded as" disability, by suggesting that concerns about an impairment that are sufficient to prompt negative employment action may not be sufficient to establish that the employer is regarding the individual as disabled. Cases decided before *Sutton* had

interpreted the definition of disability to encompass a third party's perception that disability was likely in the future, if that perception influenced present action. In *Doukas v. Metropolitan Life Insurance Company*, a federal district court noted that limiting ADA to perception of present disability would violate congressional intent and "allow an employer to refuse to hire an epileptic as long as the job applicant was not having a seizure at the time" (25). The court in *Winslow v. IDS Life Insurance Co.* found this reasoning persuasive (26). These cases, which concerned mental illness, and analogous cases involving genetic predisposition to disease, might be distinguished from *Sutton* by the unavailability of mitigating measures and the severity of the potential impairment. *Sutton* was, after all, a case about mitigating measures, despite language dismissive of probabilistic calculations and fears about the future as elements in the construction of disability. In *Cook v. State of Rhode Island Dept. of Mental Health*, a case decided under the Rehabilitation Act, the First Circuit Court of Appeals found that an employer's fears about risks associated with the plaintiff's morbid obesity were sufficient to establish perceived disability (27).

Is it good policy to adopt a generous interpretation of the ADA definition of disability and so extend ADA protections to unaffected carriers of recessive and X-linked disorders and individuals with unexpressed late-onset genetic conditions or genetic predispositions to disease? The answer to this question would appear to rest on a careful analysis of the fit between the purposes of ADA and the experience of persons falling within the particular category under consideration. Unaffected carriers of sickle cell trait can point to a history of isolation and segregation. Individuals with a genetic predisposition to cancer can point out that when employers turn them away, their contributions as productive members of society are diminished, perhaps unnecessarily and certainly prematurely. It is worth noting that state antidiscrimination laws may contain definitions of disability which are broader than the ADA's. For example, New York's highest state court has interpreted the New York State Human Rights Law to include "diagnosable medical anomalies which impair bodily integrity and thus may lead to more serious conditions in the future" (28).

COLLECTION OF INFORMATION BY EMPLOYERS

Title I of the ADA regulates medical examinations and inquiries conducted by employers (29). What is permissible varies according to the stage in the hiring process. To make sense of ADA, one needs to view hiring in terms of three stages: a pre-employment or interviewing stage prior to an offer of employment, a pre-placement or entrance examination stage after an offer of employment has been made but before commencement of employment duties, and a postemployment stage initiated with employment duties. The only acceptable *pre-employment* inquiries concern the ability of an applicant to perform job-related functions. Examinations and inquiries are generally prohibited *postemployment*, unless they can be shown to be "job-related and consistent with business necessity." However, two kinds of information-gathering activities relating to existing employees are expressly permitted: (*1*) voluntary medical examinations, including voluntary medical histories, as part of an employee health program, and (*2*) inquiries into the ability of an employee to perform job-related functions.

Employers have the most freedom at the *pre-placement* stage of employment. Title I states that after an offer of employment has been made, but prior to the commencement of employment duties, an employer may require a medical examination (which may include a review of medical records) and may condition the offer of employment on the results. The two limitations on employer discretion in this area are (*1*) all entering employees must be subjected to examination regardless of disability, and (*2*) the medical information obtained in this way must be collected and maintained on separate forms and in separate files, must be treated as confidential, and can be used only as permitted under Title I. Title I permits release of information to supervisors and managers where it concerns necessary work restrictions and accommodations, to first-aid and safety personnel when appropriate if emergency treatment may be required, and to government officials conducting compliance investigations. The same rules concerning separate forms and files, confidentiality, and use, apply to information obtained through examinations and inquiries made of existing employees.

The regulations issued by EEOC explicitly state that employment entrance examinations need not be job-related and consistent with business necessity (30). However, if an employer withdraws an offer of employment based on the results of an examination, the criteria used must not be of a kind to screen out or tend to screen out individuals with disabilities, or must be job-related and consistent with business necessity. This restriction on employer discretion may be hard to enforce, since job applicants who have received conditional offers of employment will often have a difficult time detecting illegal uses of information. An employer is generally not required to share the employer's reasons for withdrawing an offer of employment with the affected individual, and ADA does not alter this state of affairs (18). Some state genetic privacy laws require specific consent for genetic testing and disclosure of results, but absent such legislation, an individual may be completely in the dark concerning the nature or results of any tests conducted or information reviewed as part of an entrance examination.

Once in court, job applicants or employees face several hurdles. They may be met with the argument that only individuals who meet the statutory definition of disability are protected from inquiries and examinations. (As discussed at considerable length above, this test may be difficult to satisfy.) Federal appeals courts in the Eighth, Ninth, and Tenth Circuits (covering the western states and much of the midwest) have ruled that a plaintiff need not be disabled in order to state a claim for the unauthorized gathering or disclosure of confidential information by an employer (31). Plaintiffs alleging a violation of ADA's confidentiality protections may have difficulty showing that the violation resulted in some kind of tangible injury. Further ADA's confidentiality

protections only apply to information collected through pre-placement medical examinations and the kinds of medical examinations and inquiries authorized under the ADA for existing employees. This leaves out medical information in benefit records, for example, medical information contained in benefit request forms. In *Yoder v. Ingersoll-Rand Company*, a federal district court ruled that the confidentiality provisions of ADA did not apply to a physician's statement confirming a diagnosis of HIV/AIDS in a disability benefit request form (32). The court rejected an argument for broader protection based on the general purposes of ADA. Medical information obtained before the effective date of ADA will also fall outside its protections (33). Finally, ADA does not prohibit employers from using general release forms, or soliciting consent to a broad battery of tests, at least at the conditional offer (pre-placement) and postemployment stages. Employers may compile extensive information in connection with voluntary wellness and employee assistance programs (34).

The Lawrence Berkeley Laboratory case, mentioned in the introduction, illustrates some of the difficulties associated with pursuing a remedy for unauthorized genetic testing. The employees in that case sued their employer under Title I of ADA and under other state and federal laws. The employees contended that testing for sickle cell trait and other sensitive medical conditions, allegedly without their knowledge or authorization, violated ADA because the testing was neither job related nor consistent with business necessity. They also advanced claims based on violations of privacy rights under federal and state constitutions. Finally, they argued that in singling out black employees for sickle cell testing (and female employees for pregnancy testing), the defendants violated Title VII of the Civil Rights Act. Title VII prohibits discrimination in employment based on race, color, religion, sex, or national origin. The employees did not allege that any employment-related action was taken on the basis of their test results or that their tests results were disclosed to third parties.

In *Norman-Bloodsaw v. Lawrence Berkeley Laboratory*, the Ninth Circuit Court of Appeals concluded that there is no remedy under ADA for unauthorized testing or testing lacking a job- or business-related justification at the pre-placement stage of employment (35). The court suggested that the only viable claim given the facts would concern a failure to properly maintain medical records according to ADA requirements but that something more than a general allegation of inadequate safeguards would be necessary for that purpose. The court found that the plaintiffs were entitled to a trial on their other claims.

It is important to note that ADA does not preempt state laws that provide greater or equal protections, and plaintiffs may have greater success in pursuing claims for breaches of confidentiality or unauthorized testing or inappropriate inquiries under state law. Title I of ADA must also be considered together with other federal laws affecting employment. For example, the Occupational Safety and Health Act (OSHA) requires medical monitoring of employees who may be exposed to hazardous chemicals. The implementing regulations require physical examinations before workers are assigned to certain sites, including a family history addressing genetic factors, but they do not require genetic testing (36). EEOC has stated that Title I does not halt the performance of such examinations, which might in any event be justified as job related and consistent with business necessity (37).

DISCRIMINATION IN EMPLOYMENT AND EMPLOYER-PROVIDED BENEFITS

Elements of Prima Facie Case and General Defenses

The general prohibition of discrimination in Title I is broadly stated to encompass all aspects of employment (38). So long as an individual with a disability is "qualified," that is, can perform all essential job functions, he or she is protected from discrimination because of or on the basis of the disability with respect to job application procedures, hiring, advancement, discharge, compensation, training, and other terms, conditions, and privileges of employment. Limiting, segregating, or classifying a job applicant or employee in a way that adversely affects his or her employment opportunities or status, and failing to make reasonable accommodations (i.e., accommodations that could be accomplished without undue hardship to the employer), are instances of discrimination, as are failures to abide by the rules concerning medical examinations and inquiries. Individuals with dependents with disabilities are also protected, because discrimination is defined to include the denial of equal jobs or benefits to a qualified individual because of the known disability of an individual with whom the qualified individual is known to have a relationship or association.

Even if a plaintiff makes a prima facie case, that is, establishes the elements of disability and disability-based discrimination, the employer can avoid liability by showing that a particular use of qualification standards, tests, or selection criteria was job-related and consistent with business necessity, and that no reasonable accommodation was possible under the circumstances. Qualification standards may include a requirement that an individual not pose a "direct threat" to the health or safety of other individuals. EEOC's Title I regulations specify that a direct threat is "a significant risk of substantial harm" (to self or others) that cannot be eliminated or reduced by reasonable accommodation (39). The assessment of risk must be based on "the most current medical knowledge and/or on the best available objective evidence." The likelihood and imminence of the potential harm are among the factors to be considered. This fairly stringent standard should preclude employers from making employment or job assignment decisions based on genetic susceptibilities or predispositions to disease, unless the science is good and tests of significance and substantiality are met. ADA does not prevent employers from seeking to understand and reduce hazards in the work environment. Nor does it prohibit the offer of accommodation to an employee with a condition that greatly increases the likelihood that the employee will suffer harm from a particular activity, or become incapacitated in a way that puts others at risk, assuming information about the condition is acquired by legal means. The fall-back defense of "undue hardship"

requires a showing of "significant difficulty or expense," considering the nature and net cost of the accommodation, the financial resources of the facility and the covered entity, the impact of the accommodation on operations, and so on.

Employer-Provided Insurance and the Insurance Safe Harbor

While an employer cannot refuse to hire, or fire, a qualified individual with a disability due to fears about increased health care costs, or exclude the individual from benefit programs available to other employees, the employer is given considerable latitude in the area of insurance. The key provision in this area is what has become known as ADA's "insurance safe harbor" (40). The relevant subsection states that Titles I through IV should not be construed to prohibit or restrict (*1*) an insurer or other entity that administers benefit plans from underwriting risks, classifying risks, or administering such risks in a manner based on or not inconsistent with state law; (*2*) a person or organization from establishing, sponsoring, observing or administering the terms of a bona fide benefit plan based on underwriting risks, classifying risks, or administering such risks in a manner based on or not inconsistent with state law; or (*3*) a person or organization from establishing, sponsoring, observing or administering the terms of a bona fide benefit plan that is not subject to state laws that regulate insurance. (The Employee Retirement Income Security Act of 1974, known as ERISA, prevents the application of state insurance laws to employers' self-funded health plans.) However, ADA states that this provision cannot be used as a "subterfuge" to evade the purposes of Titles I and III.

EEOC's interpretive guidance on Title I addresses medical underwriting and preexisting condition clauses and benefit design issues. The guidance document states that medical underwriting and preexisting condition clauses included in health insurance policies offered by employers are not affected by ADA, except to the extent that practices are found to be inconsistent with applicable state law (21). In the area of health insurance, the Health Insurance Portability and Accountability Act of 1996 (HIPAA) may offer more extensive protections. (HIPAA focuses on group policies, but those with individual polices may benefit to a limited extent from provisions governing the transition from group to individual coverage and guaranteed renewability.) HIPAA *permits* issuers of group policies to impose limited preexisting condition exclusions *but only* if these relate to conditions for which medical advice, diagnosis, care, or treatment was recommended or received within the six-month period ending on the enrollment date. Genetic information cannot be treated as a condition in the absence of a diagnosis of the condition related to such information (41). HIPAA *prohibits* issuers of group policies from excluding an individual within the group from coverage on the basis of a health status-related factor relating to the individual or a dependent. The list of health status-related factors includes genetic information. HIPAA also prohibits variation in benefits, premiums, and contributions for similarly situated group members on the basis of these factors, but neither HIPAA nor the ADA requires that employers offer any insurance at all.

Employers can affect health care costs through benefit design as well as through medical underwriting and preexisting condition clauses. EEOC has concluded that ADA does not prohibit employers from placing limits on coverage for certain procedures or treatments (e.g., visit limits), even if these adversely affect individuals with disabilities, so long as the limits are applied equally to individuals with and without disabilities (21). EEOC offers more extensive comment on health insurance in its Interim Enforcement Guidance on Disability-Based Distinctions in Employer Provided Health Insurance (42), and in the Title I Technical Assistance Manual (37). The guidelines on benefit design would permit an employer to exclude all experimental drugs or procedures from coverage, so long as this restriction is applied evenhandedly to all insured individuals. It follows that employers would have no obligation to arrange for coverage of gene therapy and other interventions still in the research phase. Indeed, EEOC states that broad distinctions that apply to a range of dissimilar conditions and constrain individuals with and without disabilities are not distinctions based on disability. A term or provision is disability-based only if it singles out a particular disability or discrete group of disabilities or disability in general for inferior treatment. Cancers, muscular dystrophies, and kidney diseases are given as examples of discrete groups of disabilities. Genetic disorders, or inherited genetic disorders, would arguably constitute a discrete group of disabilities, meaning that it would not be permissible for an employer or plan administrator to single out interventions targeting genetic conditions for more limited coverage than other conditions. Picking and choosing among genetic conditions might also run afoul of the ADA. In *Henderson v. Bodine Aluminum, Inc.*, a woman with breast cancer argued that an insurer's policy of paying for bone marrow transplants for some cancers, but not breast cancer, violated the ADA. The Eighth Circuit Court of Appeals ruled in her favor, stating that "if the evidence shows that a given treatment is non-experimental ... and the plan provides the treatment for other conditions directly comparable to the one at issue, the denial of that treatment arguably violates the ADA" (43).

Courts are still struggling with the "subterfuge" language in ADA. EEOC's Interim Enforcement Guidance states that this language refers to "disability-based disparate treatment that is not justified by the risks or costs associated with the disability" (42). If an employee can make a prima facie case of discrimination, then EEOC puts the burden on the employer to produce evidence that "the disparate treatment is justified by legitimate actuarial data, or by actual or reasonably anticipated experience, and that conditions with comparable actuarial data and/or experience are treated in the same fashion," or to offer some other acceptable justification for the practice (e.g., that there is no other way to ensure the solvency of the plan or prevent a drastic increase in premiums). This suggests that ADA can be used to challenge underwriting decisions that are based on outdated or inaccurate information about genetics in general or specific genetic disorders, or on myths, fears, or stereotypes that have no basis in science. However,

the phrase "reasonably anticipated experience," common in state insurance laws, does appear to create some room for inference from available data. Also a few courts have adopted a restrictive interpretation of subterfuge, holding that a benefit plan cannot be a subterfuge unless the employer intended by virtue of the plan to discriminate in a non-fringe-benefit-related aspect of the employment relation (e.g., the employer set out to design a benefit plan that would discourage persons with disabilities from applying for jobs) (44).

DISCRIMINATION BY INSURERS

Threshold Question: Scope of Title III

As noted above, Title III of the ADA regulates public accommodations. Rather than a true definition of the term, the statute offers a laundry list of private entities that are covered if their operations affect interstate or foreign commerce. These include an "insurance office, professional office of a health care provider, hospital, or other service establishment" (45). Because insurance is mentioned, individuals who have experienced discrimination in insurance have turned to Title III for a remedy. Resort to Title III is most common where insurance is not provided through an employer (or relief under Title I is unavailable for some other reason). Under Title III, as under Title I, the individual seeking a remedy for discrimination must first establish that he or she is an individual with a disability within the meaning of ADA.

A major point of controversy at present is whether Title III extends beyond access to physical structures to address access to services such as insurance policies. The courts are divided. In its Title III Technical Assistance Manual, the Department of Justice assumes rather than argues for the broad view (46). The evidence offered in favor of the restrictive view includes the many references to "offices" in the list that defines public accommodation and the insurance safe harbor. The leading case for the restrictive view is *Parker v. Metropolitan Life Ins. Co.* (47). In *Parker*, an individual sued her employer and her insurer claiming that a shorter benefit period for mental disability than for physical disability under an employer-provided disability policy violated ADA. The Sixth Circuit Court of Appeals ruled that a public accommodation is a physical place and a disability policy not obtained in an office transaction is *not* a service or good offered by a place of public accommodation. Although it was unnecessary to the decision in the case, the court also concluded that Title III does not extend to the contents (terms and conditions) of insurance policies. *Parker* has been followed by the Third and Ninth Circuit Courts of Appeals, and the Seventh Circuit Court of Appeals has also taken the position that Title III does not apply to the contents of insurance policies (48).

The First Circuit Court of Appeals has presented the case for the broad view. In *Carparts Distribution Center v. Automotive Wholesaler's Ass'n*, the First Circuit reasoned that by including "travel service" in the list of examples of public accommodations, Congress signaled that commercial enterprises not requiring physical entry could be public accommodations (49). The court noted that neither Title III nor the implementing regulations makes any mention of physical boundaries or physical access. Further the court believed it would be irrational, and inconsistent with the purposes of ADA, to conclude that persons who enter an office are protected by ADA, but persons who purchase services over the telephone or by mail are not. As to whether ADA requires scrutiny of the contents of insurance policies, the court found that in some cases, meaningful access to a service requires a change in substance. The reasoning of *Carparts* has been adopted by the Second Circuit Court of Appeals (50).

Other Title III Issues

Even if Title III is found to apply to insurance policies, plaintiffs may have a difficult time prevailing on a claim. The antidiscrimination language in Title III is broad: "No individual shall be discriminated against on the basis of disability in the full and equal enjoyment of the goods, services, facilities, privileges, advantages, or accommodations of any place of public accommodation by any person who owns, leases (or leases to), or operates a place of public accommodation" (51). Administrative methods that have the effect of discriminating on the basis of disability are expressly included. And Title III, like Title I, addresses discrimination based on association with an individual with a disability. However, if an insurer can establish that the actions taken were in accordance with sound actuarial principles, reasonably anticipated experience, or bona fide risk classification, they will likely be sheltered by the insurance safe harbor (discussed above). Construing the Title III antidiscrimination provisions in light of the insurance safe harbor, the Department of Justice has concluded that "a public accommodation may offer a plan that limits certain kinds of coverage based on classification of risk, but may not refuse to insure, or refuse to continue to insure, or limit the amount, extent, or kind of coverage available to an individual, or charge a different rate for the same coverage solely because of a physical or mental impairment, except where the refusal, limitation, or rate differential is based on sound actuarial principles or is related to actual or reasonably anticipated experience" (46). In essence, the Department of Justice and the EEOC have chosen the same middle course.

As in the employment context, it may be difficult for plaintiffs to make a prima facie case because they lack access to key information. The Title III Technical Assistance Manual states that ADA does not require an insurer to provide a copy of the actuarial data on which its actions were based at the request of the applicant (46). Further several courts, including the Third Circuit Court of Appeals, have suggested that allegations of subterfuge do not compel insurers to come forward with evidence to justify their coverage or underwriting decisions (48).

Still a number of plaintiffs have prevailed in lawsuits against insurers brought under Title III. For example, in *Chabner v. United of Omaha Life Insurance Co.*, an individual with fascioscapulohumeral muscular dystropy sued a life insurer for issuing him a life insurance policy at a premium that was considerably higher than the standard

premium (52). As a threshold matter, the court found that Title III applies to insurance underwriting practices. Next the court held that where it is undisputed that an individual was treated differently based solely on his disability, the insurer has the burden of coming forward with evidence that the differential treatment was based on sound actuarial principles or actual and reasonably anticipated experience. The court added that even though the legal standard refers to "anticipated experience," insurers may not engage in speculation; that is to say, underwriting must always have a basis in actuarial data. The court found that in this case the defendant had failed to satisfy its evidentiary burden. As a result the court entered summary judgment in favor of the plaintiff. A recent decision from the Ninth Circuit Court of Appeals may affect the continued validity of the legal analysis in *Chabner* (48), but the facts are representative of the type of case that may become increasingly common as genetic disorders are subject to medical underwriting.

DISCRIMINATION IN PUBLIC PROGRAMS

Although the focus of discussion has been on employment and private insurance, the potential exists for discrimination in many public sector programs. In the ADA framework, public programs fall under Title II, governing agencies of state and local government. Although the bulk of Title II is devoted to public transportation, this title contains a general prohibition of discrimination by reason of disability affecting participation in or receipt of benefits of services, programs, or activities (53). Professional licensing appears to be one area where new developments in genetic technology could give rise to discrimination. Some have noted that there is also considerable potential for genetic discrimination in the public schools. (Private schools would be public accommodations covered under Title III.) It is conceivable that genetic information will someday be used to make modifications to programs to better meet the needs of children with genetic conditions. Genetic information may also be used to segregate children, where an administrator or teacher is persuaded that a certain mutation is associated with behavioral and disciplinary problems, or to attach labels to children that may become self-fulfilling prophecies (54). In these circumstances Title II of ADA could be used to challenge segregation and to combat other damaging practices based on stereotypes or hypothetical risks rather than individualized assessment. The Rehabilitation Act, which preceded ADA, remains available as an additional source of protection against discrimination by entities receiving federal financial assistance.

CONCLUSION

In sum, ADA exhibits the usual limitations of legislation: failure to adequately address situations remote from the experience of its framers, and resort to ambiguous language to achieve consensus. Few of the lawmakers who debated ADA reflected on the significance of developments in the field of genetics for civil rights and disability law, and the text of ADA contains no mention of genetic conditions or genetic testing. Accordingly it is uncertain whether genetic conditions that are known, but presently asymptomatic, are covered under the statute. The extent to which the new antidiscrimination law should change the rules for insurance companies was certainly debated, but the resolution of that debate allowed for a range of interpretations. Indeed, the language of the statute is sufficiently ambiguous to send courts in different directions on the question of whether ADA imposes any constraints on the substance of insurance policies, especially those purchased by individuals. At present, then, there is considerable uncertainty concerning the relevance of ADA to genetic discrimination. ADA is certainly not a comprehensive response to the problem of genetic discrimination. Nonetheless, unless and until comprehensive legislation is enacted at the federal level, ADA will have to serve as proxy for a more complete and considered response, supplemented by the Rehabilitation Act and HIPAA and other federal and state laws. As genetic knowledge increases, and with it the potential for genetic discrimination, the courts will inevitably have to address some of the areas of uncertainty described above. In the not-to-distant future we should have, if not more justice, then at least more clarity.

BIBLIOGRAPHY

1. Pub. L. No. 101-336 (codified at 42 U.S.C. §§12101-12213).
2. 42 U.S.C. §12101(b).
3. M.S. Marquis and S.H. Long, *Health Affairs* **18**, 161–166 (1999).
4. H.T. Greely, in D.J. Kevles and L. Hood, eds., *The Code of Codes*, Harvard University Press, Cambridge, MA, 1993, pp. 264–280.
5. Office of Technology Assessment, U.S. Congress, *Genetic Monitoring and Screening in the Workplace*, vol. 1, U.S. Government Printing Office, Washington, DC, 1990.
6. M.A. Rothstein, *Am. J. Law, Med. Ethics* **26**, 198–204 (1998).
7. L.C. Volpe, *Genet. Test.* **2**(1), 9–12 (1998).
8. M.A. Hall and S.S. Rich, *Am. J. Hum. Genet.* **66**, 293–307 (2000).
9. O. Carrasquillo et al., *N. Engl. J. Med.* **340**, 109–114 (1999).
10. H. Ostrer et al., *Am. J. Hum. Genet.* **52**, 565–568 (1993).
11. L. Goch, *Best's Review-Life-Health Insurance Ed.*, November 1, 1998, WL 11240436, Westlaw, ALLNEWS database.
12. P.S. Miller, *J. Law, Med. Ethics* **26**(3), 189–197 (1998).
13. J. Seltzer, *Hofstra Law Rev.* **27**, 411–471 (1998).
14. Pub. L. No. 92-112, as amended by Pub. L. No. 93-516 (codified at 29 U.S.C. §§701-796).
15. 29 U.S.C. §794.
16. 42 U.S.C. §§12112, 12132, 12182.
17. 42 U.S.C. §12102.
18. M.A. Rothstein, *Houston Law Rev.* **29**(1), 23–84 (1992).
19. 29 C.F.R. §1630.2(h).
20. 28 C.F.R. §36.104.
21. 29 C.F.R. Pt. 1630, App., Interpretive Guidance on Title I of the Americans with Disabilities Act, §1630.2(h).
22. Equal Employment Opportunity Commission, *Compliance Manual*, §902.8 (1995).

23. 524 U.S. 624 (1998).
24. 527 U.S. 471 (1999).
25. 7 A.D. Cases 848, 1997 WL 833134 (D.N.H. 1997).
26. 29 F. Supp. 2d 557 (D. Minn. 1998).
27. 10 F.3d 17 (1st Cir. 1993).
28. *State Div. of Human Rights v. Xerox Corp.*, 65 N.Y.2d 213, 491 N.Y.S.2d 106, 109, 480 N.E.2d 695 (N.Y. 1985); *Application of State Div. of Human Rights on Complaint of Granelle*, 504 N.Y.S.2d 92, 118 A.D.2d 3 (N.Y.A.D. 1 1986).
29. 42 U.S.C. §12112(d).
30. 29 C.F.R. §1630.14.
31. *Cossette v. Minnesota Power & Light*, 188 F.3d 964 (8th Cir. 1999); *Armstrong v. Turner Industries*, 141 F.3d 554 (5th Cir. 1998).
32. 31 F.Supp.2d 565 (N.D. Ohio 1997).
33. See *Buchanan v. City of San Antonio*, 85 F.3d 196 (5th Cir. 1996).
34. E.E. Schultz, *Wall Street J.* May 18 (1999).
35. 135 F.3d 1260 (9th Cir. 1998).
36. OSHA Instruction STD 1-23.4 (1980), 1 OSHR Ref. File 21:8212.
37. Equal Employment Opportunity Commission, *A Technical Assistance Manual on the Employment Provisions (Title I) of the Americans with Disabilities Act*, U.S. Government Printing Office, Washington, DC, 1992.
38. 42 U.S.C. §12112.
39. 29 CFR §1630.2(r).
40. 42 U.S.C. §12201(c).
41. 42 U.S.C. §300gg.
42. Equal Employment Opportunity Commission, *Interim Enforcement Guidance on the Application of the Americans with Disabilities Act of 1990 to Disability-Based Distinctions in Employer Provided Health Insurance*, U.S. Government Printing Office, Washington, DC, 1993.
43. 70 F.3d 958 (8th Cir. 1995).
44. *Krauel v. Iowa Methodist Medical Center*, 95 F.3d 674 (8th Cir. 1996).
45. 42 U.S.C. §12181.
46. U.S. Department of Justice, *Title III of the Americans with Disabilities Act Technical Assistance Manual*, U.S. Government Printing Office, Washington, DC, 1993.
47. 121 F.3d 1006 (6th Cir. 1997), *cert. denied*, 522 U.S. 1084 (1998).
48. *Ford v. Schering-Plough Corporation*, 145 F.3d 601 (3d Cir. 1998), *cert. denied*, 119 S.Ct. 850 (1999); *Weyer v. Twentieth Century Fox Film Corp.*, 198 F.3d 1104 (9th Cir. 2000); *Doe v. Mutual of Omaha Ins. Co.*, 179 F.3d 557 (7th Cir. 1999), *cert. denied*, 120 S.Ct. 845 (2000).
49. 37 F.3d 12 (1st Cir. 1994).
50. *Pallozzi v. Allstate Life Ins. Co.*, 198 F.3d 28 (2d Cir. 1999), *amended on denial of reh'g*, 204 F.3d 392, (2d Cir. 2000).
51. 42 U.S.C. §12182.
52. 994 F.Supp. 1185 (N.D. Cal. 1998).
53. 42 U.S.C. §12132.
54. L.F. Rothstein, in M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997, pp. 317–331.

See also GENETIC INFORMATION, LEGAL, REGULATING GENETIC SERVICES.

GENETIC INFORMATION, LEGAL, REGULATING GENETIC SERVICES

GEORGE C. CUNNINGHAM
State of California-Department of Health Services
Berkeley, California

OUTLINE

INTRODUCTION

Genetic testing is expanding at an accelerating rate (1–5). Genetic tests were first used to screen for or diagnose a hereditary disorder in a given individual. This early testing was based on the determination of abnormal protein products or metabolytes produced by the mutant genes, such as sickle hemoglobin in sickle cell anemia or phenylalanine in phenylketonuria. Subsequently this was expanded to include testing for carriers of mutations with potential expression in subsequent offspring by identification of sequences of deoxyribonucleic acid (DNA) representing the actual mutation, such as sickle cell and cystic fibrosis carrier screening. Most recently it now includes testing for genes which, based on other genetic and environmental interactions, predispose to a wide variety of disorders, such as breast/ovarian cancer (BRCA1 and 2), cardiovascular disease, schizophrenia, and obesity. Genetic tests now include tests for gene products, actual genes, that is, DNA sequences, and abnormalities of chromosomal number and morphology. The field of gene therapy and preventative genetic engineering is just beginning to develop effective interventions.

The broadening of the scope of genetic medicine has not yet been recognized by public policy makers in legislative bodies, public health agencies, and the courts. As a

consequence there is only a relatively limited body of law and regulation that addresses the many unresolved problems this new technology presents.

The legitimate interests of government include protecting the liberty, privacy, health and safety of its citizens, promoting the welfare of the community, and arbitrating disputes between competing values and interests in the society so as to obtain the greatest good for the greatest number with due consideration of minority rights and values. A great deal of interest has been generated and legislation enacted that addresses the issues of the use of genetic information in insurance, employment, and research (6–8). Much of the preoccupation with this issue is due to the unique market-based philosophy applied to health care delivery in the United States. Discriminatory use of genetic information to deny access to or increase the cost of insurance is a symptom of this broader problem, and its importance diminishes greatly in other developed nations where health care is not dependent on employment and is available as a universal social benefit. Discrimination has been discussed in the press and in professional journals for some time and will not be discussed in this article. Rather, we will try to address the neglected issue of ensuring equitable access to genetic services that meet a minimal standard of quality at reasonable cost. The issues of how best to use this new knowledge to alleviate the burden of genetic disorders on individuals and the society requires a response irrespective of the problems associated with discrimination.

PUBLIC HEALTH ROLE

The core functions of public health have been listed by the National Academy of Science Institute of Medicine as assessment, policy development, and assurance (9). Federal and state public health agencies have made only a small beginning in fulfilling these functions in genetic services.

ASSURANCE OF QUALITY GENETIC SERVICES

Licensure

Federal and state healthcare agencies are charged with the general mission of protection of public health, safety, and welfare. One of the major administrative tools used to ensure this general objective is achieved is regulation. This includes regulation of persons and facilities providing services, that is, governmental licensure. A license is an official governmental document that allows the holder to perform certain actions that are prohibited to the general public. It is therefore a restriction of personal liberty. Any legislator who proposes a law that restricts personal liberty needs to convince the legislature and the public that the restriction is necessary to protect the public from a greater harm than the restrictions would impose and that there is no other effective or less restrictive way to prevent the harm. Based on these basic considerations, is there a case for governmental licensure of genetic personnel such as doctoral-level geneticists in cytogenetics, genetic counselors, genetic laboratory technology, genetic nurse specialists, and genetic facilities such as laboratories?

There are different concepts of what evidence would be sufficient to demonstrate public harm. Some would espouse a proactive stance and would be satisfied with evidence that supports a reasonable probability of harm.

Others favor a reactive philosophy and want to "count the bodies," such as document-specific instances of real harm.

PUBLIC EDUCATION

If citizens can protect themselves, there is no need for governmental involvement. There is little reason to believe the public can protect it self. The National Academy of Sciences in 1975 concluded: "It is essential to begin the study of human biology, including genetics and probability, in primary school, continuing with a more health-related program in secondary school.... Sufficient knowledge of genetics, probability, and medicine leading to appropriate perceptions of susceptibility to the seriousness of genetic disease and of carrier status cannot be acquired as a consequence of incidental, accidental, or haphazard learning..." (10). This has been reinforced by the Presidential Commission, the Institute of Medicine, the National Science Foundation, and other public and private groups in the intervening years. However, studies continue to document the low level of scientific literacy in the United States. The public at large is not well informed about genetics and genetic disorders and needs to rely on the services of experts. Five out of six never heard of genetic engineering. Only one-third of college graduates can correctly describe deoxyribonucleic acid (DNA). Surveys of scientific literacy have validated this deficiency (11–16). Erroneous genetic information or misinterpretations can lead to decisions not to get pregnant, terminate a pregnancy, stigmatization, loss of self-esteem, marital or familial disruptions, inappropriate and risky interventions, and the like.

PRIMARY CARE

Can we rely on the usual sources of care to protect the public from unnecessary services while assuring access to appropriate, high-quality services? With respect to genetics, the high probability of harm resulting from services provided by persons without specific training and experience in genetics is documented in several studies of genetic knowledge and practice of nongeneticists, such as, primary care practitioners (17–22). The inclusion of modern human genetics in the training programs for physicians and nurses is a relatively recent event. The number of human chromosomes and their relationship to disorders such as Down syndrome was not known until 1959. Prenatal diagnosis was introduced in the early 1960s. Use of DNA based tests in clinical medicine only began in the late 1970s. There are a large number of practicing physicians who have had little or no genetics in their training. Even after medical schools began to introduce genetics, the hours were minimal and frequently

elective. Few questions on genetics are included on medical licensing or specialty board examinations. The result is a generally unsatisfactory understanding and application of clinical genetics in primary care. This is not a reflection on the dedication and concern of primary care physicians, who could not reasonably be expected to keep up with the explosion of genetic knowledge, but simply a factual description of the current situation.

The occurrence or high probability of occurrence of harm in the absence of access to qualified genetic personnel has been demonstrated and accepted by most knowledgeable professionals who have studied this area, but the questions remains, Is licensing the only or best solution to the problem? What are the alternatives? One alternative solution would be improved education of primary care providers. This is certainly necessary in any event and could contribute to improvement of services but does not appear to be practical or effective as the only response. Primary care physicians have limited time for continuing education and many areas of clinical practice other than genetics are competing for this limited time. In addition they have limited time with their patients. Studies indicate an office visit includes approximately 11 minutes of face-to-face contact with the physician. The average genetic counseling visit is approximately 50 minutes (23–25).

PROFESSIONAL CERTIFICATION

Another alternative is self-certification by professional societies (26). The first effort in certification of genetic counselors was the American Society of Human Genetics (ASHG) program which began in 1979. This was followed in 1980 by the recognition by the American Board of Medical Specialties of the American Board of Medical Genetics (ABMG) with the first certificates issued in 1981 under the auspices of ASHG. Certificates were issued to genetic counselors, doctoral-level geneticists, and physicians. The American College of Medical Genetics (ACMG) was established in 1992 and recognized as a component society by the American Medical Association (AMA). In 1993 the American Board of Genetic Counselors (ABGC) was established to continue the certification of genetic counselors. There are now 1006 medical genetic specialists, 150 Ph.D. medical geneticists, and 779 genetic counselors with board certification in the United States. While this process was a major contribution to the resolution of the problem, it is again not completely satisfactory. The medical specialty is still not listed in telephone directories nor recognized by the majority of third-party payers as one to be included in panels of specialists or listed in directories of providers. Utilization guidelines and reimbursement for services for these specialists are still being developed. Failure to authorize or reimburse for genetic counseling done by nonphysicians has limited the use of genetic counselors and Ph.D. geneticists. Furthermore, in most states, any physician can legally provide genetic services without special training or qualifications or specialty board recognition. This situation allows any physician, nurse, or counselor to be self-designated as a genetic specialist or subspecialist, which is misleading to the public.

FUNDING

Another alternative to prevent harm is the use of federal and state payment for services as a means to require the use of qualified staff in delivery of high quality genetic services. However, this is only effective for those services eligible for state payment and leaves the citizens who are not eligible unprotected. The reimbursement requirements of Medicaid or children with special health care needs are examples of this approach.

LICENSURE

As a result of this analysis, the only effective solution is to legally recognize the training and expertise of genetic professionals by licensure. There are additional arguments for licensure in terms of development of quality genetic services. Frequently quality genetic services require consideration of complex questions of risks, benefits, conditional interpretation, various interventions of varying effectiveness, and the like, in brief, information that would be beyond the reasonable expectation of the scope of primary care. If genetic personnel were licensed, the state could require the use of licensed or otherwise qualified personnel for complex genetic problems in order to maintain the quality of services. The existence of a license would allow genetic counselors as recognized professionals to have some say in governmental policy affecting their field of interest. Licensure would also promote the creation and funding of positions needed for expansion of new genetic services. In order to expand training programs, it is necessary to create a defined pool of positions for the graduates. Licensure would allow the public to specify and request referral to these professionals as recognized by the state.

SCOPE OF PERSONNEL LICENSURE

What kind of genetic personnel need licensure? Physician geneticists are licensed as physicians and only require better recognition and utilization of their specialty board certification. There are a group of doctoral geneticists trained in human clinical genetics who are certified by ABMG as Ph.D. medical geneticists. The laws of most states do not allow such persons to obtain a physician's and surgeon's license, and therefore prohibit their clinical utilization. For example, a cytogeneticist is prohibited from diagnosing chromosome anomalies as the practice of medicine. Provided their clinically related services are limited to their area of training and expertise, provision should be made to legally recognize them as practitioners by certification or licensure. Masters and doctoral level genetic counselors are an essential part of quality genetic services and need to have their practice made legal, including registration or licensure. In the laboratory area there are four subspecialist certificates issued by ABMG: clinical cytogeneticist, clinical biochemical geneticist, clinical molecular geneticist, and clinical biochemical molecular geneticist. The 1997 directory lists 425 clinical cytogeneticists, 137 clinical biochemical geneticists, 143 clinical

molecular geneticists, and 49 clinical biochemical molecular geneticists (27). These are doctoral level categories, and they require candidates to pass an examination in general medical genetics as a precondition to taking the specialty examination. Individuals in these classifications are intended to function as laboratory directors of genetic specialty laboratories. The actual bench-level performance of tests is the job of the laboratory technologist. The complex techniques used by these technologists are not a part of traditional laboratory technology training programs. Recognizing the need for special training and certification in the field of cytogenetics, five California cytogeneticists organized an Association of Cytogenetic Technologists (ACT) in 1975. In cooperation with the National Certifying Agency for Medical Laboratory Personnel (NCAMLP), a national technologist certification program was developed in cytogenetics, which issued its first certificates in 1981.

The field of molecular biology, namely DNA analysis, was largely research oriented until the early 1990s. The California Department of Health Services contacted ACT and NCAMLP and requested a certification process be established in molecular biology. As a result ACT expanded its area of interest and in 1996 changed its name to Association of Genetic Technologists (AGT). NCAMLP responded by establishing a certification program for Certified Laboratory Specialist in Molecular Biology with the first examination given in July 1997. There is no specialty certification for technologists in biochemical genetics at this time.

FACILITIES LICENSURE

Federal and State Licensing Responsibilities

The federal government has not assumed responsibility for licensure of personnel. The basic law regulating laboratory practices, the Clinical Laboratory Improvement Act of 1988 (CLIA), does provide for certain minimum educational and experience qualifications for laboratory directors, technical supervisors, and testing personnel and describes their responsibilities (28). The federal regulations first require a current license issued by the state in which the laboratory is located. The only specific reference to genetics is the area of cytogenetics where the technical supervisor must have four years of genetic training or experience, two of which must have been in cytogenetics. There is currently an advisory committee working on improving the coverage of genetic laboratory personnel. All laboratories must be in compliance with CLIA, including the genetic specialty laboratories. The federal regulations cover areas of staffing, patient test management, quality control, proficiency testing, inspections, and sanctions.

States generally have laws that require licenses for clinical laboratories and laboratory personnel. New York (1972) (29) and California (1995) (30) recognize genetics as a laboratory specialty area. New York defines the qualifications for laboratory director and makes the director responsible for using qualified technologists and maintaining quality control. Specific standards are detailed for cytogenetics (1972), and genetic testing (1990). Proficiency testing and site visits are required. California is currently implementing a similar program. However, in addition California requires licensure of genetic technologists.

LICENSURE CONSIDERATIONS

Public agencies and legislatures, in considering a proposed licensure program, need to collect information in a variety of areas before the full societal impact can be assessed. This includes the numbers of personnel and their professional representation; what segment of the public is served, what is the position of public advocacy groups, is there duplication or competition with existing licenses, what is the nature and severity of harm to be prevented? Are there alternatives? What will licensure cost? What will be the effect on supply? What is the limit on the scope of practice? Are knowledge and skills testable? Are there approved schools to provide training? What is the economic impact?

LEGISLATION AND REGULATIONS

With or without licensure, laws can be passed regulating the provision of genetic services. Legislation has been used in California to regulate prenatal serum screening (31). The state law permits the Department of Health Services to specify standards for vendors participating in the statewide birth defect screening program, which is called the Expanded AFP program. This program is based on the well-documented association of specific patterns of analytes (alpha feto-protein, human chorionic gonadotropin, and unconjugated estriol) with increased risk of birth defects (neural tube defects, abdominal wall defects, Down syndrome, and other chromosomal defects). The law also requires that all women seen before the twentieth week of gestation be provided information about the screening and be offered an opportunity to be screened by the program. If the woman elects to be tested, she signs an informed consent. Specimens are collected and transmitted to the laboratory. Specimens are analyzed in one of eight regional private laboratories under contract to the department. These laboratories use uniform methodology and are subject to daily quality control by the state. All data and laboratory results are communicated to a central computer in Berkeley. All persons who are judged to be high risk by the central computer algorithm are authorized, at no additional charge, to receive follow-up diagnostic services at one of 29 state-approved Prenatal Diagnostic Centers. The follow-up services include genetic counseling, ultrasound examination and, if necessary, amniocentesis, amniotic fluid analysis, and karyotyping. Any facility that meets state standards can be designated an approved vendor. The standards require that the prenatal diagnostic center be directed by a board-certified medical geneticist, that genetic counseling be provided by board-certified genetic counselors, that ultrasound examinations be performed by specially skilled and experienced ultrasonologists, and that amniocentesis, if indicated, be performed by experienced perinatologists/obstetricians. The state

collects a participation fee from third-party payers or from the participant, which covers all operating costs and is used to reimburse vendors. This public–private partnership design has succeeded in providing universal access to high-quality services.

TECHNOLOGY ASSESSMENT

In addition to prevention of adverse consequences of genetic disorders by promotion and regulation of personnel, the state is obligated to provide another kind of protection, namely protection from premature promotion of tests and substandard services that could adversely affect individual citizens or the community at large. This establishes the public health department as the primary technology assessment agency. Technology assessment really is a process of reviewing the scientific evidence and the information and opinions of experts to determine if a given technology should be applied in clinical practice and under what circumstances and conditions.

FEDERAL ROLE

Decisions on the appropriate implementation of any new genetic testing program are currently not centralized (32). There was an Office of Technology Assessment (OTA) established in 1972 to conduct assessments for the Congress. The mandate was broad, encompassing any technological problem, and OTA did publish some studies of heath technologies. OTA was abolished in 1995. In 1989 the Institute of Medicine published a monograph recommending a national technology assessment agency (33). The Congress established the Agency for Health Care Policy and Research (AHCPR) in 1989, but again, it has a broad area of responsibility and has not included many genetic technologies in its reviews. The National Institutes of Health (NIH) has established a mechanism called a consensus conference where a panel hears presentations from experts, reviews the literature, and publishes consensus statements on technologies. The federal Food and Drug Administration (FDA) has been proposed to play this role through their regulation of diagnostic kits and devices (34,35). While they could regulate clinical accuracy and utility, they do not have the authority to regulate the ancillary clinical setting in which the test is used.

STATE ROLE

Given the numbers of technologies being proposed or currently in use without rigorous analysis, all these efforts contribute useful information and should be encouraged. However, it is important to establish technology assessment capacity at the state level since the states play such a critical role in regulation and funding of health care. The legislature of the state of Maryland established a Maryland Commission of Hereditary Disorders in 1973, which reviews genetic tests, but this model has not been adopted by other states. There is a legitimate role for the state regulatory process using input from both experts and the public. The regulations can prohibit unvalidated testing, when the preponderance of evidence indicates that the public either individually or collectively could be harmed, except as part of a research project. The state can impose conditions on genetic testing by regulation when the evidence indicates such conditions are necessary. These conditions might include specialized informed consent, confidentiality, pre- and/or post-test counseling, protective measures, quality assurance requirements, record keeping requirements, availability of diagnostic and intervention resources, and so on. This approach could include use of state accredited or registered personnel who had training essential for the proposed testing program. Finally, once a technology is accepted by the experts and the public using the evidentiary process, the state has an obligation to promote equitable access. This could involve mandating that all public and private payment sources pay for any cost-effective technology, funding screening centers, and conducting public and professional education or outreach with the at-risk public. One example of the states' carrying out this function is the newborn screening for genetic disorders.

PRIVATE SECTOR

In addition to state regulatory practices, most insurance companies and managed care organizations have technology assessment groups. The criteria used by these private groups can be unduly influenced by cost considerations and it is not unusual to find a test accepted by one payer and regarded as experimental and not accepted by another.

PROBLEMS WITH LEGISLATION AND REGULATION

While laws and regulations have undoubtedly saved lives, prevented disease and disability, and increased the value of the goods and services, the potential for regulatory abuse and damage is all too apparent to the public. The arguments made in favor of an expanded activist role for public health in regulation of genetic services could be seriously undermined by failure to avoid the situations that have contributed to the current low esteem accorded the use of this governmental tool. The first error is what is referred to as "agency capture" where the special interest group affected by the regulations controls directly or indirectly the governmental regulators. Regulation should not be used to increase incomes of specialists through unnecessary restriction of services. Licensure laws can be used to exclude qualified providers in order to maintain incomes. Facility standards can be used to monopolize services and improve their economic outcomes. The public and the regulators should be aware of this tendency and should maintain an open public process that remains focused on the goal of assuring universal access to comprehensive high-quality, cost-effective services. While recognizing the contributions and qualifications of such professional groups as ACMG and the National Society of Genetic Counselors (NSGC), the state should remain open to including others, such as pathologists with subspecialty training in cytogenetics or molecular biology and nurses with subspecialty training in genetics

as qualified providers. Special certification of training in hereditary cancer counseling, sickle cell counseling, and cystic fibrosis genetics can be used to create a pool of qualified personnel to implement specific screening programs. Finally, involvement in policy development of the increasingly active groups of individuals and families affected by genetic disorders and the general public can exert a corrective action.

Part of the process of effective regulation is the follow-up of enforcement and monitoring. Simply putting a requirement in a statute or regulation is no guarantee that the system will implement the requirement as intended. Failure to enforce regulations is another factor undermining public confidence in government's interest and ability to represent the public interest. On-site monitoring, chart reviews, records maintenance, reporting outcomes or events, and the like, are required on a continuous basis to ensure uniform application of the law and community-wide implementation. While these add to the burden of regulation they are essential if the benefits of regulation are to be real instead of imagined.

Another problem is the incompetence of some of the regulators. It is difficult to provide in the public sector the kinds of salaries that will be guaranteed to attract the kind of genetic expertise needed. It is important that regulators have access to expert consultants and adopt policies and processes that permit input from a broad variety of genetic and nongenetic professionals, as well as affected members of the public. Qualifications of regulators should include familiarity with the field of genetics, public health, law and administration, and the way the health care system operates. The regulatory agency should have sufficient resources and visibility to be able to develop and implement effective programs and formulate and enforce regulatory standards. The regulations should reflect a consensus of affected parties as to the minimum requirements of currently accepted standards of care, and not utopian efforts to provide cutting-edge technologies to anyone who might possibly benefit.

CONCLUSION

The rapid development of genetic knowledge and technology poses problems, both familiar and novel, for the society. As the representative of the public, the federal and state public health agencies need to be prepared to ensure equitable access to quality genetic testing for high-risk populations. Regulation of personnel and facilities providing testing can play a constructive role in assuring that the inevitable adverse consequences of testing are minimized and benefits are maximized with fair treatment of all the involved parties and interests.

BIBLIOGRAPHY

1. F. Collins, *N. Engl. J. Med.* **341**, 28–37 (1999).
2. M. Khoury, *Am. J. Pub. Health.* **86**, 1717–1722 (1996).
3. *Science*, The Genome Issue **286**, 443–491 (1999).
4. J. Bell, *Br. Med. J.* **316**, 618–620 (1998).
5. *Time*, January 11 (1999).
6. K. Rothenberg, *J. Law Med. Ethics* **312**, 312 (1995).
7. K. Rothenberg et al., *Science* **275**, 1755–1757 (1997).
8. A.M. Capron, in *Proc. Conf. on The Implications of Genetics for Health Professional Education*, 1999, pp. 123–162.
9. Institute of Medicine, *The Future of Public Health*, National Academy Press, Washington, DC, 1988.
10. L.B. Andrews et al., eds., *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994, p. 186.
11. J. Miller, *Report to the National Science Foundation*, NSF, Washington, DC, 1992.
12. March of Dimes Birth Defects Foundation, *Genetic Testing and Gene Therapy*, Louis Harris and Association, White Plains, NY.
13. March of Dimes Birth Defect Foundation, *News Release*, September 29, White Plains, NY, 1992.
14. K.J. Hodgkinson, *J. Med. Genet.* **27**, 552–558 (1990).
15. L.B. Andrews et al., eds., *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994, p. 186.
16. K.J. Hofman et al., *Acad. Med.* **68**, 625–631 (1993).
17. N. Holtzman, *J. Clin. Ethics* **2**, 1–5 (1991).
18. N. Holtzman, *J. Clin. Ethics* **2**, 1–2 (1992).
19. E.S. Tambor et al., *Am. J. Pub. Health* **83**, 1599–1603 (1993).
20. D.C. Wertz, *Am. J. Hum. Genet.* **61**, A193 (1997).
21. A.D. Kline, G. Halpin, and I.S. Mittman, *Am. J. Hum. Genet.* **61**, A189 (1997).
22. L.B. Andrews et al., eds., *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994, pp. 116, 119.
23. B.A. Bernhardt et al., *Obs. Gyn.* **91**, 648–655 (1998).
24. N.C. Blumenthal et al., *J. Family Pract.* **48**, 264–271 (1999).
25. T.G. Ferns et al., *Arch. Ped. Adoles. Med.* **152**, 227–233.
26. V.A. McKusick, *J. Am. Med. Assoc.* **270**, 2351–2356 (1993).
27. American Board of Medical Genetics, Available at: *http//www/faseb.org/genetics/abmg/stats.allyears.htm*
28. Title 42 code of Federal Regulations Part 405 et seq., Section 493.1449.
29. New York Public Health Law, Title V, Article 5.
30. California Business and Professions Code Section 1202.5 et seq.
31. California Health and Safety Code Part 5, Article 1, Hereditary Disorders Act, Section 124975 et seq.
32. N. Holtzman, *Science* **286**, 409 (1999).
33. L.B. Andrews et al., eds., *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994, p. 13.
34. B.S. Wilford and K. Nolan, *J. Am. Med. Assoc.* **270**, 2948–2954 (1993).
35. National Academy of Science Institute of Medicine, *Assessment of Diagnostic Technology in Health Care*, National Academy Press, Washington, DC, 1989.

See other entries GENETIC INFORMATION, ETHICS, AND INFORMATION RELATING TO BIOLOGICAL PARENTHOOD; GENETIC INFORMATION, ETHICS, INFORMED CONSENT TO TESTING AND SCREENING; GENETIC INFORMATION, LEGAL, FDA REGULATION OF GENETIC TESTING; GENETIC INFORMATION, LEGAL, GENETIC PRIVACY LAWS; GENETIC INFORMATION, LEGAL, GENETICS AND THE AMERICANS WITH DISABILITIES ACT.

H

HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY, ETHICS, COGNITIVE ENHANCEMENT

PETER J. WHITEHOUSE
Case Western Reserve University
Cleveland, Ohio

CYNTHIA R. MARLING
Ohio University
Athens, Ohio

OUTLINE

INTRODUCTION

Have you ever smoked a cigarette or drunk a cup of coffee and felt more alert? Do you know someone who wears a hearing aid and attends better to oral language, thus remembering things better? Do you want your surgeon to be using the latest technological equipment in treating your condition? Should Indian chess players participating in international competition be permitted to consume bramin, which is used widely in their country for cognitive enhancement? Do you believe that electronic technology will play an increasing role in education? Do you think that biological and information sciences will lead to new knowledge that will enhance our ability to pay attention and remember? Do you believe that, particularly at this time in the history of the human race, such enhancements in our ability to think may be especially critical? Do you think wisdom is a desirable individual and social goal?

This article presumes that many readers will answer some of these questions in the affirmative. People already consume products and participate in activities that they believe will enhance their thinking abilities, and they are inclined to seek out the latest advancements in these areas (1). Perhaps this interest reflects the fast pace of life in modern times and the desire for individuals and groups to obtain competitive advantage in the worlds of business and education.

This article examines some of the underlying ethical issues that relate to cognitive enhancement. We will explore these issues after a discussion of what we mean by cognition and by enhancement (2). We believe that interest and activity in this area of life will only increase in intensity in the future.

Cognition

The term "cognition" could perhaps be replaced with the word "thinking." However, clinicians and researchers use the term to refer to a variety of intellectual skills, including attention, learning, memory, language, skilled motor behaviors, and perceptual abilities. In addition it often encompasses so-called executive functions, such as goal setting, planning, judgment, problem-solving, and decision-making.

Most of the literature on cognitive enhancement addresses methods of enhancing either attention or memory through medications. Thus drugs may help one stay awake and attend to stimuli in the environment or remember past or upcoming events better. Undoubtedly, we would be even more excited by cognitive enhancers that improve so-called higher level thinking, such as the executive functions and even wisdom (3). All intellectual abilities depend on adequate arousal. (Note that it is difficult to learn new material when in a coma.) Arousal is intimately related to attentional mechanisms. Another fundamental intellectual task is to selectively attend to important stimuli in the environment and to avoid distraction by those less critical. Interventions that affect the speed of processing, allocation of attention, and accomplishment of complex tasks might be expected to improve higher level decision-making and problem-solving as well.

Cognitive Enhancement

The term cognitive enhancement is usually used to differentiate this concept from enhancement of noncognitive or

emotional abilities. The ability to appreciate and detect the wide range of human emotion both in oneself and in others is also a critical ability that deserves consideration for enhancement. Clearly, alterations in mood, such as depression and anxiety, can impair decision-making and problem-solving. Drugs such as Fluoxetine (Prozac) can help clinically depressed patients, but whether interventions can enhance emotional abilities in normal individuals remains uncertain (4,5). Nevertheless, presumably normal grocery shoppers snap up herbal teas and substances like St. John's Wort in hopes that they can.

The distinction between cognitive and emotional capabilities is somewhat arbitrary. Fundamental psychological functions, such as arousal and motivation, suggest that basic distinctions between thought and feeling deserve scrutiny. However, for this article we will not focus on enhancement of noncognitive abilities, although we will make brief mention of the possibilities of biological intervention in this domain. We should also point out that one of the ultimate goals of enhancement technologies, that is, to improve human wisdom, would undoubtedly involve improvement in both cognitive and emotional skills, as wisdom represents that integration of high-level processing of both thoughts and feelings (6).

Biological Substrates

We are beginning to understand more about the biological substrates of human cognition, opening this as an avenue for the development of enhancement technologies through molecular biology and neurochemistry. Much of this understanding has come from the study of diseases such as Alzheimer disease and other dementias (7,8). Dementia is the medical term for loss of cognitive abilities in more than one domain, and Alzheimer disease is the most common dementia. Alzheimer disease is biologically characterized by the loss of specific populations of nerve cells in association with specific pathological features observable under the microscope. Relating the loss of particular populations of cells to clinical symptoms has been the Holy Grail of clinical pathological correlation in Alzheimer disease and related disorders.

One of the biological systems underlying learning and memory is the cholinergic basal forebrain (9). Loss of nerve cells in this structure, located deep in the brain underneath the basal ganglia, is a substrate for the cognitive impairment found in Alzheimer disease and some other disorders. The basic scientific evidence for this is fairly convincing in terms of the effects of damage to the structure in animals on learning and memory. Drugs that block cholinergic systems can cause memory problems in normal human beings. Most important from a therapeutic point of view, drugs such as Donepezil, which enhance cholinergic function (they are cholinesterase inhibitors that work by blocking the enzyme that breaks down acetylcholine), improve attention and memory in these conditions. Although the effects are moderate in size, they have been definitely and conclusively demonstrated in double-blind placebo controlled studies (8).

There is no specific biological marker to differentiate normal aging from Alzheimer disease. Some degree of dysfunction to the basal forebrain occurs as we all age. A variety of labels (10,11) have been applied to this condition, ranging from benign senile forgetfulness to aging associated memory impairment to the term most commonly used today, mild cognitive impairment. These labels are applied to individuals who are usually older but do not have significant enough cognitive impairment to be considered demented. For example, memory problems rarely affect function in daily life. Thus a variety of trials are underway to enhance memory in individuals who are labeled with these conditions. They are by definition not demented and hence, normal. Therefore these trials really represent attempts to enhance cognition in normal individuals.

In addition to interventions that enhance cognition symptomatically, science is attempting to develop therapies that may slow the progression of conditions like Alzheimer disease that are due to gradually increasing loss of nerve cells in structures such as the cholinergic basal forebrain. A variety of approaches are being used, ranging from antioxidants such as vitamin E, antiinflammatory agents, and compounds that act to enhance the viability of nerve cells, such as nerve growth factors. Trials are also underway to treat patients who have mild cognitive impairment with these agents to delay the onset of Alzheimer disease, which would occur in a significant number of these individuals. Hence two forms of treatment are currently being used in normal people to try to enhance cognition, one symptomatical and the other preventative.

We should hasten to add that the therapeutic targets biologically are more than just the cholinergic system. Neuronal loss occurs in the locus ceruleus, which uses the neurotransmitter noradrenalin and the raphae nuclei, which use serotonin. The clinical consequences of loss of cells in these populations are less clear. However, drugs that act on these transmitters can affect cognitive abilities such as attention, as well as mood (4,12). Antidepressant medications work to enhance neurogenetic and serotonergic functions. For example, drugs such as amphetamine and Ritalin can affect mood and attention in normal individuals, and are used to treat Attention Deficit Disorder in children and adults.

Cholinergic medications may improve behavioral symptoms, namely noncognitive symptoms in dementia. Thus, again, we are reminded that the distinction between things cognitive and things affective is not always easy to determine either at a clinical or biological level. Although biological approaches are promising and growing increasingly so, most of human history has focused on nonbiological approaches to enhancing thinking abilities. Schools have been widely used throughout history to enhance thinking ability in children and adults, although their effectiveness has been under increasing scrutiny. Various assistive devices have been employed, such as the Chinese abacus, which was one of the earliest. Varieties of memory assistive devices have been used probably since the advent of commerce. However, the power of information sciences is increasing as rapidly or perhaps more rapidly than that of the biological sciences, offering other forms of enhancement possibilities for cognitive abilities. As computers and personal digital assistants proliferate and become more intelligent, the symbiosis between individual people and their computers becomes greater.

We have chosen here to focus on both biological and information systems enhancement and their ethical implications, although at first glance they would appear to be quite different. However, devices are currently used clinically in which microcomputers embedded in people control the infusion of biological substances designed to improve cognition (e.g., insulin pumps in diabetics). In principal, as biochips become increasingly sophisticated, enhancement technologies would likely include both silicon- and carbon-based approaches.

THE NATURE OF ENHANCEMENT

The article by Eric Juengst (13) in this encyclopedia provides the framework for our considerations of enhancement. Moreover we have benefited from work focusing on enhancement in sports (14). The area of genetic enhancement and physical enhancement has received more attention (15), for example, in the area of cosmetic surgery (16). Enhancement has been useful to limit the domain of medical practice, as it is usually considered to lie at the fringe of the scope of medicine. However, we will focus on the second meaning that Juengst gives to enhancement, which is the notion of self-improvement. The limiting of the domain of medicine is an interesting area in cognitive enhancement. The success of the medical establishment in identifying Alzheimer disease as a treatable condition requiring medical research and intervention has been noted. Interestingly the nature of Alzheimer disease as a disease is being challenged by the approach to try to enhance cognition in people with so-called mild cognitive impairment. However, we will focus principally on the notion of cognitive enhancement as improvement rather than treatment.

IMPORTANCE OF COGNITIVE ENHANCEMENT

Individual Use Pattern

It seems evident that individual human beings have decided that cognitive enhancement is a goal worth pursuing. Not only do we put considerable energy into going to school, but we also pursue increasingly faster individual computers and frequently take biological interventions ranging from coffee to complementary alternative measures in seeking out enhanced thinking ability. Global expenditures of billions of dollars, far exceeding expenditures on prescription drugs for all conditions, let alone those designed to enhance cognition, demonstrate the importance that individuals assign to this area, as this money is spent out of pocket (17). The range of biological products that have claimed to enhance cognition is enormous (1). Perhaps the most commonly used drug for this purpose worldwide is ginkgo biloba. The scientific evidence that ginkgo biloba helps in any disease state is inconclusive. The evidence that it enhances normal cognition is even weaker.

An entire class of drugs has been referred to as noototropics, meaning mind growth. The original compound, piracetam, has been demonstrated to improve learning and memory in a variety of animal models. Its effects in human beings are limited, however. Yet compounds in the same class such as nefiracetam are under active investigation for treating stroke and dementia. The power of genomics, combinatory chemistry, high throughput screening, and a variety of other approaches available in the industry, make it reasonable to think that more effective medications will be developed in the future, not only for diseased conditions, but for enhancing normal thinking.

Understanding Complex Systems

The stakes for cognitive enhancement go far beyond the individual economic performance of human beings. The human race is facing complex challenges to its very survival (18,19). For example, it seems evident that human beings have had significant impact on their environment and that of other species. A topic such as global warming and the controversy surrounding it illustrates the difficulties that human beings have in understanding the behavior of complex systems, such as our biosphere, and projecting the effects of our behavior in the present onto the state of our biosphere in the future. Computer models can be used to make projections about the viability of life on this planet, and thus we are already using information technologies to help us analyze complex system behavior. It is obvious that if we had biological and information science interventions that could enhance our ability to understand the consequences of our own behavior in the present, this would be a tremendous advancement for future generations. This would be particularly important if the cognitive abilities enhanced were, in fact, executive functions and even wisdom, improving the human brain's ability to model the consequences of present behavior on future states.

Thus we believe that the current interest expressed by individual human beings in enhancement could be reason enough to consider the ethical issues seriously. However, the need to enhance our cognitive abilities to help ensure the sustainability of life on this planet raises the stakes even further. It seems clear to us that thinking through the ethical issues surrounding cognitive enhancement warrants serious consideration.

ETHICAL ISSUES COMMON TO BIOLOGICAL AND INFORMATION INTERVENTIONS

Goals

If the goal of cognitive enhancement using either drugs or computer systems were self-evidently a desirable outcome, then the enhancement of higher level thinking, such as wisdom, would seem to be especially desirable. However, it is quite possible that enhancement of selective areas of cognition would not necessarily result in overall improvement. We already exist in a world in which people find the pace of life rapid. Would a drug that merely improved the ability of an individual to think more things be a desirable product? Would this merely focus attention on quantitative rather than qualitative outcomes? Would we create increased unhappiness by driving people forward to greater and greater productivity? Enhancing cognition might have detrimental effects on

the broader personality. One is reminded of the warning of Spock on *Star Trek*, that there might be sacrifices to be made in the emotional life by enhancement in the cognitive sphere. Surely this is not necessarily a consequence of cognitive enhancement, but it is one worth being aware of. After all, we do live at a time that still celebrates the rational values of the enlightenment. Focusing on enhancing thought without feeling, that is, knowledge without wisdom, might in fact contribute to further self-destructive and society-damaging behaviors.

Costs

Another ethical issue common to enhancement technologies, be they drugs or computers, represents the issue of costs. How much energy should human society place into trying to develop interventions to improve cognition? Of course, this is a difficult question to answer because one can never predict the results of scientific research. If very effective interventions were developed at low costs, this might be desirable. However, it is a daunting challenge to enhance human cognition in any way, and thus societal resources could be invested out of proportion to the likelihood of success.

Justice

The fact that enhancement technologies already do cost considerable amounts of money raises the ethical issues surrounding justice and access (20,21). This topic is already of considerable interest to health care professionals and individuals who have thought about universal access to computer technology. Would the availability of even more effective enhancement technologies increase the already growing distance between the have and have not countries, as well as the have and have not populations within countries?

Risks and Benefits

The introduction of new technologies raises the issue of risks and benefits. Although we have talked mostly about the positive outcomes of enhancement technology, all technologies have potential downsides. Medications to enhance cognition that are given to healthy people would need to have a low chance of significant side effects in order to justify their use. Who would decide what level of enhancement is worth what level of risk? In the United States and in most countries, the regulatory authorities in governments attend to the issues of risk benefit and disease but have limited jurisdiction over enhancement technology designed for individuals who suffer from no illness. Clearly, at a more macro level, the Y2K problem illustrates the risk of dependence on information technology. Will we someday regret that we have become so dependent on computers that human lives can be lost as a result of power failures or other system crashes?

ETHICAL ISSUES RELATING PRINCIPALLY TO BIOLOGICAL INTERVENTIONS

Interventions designed to change the biology of an individual human being present some ethical issues, which if not unique, are at least more obvious in relationship to this form of intervention. We might have included the issue of risk-benefit in this discussion, as various adverse events are more likely to be associated with medications than with the use of computers. However, pills seem to create the ethical issue of an artificial road to enhancement. Steroids taken by athletes are viewed as creating an unfair competitive advantage. While improving athletic prowess through diet, exercise, videotape feedback, and even computer analysis of physical motion is viewed as laudatory, taking pills to improve athletic prowess is not. Thus drug tests are now de rigeur at sporting events, and athletes who are found to have taken performance enhancing drugs are disqualified. Are drugs equivalent to hours of training? Will a cognitive enhancing pill that replaced hours of toil and sweat in a classroom similarly be viewed as some kind of inauthentic perversion (22)? However, distinctions between artificial and natural are actually difficult to make and, if the pills were relatively safe, most individuals might believe that this form of enhancement would be appropriate.

ETHICAL ISSUES RELATIVELY UNIQUE TO COGNITIVE ENHANCEMENT WITH INFORMATION TECHNOLOGY

Clearly computers are a more evident technology that enhances human beings' ability to remember and problem solve. What is clear also is the rapid advance in the intelligence of computers. What are the ethical issues that relate to enhancing human cognition through the use of computers?

Confidentiality

The first issue relates to that of confidentiality (23). Computers enhance human cognition in part by allowing information to be shared more quickly between different individuals, for example, through the use of e-mail. Yet this easy distribution also allows for the possibility that more information becomes available essentially to the entire world. Who has not received an e-mail entitled something like "learn everything about everybody"? Therefore violation of privacy becomes a major concern in using information technology to enhance human thinking abilities.

Computer-Assisted Information Processing

As mentioned in the beginning, in terms of environmental issues, solving complex health and resource problems is becoming increasingly difficult. Within the area of medicine, for example, it is becoming increasingly difficult to know which medical interventions to offer which patients. It is ironic that almost a century after the beginning of so-called scientific medicine, we are now promoting the notion of evidence-based medicine and health care. We need to take seriously the moral obligation to use optimally the information that has been collected to make individual and population health decisions. One approach that has been used as part of that evidence-based medicine is meta-analysis (24). Individual clinical trials often provide useful information

that affects the behavior of clinicians, for example, which drugs to use, in which quantities, and for what conditions. However, many studies are equivocal in their interpretation. Hence the notion arises of examining conclusions that might be drawn by reviewing the entire body of information available about a particular intervention. In other words, meta-analysis synthesizes the results of studies, viewing them not just as individual protocols, but as a sum total acknowledgment in a particular domain.

Meta-analysis has, however, engendered considerable controversy, being described variously as obvious, necessary and wise, or statistical fakery. The controversy can be seen, for example, in the meta-analysis of the effects of secondary smoking, namely smoke inhaled by individuals who do not themselves smoke but who are in the environment of smokers. It seems evident that it is worthwhile examining all the evidence available on a particular topic, such as the effect of passive cigarette smoking, but this must be done in a rigorous way.

Missing Data

A further ethical issue has to do with the availability of all the information about a particular intervention. Academics and drug companies, for different but related reasons, do not like to make so-called negative studies available through publication. Reporting on something when a study does not work cannot advance one's career or one's bottom line. Yet a systematic and important bias is introduced into the domain of knowledge about an intervention if only positive trials are reported. Thus it seems apparent that there should be ethical obligations to publish negative results as well as positive results, or else clinical decision makers will be misled when they review the available literature.

Cognitive Prostheses

The branch of computer science most intimately concerned with cognitive enhancement is artificial intelligence (AI). A recent trend in AI research is toward building "cognitive prostheses," systems that amplify a human problem solver's own thought processes. As used in this context, the term "prosthesis" includes enhancement as well as treatment so that both hearing aids and stethoscopes would qualify as prostheses. The goal of building cognitive prostheses is to enhance the power of even the finest human mind (25). The vision is not to build an all purpose problem solver that makes its user generally smarter but to build a series of special purpose computer tools, each of which symbiotically interacts with its user to solve one particular type of problem. Thus one tool might help a lawyer to plan a more brilliant defense, and another might help a professor to write a more interesting lecture. There is no theoretical limit to the type of cognitive activity that might be enhanced. Clearly, if that activity were itself unethical, such as plotting a perfect crime, then using cognitive prostheses to support it would also be unethical. However, not all of the ethical issues involved are so obvious.

Computer Program Error

One new issue is that of determining who is responsible if the computer makes a mistake. Programming errors in devices designed to administer doses of radiation to treat cancer have occurred and led to human death (26). AI systems are not infallible, and they might conceivably offer bad advice or focus a user's attention away from critical information. While legal responsibility is still solely assigned to human beings, other possibilities for moral responsibility have been proposed for the case in which computers autonomously make bad decisions. One possibility is to think of a computer as an agent, which can be liable for harm, just as human medical assistants can be. Another possibility is to allow that no one may be responsible for faulty judgment rendered by machines. The rationale for the latter position is that the use of truly life-enhancing technology should be encouraged, and blame is clearly discouraging (27).

It has been suggested that the reason responsibility is so hard to assign when computers make mistakes is that the norms for building good computer decision systems are not well understood (28). There are no accepted standard practice guidelines, as there are in medicine, to help determine if system designers and programmers have done everything reasonable to ensure the goodness of a computer system. We must depend on the integrity and skill of system developers to ensure dependable, accurate systems. Fortunately, in symbiotic systems, in which people and computers work together, the human partners may serve as valuable safeguards, recognizing when computer outputs seem dubious.

Symbiosis

Still other issues arise from the symbiotic nature of the relationship between the human user and the computer system. A central goal of AI research is to create machines that think like people. Whether or not that goal is ever fully achieved is open to technical and ethical debate. However, there can be no doubt that great strides have already been made. So it is entirely possible that a cognitive prosthesis could function like a virtual colleague. How will we relate to our virtual colleagues? Will we come to depend on them, feel emotionally attached to them, and even debate ethical issues with them, as we might with real colleagues? If so, will this enhance our professional lives, or merely reduce the social interaction we might otherwise have with real colleagues? What roles will we allow our virtual colleagues to play, and what types of activities will we reserve for human beings?

CONCLUSIONS AND FUTURE DIRECTIONS

In this article, we have reviewed a broad collection of issues relating to human values and hence ethics surrounding the use of biological and information technologies to enhance cognition in normal people. We have tried to illustrate our case by examples of behavior and practice in evidence today. It seems that there are important reasons to try to enhance individual and social intelligence using both drugs and information systems. What issues might we

see emerging in the future if we momentarily take the viewpoint of the science fiction writer?

Cyberspace

First, recall that it was William Gibson who introduced the notion of cyberspace some years ago in his book *Neuromancer* (29). The overlap between biological and computer enhancement was evident in this work, as the protagonists would choose from a wide assortment of stimulants and related biological compounds before "jacking in," that is, creating a direct biological link to the computer network before they immerse themselves in the shared mental space known as "cyberspace." To some extent we are already involved in cyberspace. Multiuser domains are spaces created electronically in which people can interact with so-called avatars, where a visual image can be created and observed to interact with other individuals in that space. It is quite possible to adopt a false identity, even to change age and gender, and to interact in social circumstances that can cause benefit or harm to other participants in this space. Admittedly, the power of interaction in this kind of space is more limited than that envisioned in *Neuromancer*, but it is certainly a start. We are all aware of stories of people meeting on the Internet who either marry or kill each other.

The notion that drug enhancement can be combined with computer enhancement is clearly evident in those of us who brew a cup of coffee before answering our e-mail. The likelihood of direct biological connection to the Internet is also not so farfetched. Already we can wear headphones and goggles that permit us to enter the world of virtual reality generated by computers. From its earliest applications in space exploration and flight simulation, virtual reality has grown to encompass surgical simulation, virtual anatomy for medical education, and artificial threatening environments for use in the psychiatric treatment of phobias (30). Yet the Internet is also full of virtual reality "games" featuring countless fictitious creatures, as well as representations of real people, who have been killed, maimed, or destroyed. Ensuring that this powerful technology is used only for societal benefit becomes a new moral imperative.

It is a small step to recognize in an individual with a cochlear implant or some visual assistive device that involves the interaction between their own individual assistive device and their nervous system to directly connect this device to the Internet. It is not so farfetched to imagine a plug on a cochlear implant that would allow a direct connection to music obtained and downloaded from the Web.

Computers as Moral Players

We will conclude on one, perhaps most distant and yet provocative issue that exists in *Neuromancer* but not in reality yet. We do already meet intelligent entities on the World Wide Web that are not human beings or even the manifestations of live human beings. Many have played a game of chess where the opponent is a computer. At what level of intellectual capacity would a computer have to exist before we would give it moral status? Perhaps the answer to that question is an infinite amount, since moral status is not granted on the basis of cognitive abilities alone.

Yet other types of human abilities are already being envisioned for computers and built into working computer systems. Kurzweil has written a book called *The Age of Spiritual Machines* (31,32), which asks whether emotional and other more affable human abilities can be programmed into a computer. AI researchers are currently building systems that can understand and model human emotions (33). Initial results in the field of affective computing may seem modest. Computers sense human emotions, like frustration, in computer users so that they may better respond to user needs. Virtual animals and cartoon characters are set in virtual worlds, where they act in accordance with their levels of hunger, fear, playfulness, aggression, and desire for affection, rather than according to programmed scripts. However, the research goals are far from modest. Following neurological findings that rational thinking may be affected by too much or too little emotion, AI researchers seek to enhance computers with the abilities to recognize, possess, and express human emotion. Will these new capabilities enhance computers to the point where we might afford them moral status?

Certainly science fiction writers have already addressed this topic. We could be forming the moral relationships between silicon-based information entities and carbon-based entities. Isaac Asimov developed an entire world based on the three principles of robotics that define the moral obligation of intelligent robots towards human beings (34). The first principal of robotics is that robots must not harm human beings nor allow them to come to harm through inaction. Who would determine the three laws of human beings that would govern their behavior toward complex computer systems? Destroying a computer system would certainly cause significant moral harm to human beings dependent on it, but at what point would we raise concern about destroying the computer itself?

Yes, this does seem farfetched, but it is not unrelated to growing concerns in bioethics about the moral relationships between human beings and other biological information processing entities. As concern about the environment continues and the relationships between human beings and other life forms become more fully understood, it seems evident that we should have moral responsibilities to other creatures in our biosphere and to the biosphere itself. We can ask whether we have a greater moral responsibility to a chimpanzee than an amoeba. We can ask whether we have a greater moral responsibility to a chimpanzee than an anencephalic child. We can ask at what point do we have a moral responsibility to a highly complex information system compared to a simple biological entity. Would you be willing to consider that the Internet has distributed intelligence and good moral purpose in its own right, beyond the effects it has on human beings? Would you sacrifice a single amoeba and give it less moral status than a distributed information system being considered for termination? If you are willing to take this moral step, then when in the process of the evolution of a biological and a computer information entity does this moral shift occur, if ever?

Importance Revisited

Life will be very different in the future because biological life itself will be changed, and there will be change in large part because of the availability of human beings to manipulate and create life forms that are biological. The future will be also dramatically different because of our ability to create different information processing entities. In fact the very survival of the human race depends on the responsible use of powerful biological and information technology. Clearly, we should give thought to the ethics of cognitive enhancement now and hope that some of these improvements in biology and information technology will someday assist us in being wiser about the use of technology to enhance human thinking.

BIBLIOGRAPHY

1. D.J. Ward, *Smart Drugs II: The Next Generation*, Health Freedom Publications, Menlo Park, CA, 1993.
2. P.J. Whitehouse, E. Juengst, M. Mehlman, and T.H. Murray, *Hastings Center Rep.* **27**(3), 14–21 (1997).
3. P. Whitehouse, J. Ballenger, and J. Gute, The Neuropsychology of Wisdom. I Virtual Congress of Neuropsychology in Internet (http://www.uninet.edu/union99). *Journal Revista de Neuropsichogia Espanola*. (in press).
4. P.D. Kramer, *Listening to Prozac*, Penguin, New York, 1993.
5. G. Klerman, *Hastings Center Rep.* **2**(4), 1–3 (1972).
6. J.E. Birren and L.M. Fisher, in R.J. Sternberg, ed., *WISDOM — Its Nature, Origins, and Development*, Cambridge University Press, New York, 1990, pp. 317–332.
7. P.J. Whitehouse, *Dementia*, FA Davis, Philadelphia, PA, 1993.
8. M.J. Knapp, *J. Am. Med. Assoc.* **271**, 985–991 (1994).
9. P.J. Whitehouse et al., *Science* **215**, 1237–1239 (1982).
10. T. Crook et al., *Developmental Neuropsychol.* **2**, 261–276 (1986).
11. R. Levy et al., *Int. Psychogeriatrics* **6**, 63–68 (1994).
12. L. Diller, *Hastings Center Rep.* **26**(2), 12–18 (1996).
13. E. Juengst, in T.H. Murray and M.J. Mehlman, (eds.), *The Encyclopedia of Ethical, Legal and Policy Issues in Biotechnology*, Wiley, New York, 2000.
14. T.H. Murray, *Drugs, Sports and Ethics in Feeling Good and Doing Better: Ethics and Nontherapeutic Drug Use*, Humana Press, Clifton, NJ, 1984.
15. Ad Hoc Committee on Growth Hormone Usage, *Pediatrics* **72**, 891–894 (1984).
16. E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998.
17. Y. Ku and P.J. Whitehouse, *Alzheimer Disease Assoc. Disorders* **10**(2), 61–62 (1996).
18. V. Potter, *Global Bioethics Building on the Leopold Legacy*, Michigan State University Press, East Lansing, 1998.
19. V. Potter and P. Whitehouse, *Scientist* **12**(1), 9 (1998).
20. D. Norman, *Just Health Care*, Cambridge University Press, New York, 1986.
21. J. Sabin and N. Daniels, *Hastings Center Rep.* **24**(6), 5–13 (1994).
22. C. Taylor, *The Ethics of Authenticity*, Harvard University Press, Cambridge, MA, 1991.
23. S.A. Alpert, in K.W. Goodman, ed., *Ethics, Computing, and Medicine informatics and the Transformation of Health Care*, Cambridge University Press, Cambridge, UK, 1998.
24. K.W. Goodman, in K.W. Goodman, ed., *Ethics, Computing, and Medicine informatics and the Transformation of Health Care*, Cambridge University Press, Cambridge, UK, 1998.
25. W.E.L Grimson, in Anonymous, 1999.
26. N.G. Leveson and C.S. Turner, *IEEE Comp.* **26**(7), 18–41 (1993).
27. J.W. Snapper, in K.W. Goodman, ed., *Ethics, Computing and Medicine informatics and the Transformation of Health Care*, Cambridge University Press, Cambridge, UK, 1998.
28. D.G. Johnson and J.M. Mulvey, *Comm. ACM* **38**(12), 58–64 (1995).
29. W. Gibson, *Neuromancer*, Ace Books, New York, 1984.
30. R. Zajtchuk and R.M. Satava, *Comm. ACM* **40**(9), 63–64 (1997).
31. R. Kurzweil, *When Computers Exceed Human Intelligence: The Age of Spiritual Machines*, Viking, New York, 1999.
32. D. Gelertner, *The Muse in the Machine: Computerizing the Poetry of Human Thought*, Free Press, New York, 1994.
33. R.W. Picard, *Affective Computing*, MIT Press, Cambridge, MA, 1997.
34. I. Asimov, *The Foundation Trilogy: Three Classics of Science Fiction*, Doubleday, Garden City, 1951.

See other entries BEHAVIORAL GENETICS, HUMAN; see also HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY entries.

HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY, ETHICS, HUMAN GROWTH HORMONE

CAROL A. TAUER
College of St. Catherine
St. Paul, Minnesota

OUTLINE

INTRODUCTION

The appropriate use of drugs that supplement endogenous production of growth hormone (GH) in children and adolescents is the subject of both ethical and medical controversy. From a scientific standpoint, questions about the long-term efficacy of GH treatment for children who are not GH-deficient (GHD) remain unresolved (1–8). Yet GH supplementation is proposed for a variety of indications, ranging from the child who lacks naturally produced GH to the average-height child whose parents want him to star in basketball. A regimen that is clearly therapeutic for some indications is proposed for questionably therapeutic or enhancement purposes in other situations. Can we draw a line between applications that are directed toward appropriate goals for the practice of medicine and those that fall outside its scope? This question raises conceptual issues because of the difficulty of distinguishing "therapy" from "enhancement," professional ethics questions related to defining the goals of medicine, and social ethics questions about what is a just or fair allocation of medical resources. Growth hormone therapy has become a paradigm case for testing theoretical analyses of issues like these.

THE THERAPY/ENHANCEMENT DISTINCTION

Debates on the ethical scope for genetic manipulation often cite the insertion or correction of the growth hormone gene as a test case. In discussions of therapeutic versus enhancement applications of genetic engineering, the manipulation of the GH gene in short to average healthy children is used as a typical example of enhancement engineering. W. French Anderson, pioneer in gene therapy research who advocates drawing a line to preclude enhancement applications, states:

> The most obvious [enhancement] example at the moment would be the insertion of a growth hormone gene into a normal child in the hope that this would make the child grow larger (9, p. 22).

While insertion of the GH gene itself is still a hypothetical possibility, the modification of height that is sought through biochemical GH raises similar questions. Insertion of the gene is a more drastic measure and most likely carries more risks, yet the administration of the drug also leads to questions about appropriate uses of medical technology for purposes that go beyond therapy, or for enhancement purposes. In fact John Robertson cites the presumed right of parents to increase a child's height through injections of growth hormone to support the claim that parents have a similar right to enhance a child's height through gene insertion at the time of conception (10).

Yet while conceptual and ethical debates about the goals of medicine and therapeutic versus enhancement applications of biotechnology continue to cite growth hormone as a test case (10–13), the literature of pediatric endocrinology has become increasingly skeptical of the potential success of expanded uses of GH. As studies of long-term height gains from GH treatment appear to show less promising results than were anticipated from short-term studies, researchers are showing decreased optimism and professional societies are advising great caution in the prescription of GH (2–5,15–19). If these unpromising results are confirmed through long-term controlled studies, then the ethical questions may become moot, at least for the specific case of enhancing height through GH. But the questions will remain, even if transferred to other technologies.

HISTORY OF GROWTH HORMONE

Treatment of children who are deficient in GH began in 1958, when hormone taken from the pituitaries of human cadavers was shown to significantly increase the growth of treated children. A child who does not secrete GH is described as a pituitary dwarf, and if untreated, is not likely to reach an adult height greater than 4 ft 6 in. However, a two-year course of treatment with natural GH required hormone from 50 to 100 pituitary glands, so dosage was limited and selection criteria for treatment were stringent.

In 1985, after some natural GH was found to be contaminated with an infectious agent or the prion that causes Creutzfeldt-Jakob disease, the FDA halted its sale. Two companies, Genentech and Eli Lilly, had developed biosynthetic versions of GH through recombinant DNA techniques. Genentech's drug Protropin was approved in 1985, Lilly's Humatrope soon after. FDA approval, however, was limited to the population of children who had been treated with cadaveric hormone, that is, children with classical GH deficiency or pituitary dwarfism.

The cost of treatment with biosynthetic GH is very high; it is currently estimated to average $18,000 a year, varying with weight of child, dosage, and frequency of injections (range estimated from $10,000 to $50,000 a year) (19,20). The usual course of treatment is four to five years, yielding an average cost per child of $80,000 to $100,000 (21,22).

From a situation of scarcity we have moved to a situation of highly plentiful (but very costly) growth hormone. Thus it has become possible to consider higher dosages, more frequent injections, and a more extended course of treatment, as well as extension to populations other than children with classical GH deficiency, such as girls with Turner's syndrome and children with chronic renal failure, or previous cranial irradiation or craniopharyngioma of the pituitary. In addition GH has been prescribed for children who are very short but who exhibit no definable medical condition or deficiency. While this usage is both ethically and medically controversial, if it were successful it could result in an even more extensive use of GH for healthy children of average height whose parents for some reason wanted them to be taller.

CONTROVERSY ABOUT APPROPRIATE USE OF GROWTH HORMONE

In the United States prescription of GH is primarily in the hands of pediatric endocrinologists, to whom referral is made when there is a concern about growth (20).

But endocrinologists disagree on appropriate criteria for prescribing GH beyond the population with demonstrated GH deficiency (GHD). Some reject the limitation to children with GHD, since scientifically the line between GHD and non-GHD children is by no means bright and clear, and may not even be meaningful (23).

Indications for Prescription of Growth Hormone

Clinical Criteria. It is very difficult to measure endogenous production of GH because secretion is variable or pulsatile rather than continuous. The usual method is provocative stimulatory testing through administration of substances like arginine, insulin, or Levodopa. An alternate approach is labeled physiological: sampling at frequent intervals over a 24-hour period or in a particular situation. There is no agreement on the best method; no measurement is viewed as completely reliable (24). Some authors describe three different diagnoses: complete or classic GHD, partial GHD or GH insufficiency, and non-GHD (25). Others hold that there is no discrete condition that can be labeled GHD but rather a spectrum of disorders involving GH secretion and utilization (23).

Given the pitfalls of trying to differentiate GHD from non-GHD children, some endocrinologists say the problem to be treated is short stature or low growth rate, not simply GH deficiency (26). The latter should not be the only, or even the primary, indication for GH treatment. In this view, the criterion for treating a very short child should be actual responsiveness to administration of GH, demonstrated through significantly increased growth velocity over a trial period.

For short but otherwise normal children, such use is presently experimental. An "off-label" use of a drug approved by the FDA for another population is not illegal. However, short-term acceleration of growth in normal children may not actually be a predictor of long-term gains. The drug might accelerate bone aging and result in a child's reaching the same adult height more quickly (or even a lesser height). Thus an enormous cost and the burdens of three to seven shots a week for four to five years might yield little or no long-term increase in height (2–5,15–17,27).

Current Practice Patterns. A recent national study of pediatric endocrinologists regarding their prescribing patterns showed a wide range of variation in the indications for which these doctors would recommend GH treatment (28). With a response rate of 81.3 percent (434 out of a possible 534), this study provides broad information about current practice. Physicians were given eight scenarios, all representing variations in the clinical status of a 10-year-old boy or girl (sex alternated randomly) whose current height was either 2 or 3 standard deviations below the mean. Other factors that were varied were growth rate, whether bone age was normal or delayed, and predicted adult height. Three price variations were added to each scenario (cost of $13,000/year, drop to $2000/year, drop to $100/year), as well as intensity of family desire or pressure, resulting in 32 different decision situations.

Respondents generally agreed that GH use for non-GHD children has been increasing over the last five years, and almost unanimously agreed that "short stature matters and has dysfunctional emotional impact." But there was lack of consensus regarding the perceived efficacy (both expected adult height and long-term adverse effects) of GH treatment for non-GHD children. Thus there was more agreement on the psychosocial ramifications of shortness than on the scientific and medical evidence, which is presumably the area of these professionals' expertise.

The investigators concluded that GH recommendations are currently based at least partly "on a desire to address perceived impairment rather than on a clear knowledge of patient response [i.e., medical evidence]" (28). External nonphysiological variables such as strength of family wishes and cost are apparently also significant in relation to prescribing decisions. A commentator on this study noted that until we have validation of GH treatment from controlled prospective trials, "the use of GH treatment in non-GHD children will continue to be based on anecdote and emotion rather than fact" (29).

Forces Driving Interest in Expanded Uses of GH

In the United States, thousands of short healthy children are currently receiving GH treatment which is unvalidated (30). A large number of them are part of a study that follows them at least through the period while they are receiving the drug. Genentech, whose drugs Protropin and Nutropin dominated the U.S. market in the early years of biosynthetic GH, provides for all recipients to be part of its postmarketing surveillance program, the National Cooperative Growth Study (31). While approximately two-thirds of the 12,000 children enrolled in the program by 1992 had GH deficiency or another definable medical condition, the remaining third is believed to represent primarily healthy short children (30,31).

The National Institutes of Health (NIH) are conducting several studies, mainly directed toward subject populations that have medical diagnoses, but one to test the long-term effectiveness of GH therapy in very short healthy children (27). (A study similar to the latter is ongoing in Great Britain, the Wessex Growth Study.) The NIH study calls for 80 "short stature" subjects to participate in a double-blind controlled clinical trial, with half the children receiving GH injections three times a week, the other half placebo injections. All children undergo extensive tests and examinations and will be followed to adult height, when the mean heights of the two groups will be compared.

Because of complaints that the NIH study violated ethical and regulatory requirements for research with children, an independent Review Committee was convened in late 1992. While this committee determined that continuation of the study was ethically acceptable, commentators pointed out that this conclusion was reached only by extending ethical norms on research involving sick children to research involving healthy children who were short (30,32,33). Thus the Review Committee's approval of the NIH study suggests that it is appropriate for medical science to seek remedies for

the condition of short stature, even in the absence of an identifiable medical condition.

Four types of forces are driving research programs on expanded uses of growth hormone: desire to resolve uncertainties in clinical medicine, economic interests of the drug companies, consumer demand and family autonomy, and ethical claims related to equality and justice.

Desire to Resolve Uncertainties in Clinical Medicine. The question of whether growth hormone prescribed to short normal children actually results in an increase in final adult height is considered one of the most pressing problems in pediatric endocrinology today. Thousands of short children with no diagnosable medical conditions are receiving GH, but their physicians do not know what the treatment is actually achieving, even at a purely physiological level. This widespread prescription of a nonvalidated (and costly) therapy mandates a research program of the highest rigor, according to many proponents including the NIH Review Committee (1,30).

In addition there is some concern about the safety of long-term use of GH in children who are not measurably deficient. While a history of administration to children with GHD has revealed no significant problems, we are now dealing with a different population (34,35). The dosages being administered are higher than when GH was scarce; and in order to sustain increases in growth rate in non-GHD children, it appears necessary to administer even larger dosages. Two organizations that filed court complaints opposing the NIH "short stature" study focused on what they regard as real risks of harm to healthy children (36,37). While the NIH Committee assessed these risks as hypothetical or insignificant (30), Arthur Levine of NIH defended the necessity for the research by stating that GH "could be dangerous," and its safety in short healthy children needed to be proved (38).

Many clinicians would argue that until the clinical uncertainties about benefits and risks have been resolved, it is premature to worry about what they view as theoretical ethical questions about the goals of medicine and enhancement therapies.

Economic Interests of the Drug Companies. The rationale for the NIH "short stature" protocol states that the prospect of a plentiful supply of GH leads to "the need to evaluate other potential uses for this hormone" (27). Genentech had an "orphan drug" permit for GH which expired in 1992, and Eli Lilly's expired in February 1994. After these two companies lost their protected market in the United States, three other companies, Novo Nordisk, Pharmacia and Upjohn, and Serono developed biosynthetic GH drugs that have been approved for use in the United States. All five companies manufacture drugs approved for treatment of pediatric GHD, and some are also approved for other indications (39). In addition several European manufacturers have shown interest in obtaining FDA approval to market GH in the United States.

In the early 1990s market analysts suggested that extensive competition could drive the price of GH well below the cost of the protected Genentech and Eli Lilly drugs (40). It was even suggested that the cost of human GH might eventually come near that of bovine GH, which is produced by a similar recombinant-DNA process. When the congressional ban on use of bovine GH expired in February 1994, the cost of a two-week injection was $5, or $130 a year (41). However, the price of human GH has remained relatively stable despite new manufacturers and loss of orphan drug status (20).

The threat of loss of market share and a potentially lower cash return on investment would necessarily influence a company to explore additional uses for a product it has developed at great expense. In the case of recombinant-GH, potential uses in adults have begun to be explored and are the focus of numerous research efforts. Studies indicate that adults who are GHD frequently exhibit increased fat mass, reduced muscle mass and strength, smaller hearts and lower cardiac output, lower bone density, and psychological problems, and they appear to have an increased risk of death from cardiovascular disease (42). While there are some risks of side effects, GH replacement therapy for GHD adults has been demonstrated to be beneficial through double-blind, placebo-controlled trials (42,43). This therapy has been approved by the FDA for use with GHD adults as well as to treat AIDS wasting. Estimates indicate that about 70,000 adults in the United States may be affected by GHD, while the current pediatric market includes approximately 40,000 patients (39). Thus expanding the prescription of GH to GHD adults offers the prospect for a highly lucrative market for drug companies.

Nevertheless, there will likely continue to be interest in exploring additional uses for GH with children and adolescents and in extending FDA approval beyond the current indications of GHD, Turner's syndrome, and chronic renal insufficiency (39). Since short stature and abnormally slow growth of a child are observable characteristics and are generally of concern to parents, consumer interest in remedies for short stature will continue to drive research on the use of GH for non-GHD short children.

Consumer Demand and Autonomy. As the public becomes aware of the availability of growth hormone treatment, parents of a short child increasingly take initiative in requesting that it be tried. These parents may be concerned about psychological and social difficulties that may confront a short child in later childhood and adolescence, or about the long-term economic and social disadvantages of being short. Such problems are perceived to be more serious for males than for females; approximately 75 percent of the children in the Genentech collaborative study who do not have specific medical diagnoses are male (31), as are 90 percent of the children enrolled in the NIH "short stature" study (30). But even when classical GHD is included, the percentage of black children treated is only a third of what would be expected given the percentage of black children in the U.S. population, and females with problems are identified only when their deficiency is more extreme than that of males. Moreover, as awareness of the availability of GH treatment increases, data show "an even greater tendency to refer or test males [than] females" (31).

Parental request for testing and for a trial of GH treatment is largely based on a subjective perception of inadequacy, to which the medical system responds. The result is a documented referral bias in favor of testing and treating white male children, even though females, and males of other races, may be able to demonstrate greater need. Moreover there is some evidence that treatment of non-GHD females may result in more significant long-term height gains than treatment of non-GHD males (6,8,16). Given that the risks of therapy appear to be slight and that a short-term growth benefit is likely, parental pressure for GH treatment of a short healthy child, particularly a short boy, may be difficult to resist. Note that the study of prescribing patterns described earlier found that strong family wishes for GH increased the likelihood that an endocrinologist would prescribe it (28).

In this era when patient or parental autonomy regarding medical treatment is a central or even overriding ethical value, the endocrinologist may consider the option of GH therapy a matter of parental choice. Provided that parents are well-informed that for short healthy children the treatment is still experimental with a remote possibility of risk, then their autonomy permits them to choose it. However, thoughtful physicians will also raise issues of possible psychological harm: What if expectations are not reached? Will the child be severely disappointed if the gain is only one or two inches? Does the use of drug treatment in itself suggest to the child that there is something wrong with him or her? (44) After the full range of risks and benefits has been explored, the principle of autonomy may appear to support the prerogative of parents to make the choice of treatment in the interests of this particular child, whom they presumably know better than anyone else does.

Justice: Two Aspects. As the previous section notes, the current allocation or distribution of GH treatment raises questions about fairness or justice in relation to race and gender. Prescribing patterns suggest that this very costly drug is currently prescribed in a way that is discriminatory. Even if the discrimination is purely de facto, resulting from which particular parents and children pursue GH treatment most energetically, still the discrimination and its effects remain.

But a different interpretation of justice has been invoked to support a principled extension of GH treatment and to justify providing it to short healthy children who are not GHD. David Allen and Norman Fost have devised the cases of Johnny and Billy to persuade us that a short non-GHD child has as much right to GH treatment as a GHD child:

> Johnny is a short 11-year-old-boy with documented GH deficiency. . . . His predicted adult height without GH treatment is approximately 160 cm (5 ft 3 in.). Billy is a short 11-year-old boy with normal secretion according to current testing methods. . . . He has a predicted adult height of 160 cm (5 ft 3 in.) (26, p. 18).

Allen and Fost argue that it is unjust and discriminatory to provide treatment to Johnny but not to Billy, solely on grounds that Johnny has an identifiable medical deficiency and Billy does not. If the two boys can be expected to experience equivalent psychosocial problems and to be similarly disadvantaged, both as children and adults, then it seems arbitrary to treat one with GH but not the other.

The NIH Review Committee gave a great deal of weight to the principle of equal treatment, arguing that short non-GHD children suffer the same "functional impairment and psychosocial stigmatization" as GHD children. Therefore it could be unjust to deny them access to treatment simply on the basis of an "imprecise definition of 'deficiency'" (30). On this interpretation of justice it could be discriminatory to deny GH treatment to a child, no matter what his or her medical condition, if the child's short stature is perceived to be disadvantaging, disabling, or otherwise problematic. Authors who have written in defense of the treatment/enhancement distinction acknowledge that this example presents a hard case for them (13).

GOALS OF GROWTH HORMONE THERAPY

Importance of Identifying the Goal of Therapy

In the biomedical ethics literature, attempts to discern a conceptual and ethical distinction between therapeutic and enhancement uses of biotechnology often cite the case of Johnny and Billy, sometimes with discomfort that Billy appears to be as entitled to treatment as Johnny. In the literature of pediatric medicine, the current focus is on formulating a standard of care that has a rational connection with research results (18,19). Here the aim is to develop criteria for the classes of children for whom GH treatment is appropriate, based on studies that demonstrate significant benefits to that class in proportion to risks and costs, both personal and financial.

For either enterprise, the goals for GH treatment must be explicitly identified. Possible goals cover a variety of statistical, therapeutic, pychosocial, and enhancement ends or purposes. Each of these goals presumes a different understanding of the benefit to be provided, and thus each one points to a different way of evaluating whether treatment has been successful. Some of the goals appear to be only instrumental ends, where success requires that their attainment be a means to the attainment of some other more ultimate goal. Thus a rather lengthy list of goals collapses into a shorter set of ultimate goals.

Goal 1. To Increase Growth Velocity or Growth Rate. Some clinicians emphasize potential benefits of an increased growth rate, whether or not GH treatment produces a significant augmentation of eventual adult height (45). A child with measurable GH deficiency (GHD) has a pathology of growth and a physiologically abnormal growth rate. Similarly, it may be argued that a child without measurable GHD but whose growth velocity is measurably abnormal also has a pathology of growth or a growth disorder (46,47). If such a pathology is regarded as a medical condition, then its remediation is therapeutic and falls under traditional goals of medicine.

In their clinical discussion of "disorders of stature," Hindmarsh and Brook advocate using abnormality of growth velocity as the main criterion for considering

GH treatment. It is clinically diagnosable by objective standards. It circumvents the relationship of height to societal biases and is independent of racial or ethnic height differences (48). It is also independent of male–female height differences. This view maintains that amelioration of abnormal growth rate is the correction of a physiological deficiency.

But providing GH treatment to increase growth velocity in a short child may also be regarded as a means to some other goal, for example, to prevent or remedy psychosocial or behavioral problems that short children are thought to experience. However, if increased growth velocity is perceived as instrumental to the achievement of other goals, then its success and legitimacy must be evaluated in relation to those other goals (49). Only if studies demonstrate that psychosocial gains are actually achieved could an increase in growth rate directed toward such goals be defended.

Goal 2. To Increase Eventual Adult Height. The NIH "short stature" study, the Wessex Growth Study, and a number of already-completed studies aim to determine whether short-term gains from GH treatment translate into long-term height gains (2–8,15). For many physicians, parents, and researchers, the increase in growth rate sought in goal 1 is only a means to an increased adult height.

However, while medical science can assess normal growth (or growth velocity), there is really no medical criterion for normal height per se. Height is partly genetic, and is sex, race and ethnicity dependent (as growth rate is not). While it is within the competence of medicine to investigate why a child is not growing and to ameliorate that deficiency, medical science lacks objective criteria for what is a "good" or "healthy" height (13,50,51).

Height is not the sort of thing where one necessarily desires more rather than less, and some people (given goals such as gymnastics or riding race horses) may prefer to be statistically quite short. An increase in adult height through GH treatment, if achievable, is a goal that is necessarily only a means to the attainment of other desired goals. Increased height may be sought in order to avoid perceived psychosocial problems, or because extreme shortness could be functionally disabling, or because shortness is economically and socially disadvantaging within certain contexts. Thus the success of treatment must again be evaluated in terms of whether it achieves its ultimate goals.

Goal 3. To Ameliorate a Perceived Deformity. The NIH study protocol states that extreme short stature is perceived as "a major developmental abnormality" and the Review Committee cites evidence that it results in "psychosocial stigmatization." In relating this problem to traditional medical practice, Lantos, Siegler, and Cuttler note that surgery for deforming congenital anomalies is considered proper and is generally reimbursable. They believe that the short stature resulting from classical GHD is so severe that it is generally viewed as a major deformity, and hence is treatable (21). In most cases, however, the shortness of children who are not GHD is not as extreme.

Goal 4. To Prevent or Treat Psychosocial, Learning, and Behavioral Problems. Studies have suggested that children with very short stature may experience failure in school and have psychosocial and behavioral problems (52,53). For some endocrinologists, preventing or ameliorating these problems is the major goal of GH therapy. Underwood and Rieser recommend that therapeutic goals be defined "in terms of how short-term acceleration of growth would benefit the patient socially and psychologically" (45). Increasing the rate of growth could be instrumental for this short-term purpose even if significant gains in adult height could not be demonstrated.

If the actual or potential psychosocial problems of the patient are definable under the umbrella of mental health treatment (see Ref. 54 for difficulties in such definition), then GH treatment might perhaps be viewed as a therapeutic or preventive means for dealing with them (49). There are precedents in the practice of medicine for the use of drug therapy for behavioral or psychological disorders in children. However, it would be highly unusual to prescribe drug injections to children merely because they were believed to be at risk for such problems, namely for prevention (21). But what about very short children who are actually displaying these problems?

Recent studies show that psychosocial problems are not correlated to shortness per se, but that they are more likely to be related to hormonal deficiencies (2,55,57). Thus in assessing the appropriateness of GH therapy as a means for treating mental health problems, the two populations of GHD and non-GHD children must be considered separately. However, even with GHD children, no one yet is able to claim that we have scientific evidence for the effectiveness of GH therapy in improving their psychosocial functioning (52,58). "There have been no placebo-controlled evaluations of the behavioral effects of GH treatment," according to Richard Clopper (58). Some clinicians decide to prescribe GH largely on the basis of a perception that short children experience psychosocial problems, although there is no solid evidence from research to support GH as a remedy (28,29). If the reason for GH therapy is its supposedly beneficial effect on psychosocial adjustment and school performance, then studies must measure whether these results are actually achieved.

Goal 5. To Correct a Functional Disability. Advocates of GH treatment for short stature stress its functionally handicapping effects, particularly if the shortness is extreme. The NIH Review Committee cited "functional impairment" and "difficulty with physical aspects of the culture" as problems of short people that justified a research program (30). In providing examples to show that very short people are functionally handicapped, authors frequently cite problems in driving a car, and inability to reach shelves, light switches, and elevator buttons. Since children who are still growing are not regarded as handicapped simply because they are not tall enough to do these things, attaining this goal requires that GH treatment achieve an increase in final adult height.

Medical treatment is typically provided when it is effective in remedying a functional disability that

interferes with activities of daily living or that prevents one from earning a living or living independently. The question is: what level of short adult stature can realistically be defined as that sort of disability? Often it would be simpler (and less costly) to modify the environment or the vehicle than to attempt to modify the body of a completely healthy person by means of GH injections. Moreover the criteria currently used to define "extreme shortness" are statistical and have no logical relationship to standard definitions of functional disability. Thus, while the goal of remedying a disability is consistent with therapeutic goals for the practice of medicine, its application to GH treatment of non-GHD children is questionable.

Goal 6. To Remove the Economic and Social Disadvantage of Short Stature. Data show that our society has a bias against short people, particularly short males. Allen and Fost note that "discrimination based on height — heightism — pervades American life," and that shortness "imposes a disadvantage in the competition for schools, jobs, income and mates" (26).

Society's attitude toward short people is a prejudice that often leads to discriminatory consequences. Lantos et al. note that "we do not usually call prejudice-induced conditions . . . diseases" (21). In his analysis of prevention as a legitimate goal of medicine, Eric Juengst argues that this goal should be limited to efforts to defend people from "robust pathological entities [i.e., genuine disease entities], rather than changing their bodies to evade social injustices" (51). In other words, it is not appropriate for medicine to respond to discrimination by modifying the physical characteristics against which society is prejudiced. Other authors describe the use of medicine to respond to society's preference for taller people as social engineering. In attempting to make some people taller, medicine may simply be reinforcing "heightism" in our society.

Goal 7. To Enable Parents to Seek a Preferred Height for Their Children. This goal suggests that any height could be achievable, which is almost certainly incorrect. Ann Johanson, former Director of Clinical Affairs for Genentech, acknowledges that children with normal stature and growth rate are unlikely to gain significant additional height or growth velocity from GH treatment (59). Recently completed studies confirm this hypothesis (2–7,15,60). However, if a safe method of supplementing GH, either biochemically or through genetic manipulation, were shown to cause significant height increases in normal-height children, this goal would represent a pure enhancement use of GH. In the literature specifically focused on growth hormone, no one appears to advocate such enhancement use, even though parents may request it. But discussions of enhancement therapies in general sometimes defend providing them, based on the constitutional right of parents to autonomy and discretion in rearing their children (10).

GH TREATMENT FOR CHILDREN WITHOUT MEDICAL CONDITIONS

The Argument Invoking Equal Treatment

The Case of Johnny and Billy. The only clearly therapeutic goal for GH treatment is to increase or normalize growth rate in children who have a definable medical condition that is related to a pathology of growth. Secondarily this increased growth rate ought to produce an increment in final adult height over what was predicted. But what of Allen and Fost's case of Johnny and Billy? Recall that Johnny and Billy are both short 11-year-old boys, and that each has a predicted adult height of 160 cm (5 ft 3 in.). The difference between them is that Johnny has a documented GH deficiency, while Billy's tests show normal GH secretion. Allen and Fost argue that it is discriminatory to offer GH treatment to Johnny but not to Billy (26).

In applying the concept of justice to this situation, we are invoking the principle that equal cases should be treated equally. Since no two cases are ever exactly alike or equal, we must identify the relevant factors in which similarity is required. With Johnny and Billy, the similarities stated are short stature at present and equal predicted adult heights. Similarities assumed are that shortness will cause them similar psychosocial problems and will disadvantage them equally, and that GH treatment will benefit them equally and carry equivalent risks. If these assumptions are correct, then fairness seems to mandate that the two cases be treated equally in terms of provision of GH. In brief, a non-GHD child predicted to reach the same adult height as a GHD child has an equal right to GH treatment.

Justice entails this conclusion, however, only if the two cases are truly similar in relevant respects. But actual data show that they are different in important and often overlooked ways, so that the principle of equal treatment of similar cases does not apply to these two situations.

Dissimilarities between Johnny and Billy. First, a lack of endogenous GH appears to be linked to physiological and functional consequences in addition in growth. While earlier studies of short-statured children had shown them more susceptible to learning difficulties and academic failure than children of normal height, these studies did not distinguish according to the etiology of the short stature. Newer studies which concentrate on short children with GHD suggest that these children "show significant deficits in several specialized cognitive domains including those requiring complex visuoconstructional skills, orientation in space, long-term memory, and attention span" (61).

Some of this research fails to distinguish children with isolated GHD from children who lack not only growth hormone but are multiply hormone deficient or even panhypopituitary (lacking all pituitary hormones). When these distinctions are made, children with more severe endocrine problems do demonstrate more severe cognitive deficits and psychological disturbances (61,62). But authors of the Michigan longitudinal study, which followed a group of GHD children for seven years, reached a preliminary conclusion that even some children with

isolated GHD "may have cognitive profiles similar to those described as learning disabled" (61,63).

Two studies of the psychological adjustment of GHD adults show that their profiles differ from those of matched controls of equally short stature, and that they are at risk for anxiety and depressive disorders, frequently displaying symptoms seen in patients with clinically diagnosed social phobia (64,65). Again, subjects with multiple endocrine deficiencies appear more vulnerable than those with isolated GHD. But Pine, Cohen, and Brook found that "A blunted growth hormone response to physiologic challenges remains perhaps the best-replicated biological correlate of emotional disorder," while not claiming which is cause and which is effect (66).

In summarizing the results of these studies, Brian Stabler observes that short stature alone is not responsible for the low quality of life of many GHD patients, but that "neuropsychologic functioning is fundamentally impaired in many GHD children." The more deficient in GH a child is, the greater the cognitive and behavioral difficulties, suggesting a "relationship between psychosocial functioning and degree of endocrine deficiency" (62).

On the other hand, studies of the adjustment of short children in general show that their level of social and academic functioning is "reasonably indistinguishable from that of average-statured peers" (55–57). After reporting their study that showed minimal long-term height gains from GH treatment of short normal children, Hindmarsh and Brook concluded:

> It has been alleged that short stature adversely affects children and the short-term effects of r-hGH on height ... might therefore find advocates for this therapy. Neither we nor others have been able to document markers of adverse psychological effects in normal short children (2, p. 16).

Thus short stature is not the primary factor underlying the behavioral, academic, or emotional difficulties of some short individuals. Rather, these problems appear linked to an underlying medical condition of which short stature is only one feature (55). While it may be reasonable to provide GH to Johnny to help him overcome these types of deficits in addition to growth failure, it is unlikely that merely increasing Billy's stature will offer comparable psychosocial benefits to him.

Second, there is skepticism as to whether a non-GHD child could make comparable height gains on comparable doses of GH. While some authors believe that cumulative growth response of non-GHD subjects is comparable to that of GHD patients (67), others cite strongly conflicting data (68). The latter argue that in order to achieve comparable height gains you would have to treat over a longer time period, at higher dosages, with correspondingly higher costs. At some higher dosage toxicity could be expected, possibly leading to side effects such as diabetes.

Recent studies that compare two groups of non-GHD children, one group treated with GH and the other untreated, also come to differing conclusions (2,3,6,7,17). Hindmarsh and Brook found that both treated and observation groups had some increases in final height over that predicted, but that the change in the treated group was not significantly greater than that in the observation group (2).

While studies show that GHD children like Johnny can be expected to make significant height gains through GH treatment, there is no comparable consensus among researchers as to whether Billy could anticipate similar gains, especially in the long term. It is unlikely that an expenditure on Billy equivalent to that on Johnny would bring Billy equal benefits, even in relation to height itself.

Third, the case description makes no mention of the two boys' growth velocities. It is almost certain that Johnny has subnormal growth velocity, while Billy's may well be within the normal range. If so, then the two boys differ significantly with respect to pathological versus normal growth, and Billy would be regarded as developmentally normal.

Fourth, because Billy does not have an identifiable medical or developmental problem, he is apt to be at greater psychological risk as a result of prescription of drug therapy. In his article "Is Taller Really Better?" Douglas Diekema argues that treatment itself may have a stigmatizing effect:

> When we seek to change the height (or physical appearance) of a child, he may perceive that he is incomplete and unacceptable. His peers may have suggested this on the playground. Now his parents seem to have confirmed it through their efforts to make him taller (69, p. 114).

Cosmetic interventions after an illness or accident are "less likely to make the child perceive himself as undesirable," since they restore something that was taken away. Interventions which repair abnormalities that interfere with functional capacity offer the child a benefit for him or herself, not merely in relation to how the child is perceived by others or in comparison with others. But GH treatment aimed solely at changing a child's body or appearance could have adverse psychological effects and do more harm than good (69).

C.G.D. Brook of the London Centre for Paediatric Endocrinology warns doctors who are pressured into providing GH for non-GHD children to be mindful of "the problems of stigmatising otherwise normal children" (60). Billy is at more risk of this stigmatization through the medicalization of his stature than is Johnny, who has a definable medical problem.

Fifth, while Johnny and Billy as individuals may both experience psychosocial problems, Billy's are more apt to be treatable without use of GH. Stabler notes that short individuals who do not have neuroendocrine impairment are very adaptable, and he cannot say whether GH or psychotherapy is more effective in treating them (70). Since Stabler advises that GH should never be given without supplementary counseling or psychotherapy, it would be significantly more cost effective to choose psychotherapy rather than GH plus psychotherapy for Billy. Brook, in noting the lack of evidence that short children suffer psychosocial disadvantage solely because of their stature, as well as the impossibility of testing whether an increase in growth rate would provide psychosocial benefits, concludes:

> It is much more important for a short child to acquire coping skills than to buy [socially unimportant] inches through pharmacological means (60, p. 692).

These five arguments show that in the real world of GH treatment, the given data about Johnny and Billy are entirely consistent with the conclusion that they differ in a number of ways that are relevant with respect to treatment with GH. The fact that GH is provided for GHD children does not entail that it must be provided for equally short non-GHD children.

Short Stature as Disability or Disadvantage

Even if Johnny and Billy are not completely equivalent with regard to the provision of GH therapy, still it may be argued that because of the handicapping effects of very short stature, or the social and economic disadvantages of shortness, it follows that increased height (and hence GH treatment) is a legitimate medical goal for Billy. Even advocates of a restrictive definition of legitimate goals for the practice of medicine generally support treatment to remedy or ameliorate a disability. For example, Sabin and Daniels hold that "the central purpose of health care is to maintain, restore, or compensate for the restricted opportunity and loss of function caused by disease and disability" (54). The application of medical science in order to level social and economic disparities is more controverted. But even the goal of treating "disability" has problems when invoked to justify the extension of GH treatment.

Is Short Stature a Handicapping Condition? Allen and Fost classify extreme short stature as a handicap, that is, "a physical . . . disability that prevents or restricts normal achievement." They suggest that height below the 1st percentile is "likely to be handicapping" (26). (For North American adult males this would be a height about 5 ft 3 in.; for adult females about 4 ft 10.5 in.)

As noted earlier, difficulty in driving a car or in reaching shelves or switches usually can be ameliorated through modification of a vehicle or environment. While it is possible to imagine statures so short that a person is unable to function outside a radically adapted environment, the heights suggested do not appear to be that restricting.

Moreover the appeal to short stature as a functional disability is belied by height criteria that are significantly different for males and females. Since the mean height for adult males is about 4.5 in. greater than that for females, treatment criteria based on these mean heights cannot really be aimed at overcoming a functional handicap. Males and females do not drive different cars according to their sex, nor use different light switches. If a 5 ft 1 in. female is not functionally impaired in activities of daily living, why is a 5 ft 1 in. male?

Much publicity has been given to fatalities allegedly caused by air bags at the time of an automobile crash. More than half of such deaths have involved children, and almost all adult deaths have been women. Evidence indicates that "women of short stature are particularly susceptible to head and neck injuries from air bags" (71,72). Women under 5 ft 2 in. are statistically at particular risk; however, that height, while below the first percentile for males, is well above it for females. Moreover federal crash-worthiness standards call for testing that uses 5 ft 8 in. dummies, the height and weight of an average male. In September 1998 the National Highway Transportation Safety Administration proposed new rules that would require testing to minimize the risks of air bags to infants, children, and short adults. However, comment on these rules was extended to December 30, 1999, and as of early 2000, the rules had not yet been finalized (73). The reason that the driving environment is more hazardous to short people is that safety tests and modifications are premised on drivers' being average-sized males. This example illustrates how the "handicap" of shortness results from the way equipment is tested and designed, not from shortness per se.

Short people must live in the world as it exists at present, however. A careful study of daily life activities and a range of occupational choices might identify a height level below which a person really is handicapped, that is, unable to function without major environmental accommodations. (This height level might turn out to be around 4 ft 6 in.) Such a result would base the criteria for "handicapped by reason of stature" on actual data, rather than on statistical population norms. It might then be possible to consider offering GH treatment to remedy the disability of short stature.

In determining the allocation of treatment, however, the resources that are generally available for the treatment and rehabilitation of people with disabilities must also be taken into account. Fairness requires that equally debilitating handicaps be provided equal resources. Economist Mary Ann Baily suggests that money spent on expanding access to GH could better be used instead to help severely handicapped children through rehabilitative care, training, and high-tech devices, since at present these benefits are "not well-covered by private or public payers" (22).

If we apply Allen and Fost's suggestion that children in the first percentile be considered handicapped, then with about 39,000 U.S. children in the first percentile at any given age, to provide all of them with five years of GH treatment at $18,000 a year would cost over $3.5 billion a year. Compared with the defense budget this amount may seem small. However, total 1996 Medicaid payments for dependent children under 21 were only $17.5 billion, and the 1999 allocation to NIH for medical research for infants and children through the National Institute of Child Health and Development was only $752 million (74,75). Many children in the first percentile may have pathologies of growth or other medical conditions. But for those who do not, the use of scarce resources to provide GH treatment appears to be a low priority in relation to other possible expenditures related to the needs of disabled and sick children.

Should Medicine Remedy the Disadvantage of Short Stature?. Short stature has been shown to carry both economic and social disadvantages. One study suggests that there may be an increased income (height bonus) of

roughly $1150 a year (1999 dollars) per inch of greater height (76). Over a work life of 40 years, the total increment would be $46,000 per inch. If GH treatment were evaluated purely on an economic basis, a gain of two inches or $92,000 through a five-year course of treatment costing $90,000 would just about balance the investment in GH treatment. However, the anticipated increment in income would be gained through an increase in height relative to others who would become comparably shorter and presumably earn less. Thus an individual's interest in greater height is not correlated to a societal interest when medical resources are expended for GH treatment.

Most discussions of the disadvantaging effects of short stature focus on males. Studies show that a 12-inch height reduction is a significant predictor of lower economic status among men but not among women (77,78). (Obesity is a significant predictor of lower economic status for women but not for men.) These data could be taken to mean that women suffer less economically from shortness, and hence the disproportionate treatment of males with GH is not a concern. On the other hand, there is evidence that the typical size difference between males and females plays a major role in the status and income differential between the sexes. One study found that women who are 5 ft 7 in. earn on average the same salary as men who are 5 ft 7 in. (79). A study of the effects of both sex and size on status ranking found, as expected, that both males and people of greater size are perceived as having higher status. More surprisingly, this study found that the correlation between sex and status rank is largely (65 percent) accounted for by the correlation of sex with size, rather than by socioculturally influenced factors (80).

When the practice of medicine engages in trying to make some people (mostly white males) relatively taller, it is not only contributing to the "heightism" of our society. Given the correlations among sex, size, and status, it is also reinforcing discriminatory attitudes in relation to women (and toward people of typically shorter racial and ethnic groups) (79).

Moreover we do not know whether persons whose height is increased through GH treatment acquire the same advantages as people who are taller to begin with (69). Demographic data on adults who were GHD and were treated with GH are limited and often disappointing. Subjects in some studies, while reaching educational levels comparable to the population as a whole, were not comparably employed, married, or living independently of parents (62,81,82). Other studies, however, have identified more positive outcomes (83,84). A well-controlled study at Children's Hospital of Buffalo found a high educational level in previously treated adults, and also a low unemployment rate of 8 percent, at a time when overall unemployment in the area was 9 to 10 percent. None of these adults were experiencing significant emotional or adjustment problems (84).

Thus the data on the effectiveness of GH treatment for overcoming economic and social disadvantages are mixed, and placebo-controlled studies are completely unavailable (58,85). Even for GHD individuals, studies do not provide convincing evidence that GH treatment results in alleviation of the economic and social disadvantages of short stature.

These arguments provide three reasons for questioning the use of GH treatment to overcome economic and social disadvantages: (*1*) We cannot show that it is effective for this purpose. (*2*) Even if it were, providing some persons a height advantage relative to others offers no net benefit to society. (*3*) Not only is there no societal benefit, there is a negative impact through the reinforcement of societal biases and prejudices.

Liberty Rights and Personal Preferences

It could be argued that if an individual has a particular personal goal for which greater height would be an asset, then that individual should have the freedom to seek professional assistance to achieve that increased height. Since physicians, and specifically pediatric endocrinologists, control the prescription of biochemical growth hormone, physicians would necessarily be involved, and their prescription of GH to satisfy personal preferences would be no different from other uses of medical expertise for cosmetic purposes. Arguments about a just or fair allocation of societal resources would carry less weight if the patient either paid for, or contracted with an insurer who would pay for, such treatment.

This argument overlooks the fact that GH treatment to increase height must be provided during the child's growing years, and is most effective when begun early and continued until growth is completed. Hence (almost) all treatment occurs while the patient is a minor and unable to give an autonomous consent. Given the uncertainties about the clinical outcomes and the psychosocial benefits of GH treatment for non-GHD children, as well as the unpromising results of long-term studies, even adults would find it difficult to weigh these supposed benefits against the risks, inconveniences, and costs of years of medical treatment.

The endocrinologist has been designated as the customary gatekeeper for prescription of GH because of the need for a well-informed person to weigh risks against benefits, and to safeguard against the possibility of nonbeneficial (or even harmful) usage (19). Pediatricians, of whom pediatric endocrinologists are a subset, are given authority regarding child health issues because they are regarded as "guardians of and spokespersons for the well-being of children" (86). This charge requires that at minimum they base prescription of GH on evidence from studies and on practice guidelines developed by professional societies (18,19,29). It also requires them to recognize that while adults may choose to utilize medicine to change their bodies according to their preferences, adults (even parents) do not have a right to impose arduous treatment on healthy children simply to change their appearance.

The use of GH to treat short stature per se is a "medicalization" of problems whose sources lie within society rather than the short individual. The search for drug treatment to ameliorate psychosocial problems experienced by the individual patient often overlooks a wide range of other therapies, resources, and supports that are apt to be more effective in dealing with these problems. In fact some practitioners argue that focus on

short stature "is in some ways an obstacle to gaining truly comprehensive care for [short] children" (62).

Moreover treatment that is provided by physicians, even if paid for by individual parents, increases the pressure to expand reimbursement through public funds and through health plans. What is at first regarded as a consumer issue, namely providing what clients want and will pay for, eventually comes to be expressed in terms of fairness: Why should the wealthy have access to a medical treatment that most people cannot afford? History shows that entitlements are often expanded in this way, without other justification.

The use of medical technologies to modify people according to their preferences is apt to increase public concern about emerging genetic therapies. While most genetics researchers and clinicians stress their interest in applying gene therapy only to severe and lethal diseases (9), yet the public can easily be aroused to fear and oppose all interventions that involve genetics. Avoiding the use of GH for what can only be regarded as enhancement purposes upholds a consistent policy of focusing genetic science on therapeutic applications.

Finally, no patient or parent has a constitutional or liberty right to demand any particular medical treatment. Even the right to procreative liberty, a highly protected right in the United States, does not permit one to demand RU-486 as an abortion option, nor to insist on Depo-Provera as a contraceptive before it was approved for that purpose. While parents have a wide range of choices as to how they wish to rear their children, they cannot insist that the medical profession cooperate with whatever plan they may have.

CONCLUSION

The development of recombinant human growth hormone has resulted in a plentiful supply of the GH drug. It has been shown to be safe and effective in improving growth in children who are growth hormone deficient, and is also approved for treatment of girls with Turner's syndrome and children with chronic renal insufficiency. In addition administration of GH has been recognized as therapeutic for GHD adults and is effective to treat AIDS wasting. The question of whether the drug ought to be prescribed beyond these categories, particularly for children who are healthy but short, remains an unresolved question. For many endocrinologists, abnormally slow growth, but not simply shortness, is an indication for treatment, even in the absence of other medical conditions.

The prescription of GH to short but otherwise normal children represents the use of medicine for enhancement purposes. Beyond the as-yet-unanswered question of whether such treatment actually achieves a significant long-term gain in height, there are unresolved ethical questions regarding the appropriateness of directing medical and societal resources to the amelioration of short stature. If amelioration of shortness could be shown to achieve other appropriate goals, such as improvement in psychosocial functioning, there might be justification for the prescription of GH to short normal children. However, studies do not demonstrate that shortness in itself is problematic, and other solutions for psychosocial and similar problems appear to be preferable, especially given the high cost and the burdensome administration of GH treatment.

Because of a variety of factors associated with GH treatment for idiopathic or unexplained short stature, many arguments can be marshaled against allocating medical resources to this form of enhancement. However, these arguments are not necessarily transferable to other biotechnologies in which the enhancement versus therapy debate arises. Exploring these biotechnologies on a case-by-case basis provides one means for clarifying the enhancement or therapy distinction and its ethical implications, or for supporting the position that the distinction is not a helpful one.

BIBLIOGRAPHY

1. M.D.C. Donaldson, *Lancet* **348**, 3–4 (1996).
2. P.C. Hindmarsh and C.G.D. Brook, *Lancet* **348**, 13–16 (1996).
3. S. Loche et al., *J. Pediatr.* **125**, 196–200 (1994).
4. J.M. Wit et al., *Clin. Endocrinol.* **42**, 365–372 (1995).
5. L.T.M. Rekers-Mombarg et al., *J. Pediatr.* **132**, 455–460 (1998).
6. J.G. Buchlis et al., *J. Clin. Endocrinol. Metab.* **83**, 1075–1079 (1998).
7. R.L. Hintz et al., *N. Engl. J. Med.* **340**, 502–507 (1999).
8. E.S. McCaughey et al., *Lancet* **351**, 940–944 (1998).
9. W.F. Anderson, *Hastings Center Rep.* **20**(1), 21–24 (1990).
10. J. Robertson, in J. Robertson, ed., *Children of Choice*, Princeton University Press, Princeton, NJ, 1994, pp. 149–172.
11. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997, pp. 110–113.
12. N. Daniels, *GROWTH. Genet. Hormones* **8**(suppl. 1), 46–48 (1992).
13. E. Parens, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998, pp. 1–28.
14. D.W. Brock, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998, pp. 48–69.
15. J. Coste et al., *Br. Med. J.* **315**, 708–713 (1997).
16. M. Kawai et al., *J. Pediatr.* **130**, 205–209 (1997).
17. S.E. Oberfield, *N. Engl. J. Med.* **340**, 557–559 (1999).
18. Drug and Therapeutics Committee, Lawson Wilkins Pediatric Endocrine Society, *J. Pediatr.* **127**, 857–867 (1995).
19. Committee on Drugs and Committee on Bioethics, American Academy of Pediatrics, *Pediatrics* **99**, 122–129 (1997).
20. B.S. Finkelstein et al., *J. Am. Med. Assoc.* **279**, 663–668 (1998).
21. J. Lantos, M. Siegler, and L. Cuttler, *J. Am. Med. Assoc.* **261**, 1020–1024 (1989).
22. M.A. Baily, *GROWTH. Genet. Hormones* **8**(suppl. 1), 54–57 (1992).
23. R. Stanhope, *GROWTH. Genet. Hormones* **8**(suppl. 1), 6–8 (1992).
24. B. Lippe and S.D. Frasier, *J. Pediatr.* **115**, 585–587 (1989).
25. A.W. Root, *GROWTH. Genet. Hormones* **8**(suppl. 1), 1–5 (1992).
26. D.B. Allen and N.C. Fost, *J. Pediatr.* **117**, 16–21 (1990).

27. National Institutes of Health, Clinical Research Project No. 84-CH-148 (later renumbered 91-CH-0046), September 18, 1990.
28. L. Cuttler et al., *J. Am. Med. Assoc.* **276**, 531–537 (1996).
29. B.B. Bercu, *J. Am. Med. Assoc.* **276**, 567–568 (1996).
30. National Institutes of Health, Report of the NIH human growth hormone protocol review committee, October 2, 1992.
31. B.M. Lippe, *GROWTH. Genet. Hormones* **8**(suppl. 1), 31–34 (1992).
32. C.A. Tauer, *IRB* **16**(3), 1–9 (1994).
33. G.B. White, *Kennedy Inst. Ethics J.* **3**, 401–409 (1993).
34. J.M. Gertner, *GROWTH. Genet. Hormones* **8**(suppl. 1), 18–22 (1992).
35. S. Bertelloni et al., *J. Pediatr.* **135**, 367–370 (1999).
36. Physicians Committee for Responsible Medicine, *Concerns About Growth Hormone Experiments in Short Children*, Fact sheet, Washington, DC, 1993.
37. J. Mendelson, *Int. J. Risk Safety Med.* **4**, 61–70 (1993).
38. R. Herman, *Washington Post*, Health Section, 9, June 29 (1993).
39. M.J. Pramik, *Genet. Eng. News* **19**(1), 15, 27, 32 (1999).
40. M. Ho, *Genentech Inc.*, Series U.S. Research, Goldman Sachs, New York, October 20, 1993.
41. J. Walsh, Debate over BST expands, *Star Tribune* (Minneapolis), A1 and A9, February 5 (1994).
42. M.L. Vance and N. Mauras, *N. Engl. J. Med.* **341**, 1206–1216 (1999).
43. P.H. Sonksen and A.J. Weissberger, *GROWTH. Genet. Hormones* **14**, 41–48 (1998).
44. L.E. Underwood, *Hosp. Prac.* 192–198, April 15 (1992).
45. L.E. Underwood and P.A. Rieser, *Acta Paediatr. Scand.* **362**(suppl.), 18–23 (1989).
46. Discussion I: C & D, *GROWTH. Genet. Hormones* **8**(suppl. 1), 17 (1992).
47. Discussion III: A & B, *GROWTH. Genet. Hormones* **8**(suppl. 1), 41–45 (1992).
48. P.C. Hindmarsh and C.G.D. Brook, in A.B. Grossman, ed., *Clinical Endocrinology*, Blackwell, Oxford, UK, 1992, pp. 810–836.
49. M. Verweij and F. Kortmann, *J. Med. Ethics* **23**, 305–309 (1997).
50. E.T. Juengst, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998, pp. 29–47.
51. E.T. Juengst, *J. Med. Phil.* **22**, 125–142 (1997).
52. D.L. Young-Hyman, in C.S. Holmes, ed., *Psychoneuroendocrinology: Brain, Behavior, and Hormonal Interactions*, Springer-Verlag, New York, 1990, pp. 40–55.
53. B. Stabler and L.E. Underwood, eds., *Slow Grows the Child*, Lawrence Erlbaum, Hillsdale, NJ, 1986.
54. J.E. Sabin and N. Daniels, *Hastings Center Rep.* **24**(6), 5–13 (1994).
55. D.E. Sandberg, *GROWTH. Genet. Hormones* **8**(4), 5–6 (1992).
56. D.E. Sandberg, A.E. Brook, and S.P. Campos, in B. Stabler and L.E. Underwood, eds., *Growth, Stature and Adaptation: Behavioral, Social, and Cognitive Aspects of Growth Delay*, University of North Carolina Press, Chapel Hill, NC, 1994, pp. 19–33.
57. L.D. Voss and J. Mulligan, in B. Stabler and L.E. Underwood, eds., *Growth, Stature and Adaptation: Behavioral, Social, and Cognitive Aspects of Growth Delay*, University of North Carolina Press, Chapel Hill, NC, 1994, pp. 47–64.
58. R.S. Clopper, *GROWTH. Genet. Hormones* **8**(suppl. 1), 27–30 (1992).
59. Discussion IV, C & D, *GROWTH. Genet. Hormones* **8**(suppl. 1), 64–67 (1992).
60. C.G.D. Brook, *Br. Med. J.* **315**, 692–693 (1997).
61. P.T. Siegel, in C.S. Holmes, ed., *Psychoneuroendocrinology: Brain, Behavior, and Hormonal Interactions*, Springer-Verlag, New York, 1990, pp. 17–39.
62. B. Stabler, *GROWTH. Genet. Hormones* **8**(suppl. 1), 24–26 (1992).
63. P.T. Siegel and N.J. Hopwood, in B. Stabler and L.E. Underwood, eds., *Slow Grows the Child*, Lawrence Erlbaum, Hillsdale, NJ, 1986, pp. 57–71.
64. B. Stabler et al., *Clin. Endocrinol.* **36**, 467–473 (1992).
65. B. Stabler et al., *Psychosom. Med.* **56**, 175 (1994).
66. D.S. Pine, P. Cohen, and J. Brook, *Pediatrics* **97**, 856–863 (1996).
67. R.L. Hintz, *GROWTH. Genet. Hormones* **8**(suppl. 1), 10–14 (1992).
68. S.D. Frasier, *GROWTH. Genet. Hormones* **8**(suppl. 1), 15–16 (1992).
69. D.S. Diekema, *Persp. Biol. Med.* **34**, 109–123 (1990).
70. Discussion II, A & B, *GROWTH. Genet. Hormones* **8**(suppl. 1), 30–31 (1992).
71. *Washington Post* Staff, reprinted *Star Tribune* (Minneapolis), A12, November 3 (1996).
72. Editorial Staff, *Star Tribune* (Minneapolis), November 27 (1996).
73. National Highway Transportation Safety Administration, Proposed rule for 49 CFR 552, Available at: *http://www.nhtsa.gov/cars/rules/rulings/AAirbaqSNPRM/Index.html*
74. Health Care Financing Administration, *1998 Data Compendium*, U.S. Department of Health and Human Services, August 1998, Available at: *http://www.hcfa.gov/stats/stats/htm*
75. National Institutes of Health, Press briefing, FY 2001 President's Budget, Available at: *http://www4.od.nih.gov/ofm/budget/fy2001/Pressbriefing.htm*
76. M. Benjamin, J. Muyskens, and P. Saenger, *Hastings Center Rep.* **14**(2), 5–9 (1984).
77. S.L. Gortmaker et al., *N. Engl. J. Med.* **329**, 1008–1012 (1993).
78. J.D. Sargent and D.G. Blanchflower, *Arch. Pediatr. Adolesc. Med.* **148**, 681–687 (1994).
79. Editorial Staff, *Psych. Today* **5**(3), 102, August (1971).
80. P.W. Crosbie, *Soc. Psychol. Q.* **42**, 340–354 (1979).
81. H.J. Dean et al., *Am. J. Dis. Child* **139**, 1105–1110 (1985).
82. H.J. Dean, in C.S. Holmes, ed., *Psychoneuroendocrinology: Brain, Behavior, and Hormonal Interactions*, Springer-Verlag, New York, 1990, pp. 79–91.
83. C.M. Mitchell et al., in B. Stabler and L.E. Underwood, eds., *Slow Grows the Child*, Lawrence Erlbaum, Hillsdale, NJ, 1986, pp. 97–109.
84. R.R. Clopper et al., in B. Stabler and L.E. Underwood, eds., *Slow Grows the Child*, Lawrence Erlbaum, Hillsdale, NJ, 1986, pp. 83–96.
85. D.B. Allen, *GROWTH. Genet. Hormones* **8**(suppl. 1), 24 (1992).
86. J.D. Lantos, *GROWTH. Genet. Hormones* **8**(suppl. 1), 68–70 (1992).

See other HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY entries.

HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY, ETHICS, THE ETHICS OF ENHANCEMENT

ERIC T. JUENGST
Case Western Reserve University
Cleveland, Ohio

OUTLINE

INTRODUCTION

In discussions of ethics in biotechnology one frequently encounters the claim that there is an important moral distinction between using biotechnological tools and products to combat human disease, and attempting to use them to "enhance" human traits. Thus people argue that using biosynthetic human growth hormone to treat an inborn growth hormone deficiency is praiseworthy, but not the use of the same product to increase the height of a hormonally normal short child (1). Similarly, while the use of human gene transfer techniques to treat disease enjoys widespread support from secular and religious moral authorities, a line is usually drawn at using the same protocols to attempt to improve upon otherwise healthy traits (2,3). Even those unwilling to condemn the enhancement uses of biotechnology outright almost all concur that ethics demands that therapeutic applications of these tools be given priority for research and development (4). As a result the distinction has been enshrined in biotechnology policies at both professional and governmental levels, and continues to inform much of the public discussion of new biotechnological advances.

Despite its widespread support as a moral divide, however, treatment–enhancement distinction is not easy to characterize conceptually. It often even seems in danger of evaporating entirely under conceptual critiques even before the question of its moral merits is entertained. If "enhancement" is to keep serving as a significant policy boundary, we should at least be clear about just what it demarcates. Examining the multiple ways the distinction is interpreted within the bioethical and science policy literature can help with this clarification, and that is the goal of this entry. Ultimately it will suggest a normative point: that the interpretations that most accurately identify the moral concerns at stake in the uses of biotechnology are those that focus on the uses that would serve to exacerbate, rather than reform, the social injustices that flow from our intolerance for human biological variation.

TREATMENT–ENHANCEMENT DISTINCTION

The treatment–enhancement distinction is usually used in bioethics to argue that curative or therapeutic uses of biotechnology fall within (and are protected by) the boundaries of medicine's traditional domain, while enhancement uses do not, and to that extent are more problematic as a professional medical practice or a legitimate health care need (5). Unfortunately, making the distinction between treatment and enhancement is not without its own complexities. The distinction is explicated in at least three distinctly different ways, which have different merits as boundary markers for medical research and practice. There are accounts that rely on medicine's own understanding of its professional goals, accounts which rely on theoretical measures of "species-typical functioning" that go well beyond medicine, and accounts that turn on particular concepts of disease.

Professional Domain Accounts

One approach to the enhancement/treatment distinction is to define it in terms of the accepted limits of professional medical practice. On this view, "treatments" are any interventions which physicians and their patients agree are useful and proper, while "enhancements" are simply interventions which are considered to fall beyond a physician's professional purview. Thus physician-prescribed physical therapy to improve muscle strength would be considered legitimate medical treatment, while weight-lifting under a coach's supervision to achieve a particular physique would be considered an enhancement. This view resonates well with a number of contemporary social scientific critiques of biomedicine, which suggest that medicine has no natural domain of practice beyond that which it negotiates with society (6). It also provides a simple normative lesson for professionals concerned about their obligations in specific cases: One takes one's cues from the patient's value system, and negotiates towards interventions that can help achieve the patient's vision of human flourishing (7).

Unfortunately, however, these same features also deny this approach the ability to be of help to those attempting to use the treatment/enhancement distinction in order to regulate gene transfer research. Relying on the conventions of professional practice provides no principled way to classify technological innovations as either within or outside medicine's proper domain until after the fact of their acceptance or rejection by the medical community. To the extent that useful "upper-boundary" concepts are required at the policy level — for societies making health care research allocation decisions, for example — this impotence is an important weakness.

Normalcy Accounts

Fortunately another approach to interpreting the treatment/enhancement distinction is framed explicitly as a

policy tool for separating legitimate health care needs from luxury services. The most developed exposition of this view is Sabin and Daniel's endorsement of what they call the "normal function" standard for determining the limits of "medically necessary" (and therefore socially underwritten) health services (8). Sabin and Daniels argue that an appropriate boundary between medically necessary treatments and optional enhancements can be drawn by thinking about how to provide medical services fairly within a population. Following Daniels' earlier work (1,10), they construe health care as one of society's means for preserving equality of opportunity for its citizens, and define health care needs as those services that allow individuals to enjoy the portion of the society's "normal opportunity range" to which their full array of skills and talents would give them access, by restoring or improving their abilities to the range of functional capacities typical for members of their reference class (e.g., age and gender) within the human species. Daniels has specified this definition of health care needs further by saying that the notion of "species-typical functioning" it relies upon is not "merely a statistical notion," but implies "a theoretical account of the design of the organism," that describes the "natural functional organization of a typical member of the species." Any interventions that would take expand an individual's range of functional capacities beyond the range typical for his or her reference class would count as an (medically unnecessary) enhancement.

The "normal function" approach is a sophisticated attempt to define the limits of social obligations to provide health services for policy purposes, and comes close to accurately reconstructing the rationale behind many actual "line drawing" judgments by health care coverage plans and professional societies. Unfortunately, this approach is also semipermeable in an important way for our purposes.

The first serious problem is the problem of prevention. While efforts at generic "health promotion" straddle the border of biomedicine, efforts to prevent the manifestation of specific maladies in individuals are always accepted as legitimate parts of biomedicine, and would be automatically located on the "treatment" side of the enhancement boundary. On the other hand, one of the ways one can prevent a disease is to strengthen the body's ability to resist it long before any diagnosable problem appears. These forms of prevention attempt to elevate bodily functions above the normal range for the individual (and in some cases the species), and to that extent seem to slide into enhancement. Consider the case that LeRoy Walters and Julie Palmer make for including some genetic enhancements within the domain of legitimate medical needs. They start with the paradigm of a nongenetic preventive intervention — immunization against infectious disease — and then drive their genetic truck through the border-crossing it creates:

> In current medical practice, the best example of a widely-accepted health-related physical enhancement is immunization against infectious disease. With immunizations against diseases like polio or hepatitis B, what we are saying is in effect, "The immune system that we inherited from our parents may not be adequate to ward off certain viruses if we are exposed to them. Therefore, we will enhance the capabilities of our immune system by priming it to fight against these viruses.
>
> From the current practice of immunizations against particular diseases, it would seem to be only a small step to try to enhance the general function of the immune system by genetic means. . . .In our view, the genetic enhancement of immune system function would be morally justifiable if this kind of enhancement assisted in preventing disease and did not cause offsetting harms to the people treated by the technique (4, pp. 13–14).

This argument is bolstered by the fact that the technical prospects for such preventive–enhancement interventions already look good, given gene transfer research now underway to treat ill patients in just those ways. Thus the gene therapist summarizes the prospects for using gene therapy in oncology this way:

> Over the next few years, it appears that the greatest application will be in the treatment of cancer, where a number of genes that have been isolated have the potential to *empower* the immune system to eliminate cancer cells. . . .Human gene therapy cancer trials have also been initiated for insertion of the tumor necrosis factor (TNF) gene into T-lymphocytes in an effort *to enhance the ability* of T-lymphocytes to kill tumors. Another approach has been to insert the TNF gene into tumor cells in an effort *to induce a more vigorous immune response* against the tumor (emphasis added) (10).

Another gene transfer protocol already underway 'treats' people with an inherited high risk of heart disease by increasing the number of low density lipoprotein receptors their blood cells carry, enhancing their ability to clear their high levels of cholesterol from their blood before it causes heart disease (11). If it works to reduce their risk of heart disease, why not use it prophylacticly to reduce my more modest risk? Moreover, if human gene transfer protocols like these are acceptable as forms of preventive medicine, the critics ask, how can we claim that we should be "drawing the line" at enhancement?

Disease-Based Accounts

Probably the most common rejoinder to the problem of prevention is to distinquish the problems to which they respond. Treatments are interventions which address the health problems created by diseases and disabilities — "maladies" in the helpful language of Clouser, Culver, and Gert (12). Enhancements, on the other hand, are interventions aimed at healthy systems and normal traits. Thus, prescribing biosynthetic growth hormone to rectify a diagnosable growth hormone deficiency is legitimate treatment, while prescribing it for patients with normal growth hormone levels would be an attempt at "positive genetic engineering" or enhancement (13). On this account, to justify an intervention as appropriate medicine means to be able to identify a pathological problem in the patient; if no medically recognizable malady can be diagnosed, the intervention cannot be "medically necessary," and is thus suspect as an enhancement.

This interpretation has the advantages of being simple, intuitively appealing, and consistent with a good bit of biomedical behavior. Maladies are objectively

observable phenomena and the traditional target of medical intervention. We can know maladies through diagnosis, and we can tell that we have gone beyond medicine when no pathology can be identified (14). Thus the pediatric endocrinologists discourage the enhancement uses of biosynthetic growth hormone by citing the old adage "If it ain't broke, don't fix it" (15). This interpretation is also the one at work in the efforts of professionals working at the boundary, like cosmetic surgeons, to justify their services in terms of relieving "diagnosable" psychological suffering rather than satisfying the aesthetic tastes of their clients (16), and in our insurance companies' insistence on being provided with that diagnosis before providing coverage for such surgeries.

Unfortunately, this interpretation does also face at least two major difficulties. The first problem that any disease-based interpretation of the enhancement boundary faces is, of course, biomedicine's infamous nosological elasticity. It is not that hard to coin new maladies for the purposes of justifying the use of enhancement interventions. By interpreting the boundary of medicine in terms of maladies, this approach puts the power for drawing that boundary squarely in the profession's hands, with the corresponding potential for abuse. Moreover, the preventive powers of a give enhancement intervention can be difficult to disprove, if the targeted disease never manifests itself. Enhancing interventions would have the advantage of the man who claimed his dance was keeping dragons out of Central Park: Until a dragon lands, it is hard to argue that he's not providing a preventive service.

The more important problem, however, is that for practical purposes no matter how the line is drawn, most biotechnological interventions that could become problematic as enhancement interventions would not have to cross that line in order to be developed and approved for clinical use, because they will also have legitimate therapeutic applications. In fact, most biosynthetic biologicals and gene transfer protocols with potential for enhancement uses will first emerge as therapeutic agents. General cognitive enhancement interventions, for example, are likely to be approved for use only in patients with neurological diseases (17). However, to the extent that they are in high demand by individuals who are merely suffering the effects of normal aging, the risk of unapproved or "off-label" uses of these products will be high (18). This risk poses unique regulatory challenges. Even if, for example, the U.S. Food and Drug Administration (FDA) vigorously attempted to regulate genetic enhancement technology, under its current legislative mandate the agency is unlikely to regulate off-label uses of approved products (19). Moreover, given the current regulatory vacuum surrounding the private practice of reproductive medicine, there is little to prohibit the application of these techniques to early human embryos as well, in hopes of effecting germ-line transformations.

This last point is critical for policy purposes, because it suggests that, in countries like the United States, the real challenge to regulation in this area may not be the development of enhancement interventions or "enhancement research" but about the downstream "off-label" uses of gene therapies for nonmedical enhancement purposes. The policy problems then becomes one of controlling access and use of the technologies, not their research and development. Unless laws are changed, regulation of biotechnological enhancements ill fall under the common law — which, given the absence of FDA regulation of off-label uses remains the most potent potential source of legal regulation — will focus on physician malpractice and actions for lack of informed consent. This presents another set of challenges for the law, since the novelty of enhancement technologies will make it difficult for judges and juries to ascertain the reasonableness of physician behavior (19).

These realities have pressed those who would use the treatment/enhancement distinction for policy purposes to articulate the moral dangers of genetic enhancement more clearly. After all, personal improvement is praised in many spheres of human endeavor, and, as purely elective matter, biomedical interventions like cosmetic surgery are well accepted in our society as means to achieving personal improvement goals.

ENHANCEMENT AS A FORM OF CHEATING

There are two lines of thought that have emerged from this recent work. The first focuses on the idea that biomedical enhancements are a form of social cheating. This is the view that taking the biomedical shortcut erodes the specific social practices that would make the analogous human achievement valuable in the first place. Thus some people argue that it defeats the purpose of the contest for the marathon runner to gain endurance chemically rather than through training, and it misses the point of meditation to gain Nirvana through psychosurgery. In both cases the value of the improvements lie in the achievements they reward as well as the benefits they bring. The achievements — successful training or disciplined meditation — add value to the improvements because they are understood to be admirable social practices in themselves. Wherever a biomedical intervention is used to bypass an admirable social practice, then the improvement's social value — the value of a runner's physical endurance or a mystic's visions — is weakened accordingly. If we are to preserve the value of the social practices we count as "enhancing," it may be in society's interest to impose a means-based limit on biomedical enhancement efforts.

Interpreting enhancement interventions as those which short-circuit admirable human practices has special utility for policy analysis. To the extent that biomedical shortcuts increasingly allow specific accomplishments to be divorced from the admirable practices they were designed to signal, the social value of those accomplishments will be undermined. Not only will the intrinsic value be diminished for everyone that takes the shortcut, but the resulting disparity between the enhanced and unenhanced will call the fairness of the whole game (be it educational, recreational or professional) into question. If the extrinsic value of being causally responsible for certain accomplishments is high enough (like professional sports salaries), the intrinsic value of the admirable practices that a particular institution was designed to foster may even start to be called into question (20). For institutions

interested in continuing to foster the social values for which they have traditionally been the guardians, this has two alternative policy implications. Either they must redesign the game (of education, sports, etc.) to find new ways to evaluate excellence in the admirable practices that are not affected by available enhancements, or they must prohibit the use of the enhancing shortcuts. Which route an institution should take depends on the possibility and practicality of taking either, because ethically they are equivalent.

ENHANCEMENT AS AN ABUSE OF MEDICINE

Unfortunately, some of the social games we can play (and cheat in) do not turn on participants' achievements at all but on traits over which individuals have little control, like stature, shape, and skin color. The social games of stigmatization, discrimination, and exclusion use these traits in the same manner that other practices use achievements: as intrinsically valuable keys to extrinsic goods. Now it is becoming increasingly possible to seek biomedical help in changing these traits in order to short-circuit these games as well. Here, the biomedical interventions involved, like skin lighteners or stature increasers, are "enhancements" because they serve to improve the recipient's social standing, but only by perpetuating the social bias under which they originally labored. When "enhancement" is understood in this way, it warns of still another set of moral concerns.

On this interpretation, what makes the provision of human growth hormone to a short child a morally suspicious enhancement is not the absence of a diagnosable disease or the "species atypical" hormone level that would result: Rather it is the intent to improve the child's social status by changing the child rather than by changing her social environment (21). Enhancement interventions are almost always wrongheaded under this account because the source of the social status they seek to improve is, by definition, the social group and not the individual. Attempting to improve that status in the individual amounts to a moral mistake akin to "blaming the victim": It misattributes causality, is ultimately futile, and can have harmful consequences. This is the interpretation of enhancement that seems to be at work when people argue that it inappropriately "medicalizes" a social problem to use Ritalin to induce cooperative behavior in the classroom. In such cases the critics dispute the assumption that the human need in question is one that is created by, and quenchable through, our bodies, and assert that both its source and solution really lie in quite a different sphere of human experience.

This interpretation of the enhancement concept is useful to those interested in the ethics of personal improvement because it warns of a number of moral pitfalls beyond the baseline considerations that the enhancement–treatment distinction provides. Attempting to improve social status by changing the individual risks being self-defeating (by inflating expectations), futile (if the individual's comparative gains are neutralized by the enhancement's availability to the whole social group), unfair (if the whole group does not have access to the enhancement), or complicitous with unjust social prejudices (by forcing people into a range of variation dictated by biases that favor one group over others). For those faced with decisions about whether to attempt to enhance themselves or their children through gene transfer, this way of understanding enhancement is much more illuminating than attempts to distinquishing it from medical treatment, because it points to the real values at stake. Ideally, one should do no gene transfer that will make an existing social problem worse, even if exacerbating injustice would further one's own interests.

On the other hand, protecting these values is difficult in a pluralistic society like ours, since it means developing ways to policing individuals' complicity with suspect social norms (22). Under the historical shadow of state sponsored eugenics programs, our government is unlikely to promulgate lists of acceptable and unacceptable enhancements, even if the intent of the lists are to protect the interests of those who are unenhanced.

Moreover regulatory limits on access to genetic enhancements in the United States could be ineffectual if individuals could obtain enhancements abroad. In the past, we have seen a number of examples of persons circumventing U.S. laws to obtain medical care abroad, including seeking illegal abortions abroad prior to *Roe v. Wade*, purchasing unapproved AIDS drugs (RU486) and other pharmaceuticals abroad and returning with them to the United States (23), and traveling to foreign countries to obtain infertility treatments that were illegal or unavailable in the United States (24). Most nations' ability under domestic legal authority to control offshore access is extremely limited. Another approach would be to try to prevent people from travelling abroad in the first place in order to obtain contraband enhancements. Again, the United States arguably possesses the authority to restrict travel on grounds of national security (25). Would obtaining enhancements abroad amount to such a threat? Even if it did, travel restrictions have been imposed on travel to specific countries; it would be virtually impossible to restrict travel to a country for a specific purpose, if travel were permitted for other purposes.

CONCLUSION

Clearly, all of the ways of understanding "enhancement" as a moral concept that I've reviewed have limitations. However, all these interpretations do seem to be alive and well and mixed together in the literature on the topic. It is not possible to cleanly assign the different interpretations of "enhancement" to different spheres of ethical analysis. But there do seem to be some rough correlations that might be made. Thus the interpretations that contrast enhancement interventions with "treatments" seem most useful where it is the limits of medicine's expertise that is at issue. Whether medicine's boundary is defined in terms of concepts of disease, or in sociological terms as the scope of medical practice, or in terms of some theory of the human norm, this interpretation at least provides tools to draw that boundary. Moreover, all other considerations being equal, the line that it draws is the boundary of medical obligation, not the boundary of medical tolerance. Using

this tool, enhancement interventions like cosmetic surgery can still be permissable to perform as phsysicians, but also permissable to deny. This has important implications for social policy making about health care coverage, to the extent that society relies on medicine's sense of the medically necessary to define the limits of its obligations to underwrite care. Again, all other considerations being equal, this interpretation of the concept suggests that few enhancement interventions should be actively prohibited by society or foregone by individuals, even when they are not underwritten as a part of health care, since there is nothing intrinsically wrong with seeking self-improvements beyond good health.

By contrast, the interpretations of enhancement that focus on the misuse of biomedical tools in efforts at self-improvement seem the most relevant to issues in the personal, rather than professional, ethics of enhancement. Concerns about the authenticity of particular accomplishments are moral challenges to the individual but find little purchase in the professional ethics of biomedicine, with its focus on the physical safety and efficacy of its tools. The primary policy implications of this interpretation are for the social institutions charged with fostering particular admirable practices: Enhancement interventions that offer biomedical shortcuts to achievement force reassessments within those institutions of the values they stand for and the practices they have designed to foster them.

Finally, at the other end of the spectrum, enhancement interventions that seem to commit the moral mistake of trying to address social problems through the bodies of the potentially oppressed do seem to mark a stronger set of moral boundaries for all concerned. For biomedicine, this concept marks an epistemic limit beyond which medical approaches to problem-solving are not only unnecessary but conceptually wrongheaded. For individuals, parents, and society, these kinds of enhancement interventions risk either backfiring by exacerbating the social problems they are intended to address, or being futile, if they merely result in a shift of the normal range for a given social trait. Where the medicalization account of enhancements fits a given intervention, there does seem to be more justification for stronger warnings, protections, or prohibitions across the board, whether the interventions falls within medicine's boundaries or not (26).

ACKNOWLEDGMENT

The research conducted for this article was supported in part by NIH grant R01 HG-1446-02. I am grateful to Tom Murray, Maxwell Mehlman, Erik Parens, and David Smith for useful discussions of the text and many of the ideas presented here.

BIBLIOGRAPHY

1. N. Daniels, *Growth, Genet. Hormones* **8**(suppl. 1), 46–48 (1992).
2. W.F. Anderson, *J. Med. Philos.* **14**, 681–693 (1989).
3. P. Baird, *Perspect. Biol. Med.* **37**, 566–575 (1994).
4. L. Walters and J. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, Oxford, UK, 1996.
5. E. Parens, *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998.
6. B. Good, *Medicine, Rationality and Experience: An Anthropological Perspective*, Cambridge University Press, New York, 1994.
7. H.T. Engelhardt, *Soc. Philos. Policy* **8**, 180–191 (1990).
8. J. Sabin and N. Daniels, *Hastings Center Rep.* **24**, 5–13 (1994).
9. N. Daniels, *Just Health Care*, Cambridge University Press, New York, 1986.
10. K. Culver, *Genet. Resource* **7**, 5–10 (1993).
11. J. Wilson et al., *Hum. Gene Ther.* **3**, 179–222 (1992).
12. K. Clouser, C. Culver, and B. Gert, *Hastings Center Rep.* **11**, 29–37 (1981).
13. E. Berger and B. Gert, *J. Med. Philos.* **16**, 667–685 (1991).
14. E. Juengst, *J. Med. Philos.* **22**, 125–142 (1997).
15. Ad Hoc Committee on Growth Hormone Usage. *Pediatrics* **72**, 891–894 (1984).
16. K. Morgan, *Hypatia* **6**, 25–53 (1991).
17. P. Whitehouse et al., *Hastings Center Rep.* **27**(3), 14–22 (1997).
18. D. Kessler, *Harvard J. Legis.* **15**, 693 (1978).
19. M. Mehlman, *Wake Forest Law Rev.* **34**(Fall), 671–617 (1999).
20. T. Murray, in T. Murray, W. Gaylin, and R. Macklin, eds., *Feeling Good and Doing Better: Ethics and Nontherapeutic Drug Use*, Humana Press, Clifton, NJ, pp. 107–129.
21. G. White, *Kennedy Inst. Ethics J.* **3**, 401–409 (1993).
22. M. Little, in I. de Beaufort, M. Hilhorst, and S. Holm, eds., *In the Eye of the Beholder: Ethics and Medical Change of Appearance*, Scandanavian University Press, Stockholm, Sweden, 1997, pp. 151–167.
23. J.S. Batterman, *Hofstra Univ. Law Rev.* **19**, 191 (1990).
24. J. Sauer, *Hum. Reprod.* **12**, 1844–1845 (1997).
25. Kriemer, *N.Y. Univ. Law Rev.* **67**, 451 (1992).
26. E. Parens, *Kennedy Inst. Ethics J.* **5**, 141–153 (1995).

See other entries BEHAVIORAL GENETICS, HUMAN; see also HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY entries.

HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY, ETHICS, THERAPY VS. ENHANCEMENT

GERALD MCKENNY
Rice University
Houston, Texas

OUTLINE

INTRODUCTION

In recent years ongoing practices as well as new or anticipated developments in reconstructive surgery, sports medicine, psychopharmacology, human gene therapy, and emerging areas of biomedical engineering aim at altering the appearance of the body; increasing the efficiency, capacity or productivity of various human functions or performances; and changing features of the personality or mood. Conveniently (if imprecisely) referred to as enhancements, these phenomena, insofar as they are implemented through the institutions and practices of medicine, raise the question of whether their pursuit can be distinguished from the treatment of diseases, and if so, what normative significance, if any, such a distinction has. At stake in this distinction are important ethical and policy issues. As various authors have argued, the distinction between therapy and enhancement is commonly invoked to determine the composition of a basic health care insurance package and to limit the medicalization of human life (1).

THEORIES OF HEALTH, DISEASE, AND ENHANCEMENT

It may seem obvious that, for example, a chemotherapy regimen to treat cancer and a pharmacological agent or (should it become possible) a gene transfer to increase one's powers of concentration fall under radically different categories of intervention. It is natural for many to think of the former as the treatment of a disease and the latter as the enhancement of a capacity. However, there has been significant disagreement over how to account for such a difference and even whether it can be maintained at all in the end. The controversy derives from rival theories of disease in the philosophy of medicine. These include naturalist theories for which disease is a value-free concept grounded in the biological and medical sciences, and normativist theories for which disease is a value-laden concept referring to certain states that individuals (or groups or societies) seek to avoid or overcome. There are strong and weak versions of each type of theory.

Strong Naturalism

The most influential strong naturalist theory is that of Christopher Boorse (2,3). Analyzing what he takes to be the traditional medical understanding of disease, Boorse argues that (*1*) health (or normality) is the absence of disease (or pathology); (*2*) because the medical sciences are based on physiology, pathology is defined with reference to the survival and reproductive competence of the organism as the goals physiology studies (recognizing that other goals, e.g., survival of genes or ecological equilibrium, may be important for other branches of biology); (*3*) a pathology is a reduction of a part-function at any level of the interlocking hierarchy of functional processes (ranging from organelle to cell to tissue to organ to gross behavior) below its statistically species-typical range as determined with reference to the relevant age and sex class. While other theories of disease make use of a statistical range to define normality and pathology, Boorse's emphasis on physiology with its notion of a species design consisting of interlocking functional processes ultimately explained in terms of their contribution to survival and reproductive competence makes him a naturalist.

Boorse's theory has the advantage of enabling one clearly to distinguish pathologies from other characteristics. For example, to cite David Allen and Norman Fost's now-familiar scenario (4), one can distinguish short stature due to human growth hormone deficiency (a part-function operating below the normal range) from short stature as an inherited trait reflecting normal human genetic diversity (assuming that a currently unknown genetic dysfunction is not the cause). Similarly Boorse's theory enables one to distinguish, with James Sabin and Norman Daniels (5), a case of shyness characterized by interpersonal sensitivity and defensive withdrawal due to a bipolar disorder from shyness that simply falls on one end of a normal distribution of social adaptation. Moreover, because many deformities (including harelip and cleft palate) typically involve dysfunction as well as deformity, they may be distinguished from other features (ranging from male baldness to much more severe structural defects) that do not involve dysfunctions. And because age is used as a reference group to determine pathology, dementia among older adults (an abnormal condition) can be counted as a disease while osteoporosis among postmenopausal women (a normal condition) is not. Finally, part-functions operating below the normal range can be distinguished, using this theory, from those operating at the low end of the normal range. For each of these pairs, improvements of the first element in the pair would count as therapies, while improvements of the second element would be considered enhancements.

How would Boorse's theory fare in view of more complex cases? Actual and hypothetical developments in human gene transfer raise the question of how the theory would evaluate the enhancement of certain functions in the course of treating or preventing disease. For example, Juan Manuel Torres describes a multidrug resistance protocol that enhances the capacity of the bone-marrow cells of cancer patients to resist certain side effects of chemotherapy (6). Would this be considered therapy or enhancement? While the protocol involves raising a part-function beyond the normal range, this enhancement (as Torres argues) is simply a necessary part of a procedure aimed at the removal or mitigation of a severe pathology. So long as the various elements or phases of a course of treatment are explained and evaluated in terms of the overall process or its end, there is no reason not to describe the intervention as part of the treatment of a disease. The same reasoning would apply to measures such as immunizations, which enhance an immune function as a necessary condition of preventing specific pathologies.

But what about an intervention such as the one proposed by Leroy Walters and Julie Palmer, which would aim not at a specific pathology but at increasing the capacity of the immune system as a whole to resist pathology more generally (7)? Would this constitute treatment or enhancement?

Nothing in Boorse's theory suggests that the category of pathology in medical science is restricted by the level and scale of the intervention, the complexity of the part-function involved, or the nonspecificity of the pathology. Would enhancing the immune system in this way nevertheless fall under the somewhat dubious category of positive health, that is, a state beyond the mere absence of pathology? Probably not, since such an intervention would not seem to violate Boorse's three criticisms of notions of positive health, namely that such notions lack a clear limit (in this case, autoimmune disorders would establish the limit of improvement), involve trade-offs among incompatible positive goals (the goal in this case still refers to freedom from disease rather than to a positive ideal of somatic good), and are value laden in a way health as the absence of disease is not (8). Boorse's theory, then, seems capable of handling all of these difficult cases. However, in order to cover these cases, it would be more precise to follow Walters and Palmer in distinguishing between health-related and non-health-related enhancements.

A final question for Boorse concerns the distinction between genetic disease and genetic variation. Returning to an earlier example, how would his theory greet the discovery of one or more genetic patterns linked with short stature? The mere existence of such a pattern could simply be an instance of human diversity and would therefore be insufficient for designating that pattern as a pathology. But what if the pattern resulted in a low output or inhibited performance of some part-function that depressed the latter below a normal range? Since disrupted part-functions at any level of the organism count as pathologies, it would appear necessary to describe the condition as such and to consider an appropriate therapy as the treatment of a disease. This example illustrates how genetic knowledge could require one to reclassify as therapies many interventions that would now, on Boorse's theory, be considered enhancements.

Boorse distinguishes sharply between the theoretical level of medical science (specifically, pathology) and the clinical level of medical practice. This raises the question of what normative significance his theory has for policy and medical practice. Boorse recognizes that his is a theoretical and not a clinical concept of disease. On the one hand, many minor, merely local or compensated pathologies (e.g., small benign internal tumors, minor warts or scars, mild cirrhosis of the liver) are undiagnosable and/or have no gross effects on disability, deformity or distress; on the other hand, physicians legitimately perform functions (e.g., childbirth, male circumcision, treatment of osteoporosis) other than the treatment of disease. Moreover he recognizes that diagnostic and therapeutic acts are always subject to the moral criteria elaborated in the bioethics literature; clinical pathology, unlike pathology as a theoretical medical science, is not value free. However, Boorse describes the removal, mitigation, and prevention of pathology as the *core* of medicine while other activities (including enhancement) are *peripheral*: not necessarily illegitimate, but more controversial and lacking the objectivity and urgency of the treatment of pathology. He argues that the treatment of disease constitutes a presumption for clinical medicine because health (biological normality, i.e., absence of pathology) is almost always in the interests of patients and is neutral to most choices of activity and lifestyle. When these conditions fail to hold — when health is not in the interest of a patient or is not neutral with regard to activity or lifestyle — the presumption is defeated and other values may take priority over health.

However, it is doubtful that the treatment of disease constitutes a presumption even on Boorse's own account. The interests or values of the patient, and not the concept of disease, appear to be the normative core of clinical medicine; they determine when and to what extent health is the legitimate goal of clinical practice. Of course, it would be wrong to conclude from this alone that enhancements are justified by the interests or values of patients — that health should not be pursued in some cases does not entail that enhancement should or may be pursued — but there are some cases, such as the treatment of osteoporosis in postmenopausal women, where enhancements clearly should be pursued. It would be odd and even irresponsible to describe these cases as the overriding of a presumption and lacking in urgency and objectivity.

For all of these reasons it is questionable even on Boorse's account whether the objectivity of pathology in the theoretical sphere can simply be carried over into the clinical sphere (even if, as seems likely, biological normality will be the highest priority of clinical practice in the majority of cases). Hence, while Boorse gives a convincing argument for a theoretical distinction between treating a disease and enhancing a trait, the usefulness for policy makers and clinicians of a theory that places treatment of osteoporosis in the same category as facelifts is minimal at best. It says nothing about which health care services should be covered or what limits should be placed on medicalization.

Weak Naturalism

Not all forms of naturalism assume that science requires value- free concepts of health and disease that leave a gulf between medical theory and medical practice. Leon Kass's naturalism differs from Boorse's in two crucial respects (9). First, health is not simply the absence of disease but the wholeness and well functioning of the organism. There are degrees of health, and health can be promoted and maintained even in the absence of disease. The notion of wholeness cannot be captured in a physiological description and, especially in Kass's later essays, involves an experiential oneness of the lived body.

Nevertheless, this concept is still biological. It is manifested in phenomena such as the self-healing and pain-response capacities of organisms, and while Kass describes health as the fitness of the organism for its characteristic kinds of activity, he does not, as normativists do, define health in terms of what individuals

or societies value. The difference between health and mere absence of pathology, the emphasis on wholeness with its abstraction from specifiable part-functions, and the lack of a biostatistical normal range would all seem to encourage Kass to welcome health-related enhancements. His theory would seem especially hospitable to those health-related enhancements (e.g., improving the performance of the immune system) that target the organism as a whole. However, Kass never argues for categorizing any such enhancements under the pursuit of health, and his view of medicine as most properly cooperative with, rather than transformative of, the body's own processes seems to cast a general suspicion on the entire enterprise of technological enhancement (assuming that cooperation and transformation can be consistently distinguished).

Second, Kass formulates a biology grounded on reflection on formal and final causes, thus closing the gap Boorse opened between value-laden medical practice and the value-free science on which medicine is based. No longer limited to physiology with its narrowly focused concern with individual survival and reproductive competence, medicine for Kass is concerned with "powers and desires for work, friendship, love, learning, awareness, mobility, thought and memory, self-command, and the sheer enjoyment of life." Health is not the realization of these goods, but the fitness of the organism for pursuing them. The biology medicine presupposes thus constitutes a broad view of human fulfillment — one that would appear to authorize medicine to extend far beyond the domain of health and disease and would appear to endorse a wide range of non-health-related enhancements. Nevertheless, Kass is highly critical of the use of medicine to fulfill individual preferences or desires, to advance societal priorities or programs, or to alter or replace basic physical processes or social institutions.

The problems with Kass's naturalism follow from the indeterminacy, vagueness, and ambiguity of his major claims. First, it is unclear whether in the end Kass's emphases on degrees of health and on health as wholeness differ significantly from Boorse. Within a biostatistical normal range, there are degrees of difference between those at the higher and lower ends; in this sense health even for Boorse is not *merely* the absence of disease. Health maintenance and promotion, highly commended by Kass, either merely prevent or mitigate disease, or they advance one from a lower to a higher level on a normal range (i.e., they enhance one, in the health-related sense of enhancement). The difference with Boorse seems to be an optical illusion produced by fundamentally different theoretical orientations. Boorse exhibits the modern tendency to treat pathology as basic, with death as its ultimate epistemological and ontological condition. Health is defined negatively, as the absence of pathology, namely normality, a biostatistical notion. Kass aims to recover what he believes is a classical tradition in which health is basic and is not mere biostatistical normality but greater or lesser fitness for the characteristic activities of the organism. Hence what Boorse would simply describe as statistically diverse levels of a single state of being normal, Kass would describe as different degrees of health. Of course, Kass may still wish to distinguish pursuits of higher degrees of health from enhancements by distinguishing those enhancements that cooperate with nature from those that transform nature. But it is not clear what would determine such a distinction. Does it ride on the degree of alteration, on the extent to which the altered capacity exceeds the human average for that capacity, on the type of intervention involved or the technology it employs?

Second, it is unclear whether Kass can distinguish his broad list of functions that constitute health from the desires, preferences, ideals, projects, and effects that for him fall outside the pursuit of health. Kass is aware of the difficulties involved in giving an account of what would constitute a healthy power or desire for, say, work or learning without specifying the particular kinds of work or learning that are the object of individual desires, aptitudes or commitments. But without such a distinction it seems impossible to distinguish the pursuit of health from what Kass must regard as the pursuit of idiosyncracy, or health from enhancement. Health would be relative to the particular activities and goals of individuals, making Kass a normativist. One way out of this problem would be to try to distinguish general purpose interventions — those that improve capacities needed for all or nearly all ways of life — from idiosyncratic interventions that equip one for a limited range of ways of life at the expense of others. However, in addition to problems involved in making and defending this distinction beyond the obvious cases of treating serious disease, its definition of health in terms of ways of life would move Kass into the normativist camp once again.

Third, as Kass's list of human capacities and activities indicates, health can affect almost any human activity or practice. It is likely that enhancement technologies will be increasingly capable of improving "powers and desires" for these capacities and activities indefinitely. Unless, then, he can determine what levels of realization are appropriate for each capacity and what role, if any, medical intervention should play in their realization, the medicalization of these capacities and activities will go virtually unchecked. Finally, even when Kass's reflective biology leads him to approve certain procedures, such as in vitro fertilization (IVF), that are not strictly health-related (IVF does not treat infertility, and infertility, for complex reasons, is not actually a matter of disease and health for Kass) but that fulfill desires that are, according to his biology, natural, he is unsure whether physicians should perform them or not. He seems torn between his conviction that medicine should ideally restrict itself to the pursuit of health and his belief that bringing such procedures under the umbrella of medicine is the best strategy for curtailing their abuse.

In sum, despite his effort to close the gap between the natural and the normative, it is not possible to derive from Kass any clear guidelines regarding basic health care coverage or the limits of medicalization.

Strong Normativism

Strong normativist theories permit a much simpler treatment of the therapy–enhancement distinction because most of these theories deny that any such distinction

can be made in a publicly binding way. H. Tristam Engelhardt, Jr. repeatedly claims that problems become medical problems when (*1*) they appear as a failure to achieve a valued state such as a certain level of freedom from pain or anxiety, a certain level of function, a particular realization of human form or grace, or an expected span of life, and when (*2*) they are problems of a sort that cannot be willed away and are embedded in a web of anatomical, physiological, or psychological causal forces that are open to medical explanation and manipulation (10). This does not prevent Engelhardt from making distinctions among such problems. First, some conditions are likely to be disvalued in nearly all cultures and environments and in light of nearly all human purposes, while others will be disvalued only under some such circumstances. Second, the judgment that a condition is a matter for medicine rather than, say, law or religion will for some conditions (e.g., appendicitis) be a virtually unavoidable judgment, while for other conditions (e.g., alcoholism or criminal behavior) the plausibility of such a judgment will be contested. Third, there are both negative senses of health (the absence of particular diseases, deformities, and dysfunctions) and positive senses (enhancement of capacities, augmentation of pleasures, etc.).

It is natural to refer to the former phenomena in each of these three pairs as treating, describing, or diagnosing a disease and the latter phenomena as involving the enhancement of a trait. But for Engelhardt this could be misleading. To group phenomena under one or the other of these categories merely reflects, respectively, the range of contingent agreements and disagreements concerning which states are and are not disvalued; the purposes and circumstances that make one or another judgment more useful and plausible; and a particular view of which among the diversity of human traits are species typical, normative, and natural. Only within a particular substantive view of the good is it possible to determine which states genuinely constitute values and disvalues; what properly comes under the domains of medicine, law, and religion; and where to draw the line between negative and positive health. It follows that the decisions regarding the composition of insurance coverage and the proper limits of medicalization can be made only within particular communities with their substantive views of the role of medicine in the overall human good.

The plausibility of Engelhardt's account follows from the insuperable difficulties both strong and weak naturalists face in arguing, from within naturalism, for any normatively binding force for their distinctions between therapy and enhancement. Its implausibility follows from its failure to understand the relation of medicine to the sciences of physiology and pathology. In his criticism of Boorse, Engelhardt argues that to define disease in terms of what affects the individual is arbitrary in light of what evolutionary biology tells us about nature's preference for species or genes over individuals. However, Boorse argues that disease is a concept in the medical science of pathology, itself based on physiology, a science which focuses on goals of individual organisms (specifically, on their survival and reproductive competence). It is therefore a mistake to argue that concepts of disease arbitrarily privilege the individual organism over the species. Of course, it is true, as Boorse admits, that the choice medical practice (usually) makes to combat disease rather than to serve the evolutionary fitness of the human species is a normative choice. But this proves only that medicine (as a clinical practice) is normative; it says nothing about disease (11).

Similarly one may, as Engelhardt (on the basis of his views of the philosophy of science and the history of medicine) does, reject Boorse's claim that physiology and pathology are altogether value free, question whether the concept of disease they support can in difficult cases be distinguished as rigorously from other states, as Boorse assumes, and deprive the concept of disease of any normative significance for the practice of medicine — one may do all of this without denying that medicine, on the basis of physiology and pathology, can in principle distinguish disease from other states of the organism. Of course, Engelhardt could concede this and still argue that basic health care coverage and the limits of medicalization must both be established by particular communities. However, with regard to insurance coverage, Engelhardt himself admits that many conditions are regarded with near (though still contingent) universality as disvalues. If so, it is reasonable to assume for purposes of public policy that most serious pathologies would fall under this description and thus be eligible for coverage, while also allowing trade-offs for more highly contested conditions and leaving room for particular individuals and communities to determine the rest (perhaps even in the form of vouchers if one's theory of justice demands, as Engelhardt's does not, a level of public support of health coverage that extends beyond what most people regard as serious conditions).

Medicalization is more difficult. While particular communities could be left to determine for themselves what role medicine should play in their pursuit of what they understand is the human good, it will be necessary for many policy purposes and for purposes of criminal and tort law to make public decisions on the proper domain (whether that of law, medicine, or religion) of conditions such as alcoholism and socially disruptive behavior.

Weak Normativism

Not all normativist theories deny the possibility of public agreement in principle on the distinction between therapy and enhancement. K. Danner Clouser, Charles Culver, and Bernard Gert define a malady (an inclusive term covering what injury, illness, sickness, disease, trauma, wound, disorder, lesion, syndrome, etc. have in common) as a condition of an individual, other than his or her rational beliefs and desires, such that he or she is incurring or at significant risk of incurring a harm or evil (death, pain, disability, loss of freedom, or loss of pleasure) in the absence of a direct sustaining cause (12). Maladies are harms (or significant risks of harm) that are caused in a certain way and that rational persons want to avoid unless they have good reasons not to avoid them. The normative core of this concept is clear in the notion of a harm that one wants to avoid. But unlike strong normativism, the reference to what rational persons want in the absence of good reasons to the contrary stakes a claim to universality. Maladies are disvalued states, but the disvalue is objective,

assuming that the underlying theory of rationality, which Gert develops elsewhere (13), is true.

The notions of harm, significant risk, and absence of a direct sustaining cause all have significance for the distinction between therapy and enhancement. Of their various categories of harm, Clouser, Culver, and Gert devote the most attention to disability. Here they differ from Boorse in two important respects. First, rather than specify age as a reference class against which to evaluate a level of function, they highlight for each ability the stage in normal development when that ability is at its peak. Prior to that stage, to lack the ability is to have an inability; after that stage, to lack it or lose it constitutes a malady. This means that when humans of very advanced age lack or lose an ability (e.g., to walk a certain distance), they suffer a malady even if, say, 98 percent of those in their reference class lack this ability. To attempt to return such individuals to, or to maintain them at, a level of ability that is normal for the species as a whole, regardless of age, is therefore to treat a malady, not to enhance a characteristic. This denial of the relevance of age leads Clouser, Culver, and Gert to classify menopause as a malady, since it constitutes the loss of an ability women normally have.

Second, what counts as lacking or losing an ability is determined by a statistical normal range, but without a concept of species design. This means that a person who develops an extraordinary ability, such as running a marathon in fewer than three hours, and then loses that ability does not suffer a malady (unless her ability is so compromised that it falls below the normal range). It also means that statistical normality, not function, determines which reactions to environmental factors and what levels of risk are maladies. Since nearly everyone tends to become short of breath in a smoke-filled room and to attract mosquitoes, these conditions are not maladies; for the same reason, improving the performance of the immune system would constitute an enhancement.

The concept of a distinct sustaining cause is also crucial for distinguishing maladies from other conditions. For example, individuals whose stature falls below the normal range suffer a variety of discriminatory attitudes and practices. These constitute a distinct sustaining cause of the suffering: when these attitudes and practices are not in effect, the harm disappears. Thus this condition is analogous to the pain that accompanies a wrestler's lock and disappears when the hold is released, and is distinct from a genuine malady such as an allergy, which also involves factors external to the individual but which persists for a period of time even after these external factors are removed.

The harms individuals suffer due to skin color also have a distinct sustaining cause in social attitudes and thus do not constitute a malady. What, then, of the harms individuals suffer due to extreme deformities or disfigurations? Clouser, Culver, and Gert argue that unlike responses to skin color, which are learned responses and vary among societies, responses to these conditions are "universal," spontaneous, and "natural," indeed "like the natural environment." This argument is confusing, but the point seems to be that one may distinguish allergies and severe deformities from short stature and skin color by arguing that responses of the environment to the former are "normal" and intractable, while responses to the latter are variable and alterable. If this is so, the causes of the latter conditions are distinct from the individual in a way that the causes of the former conditions are not. Assuming that responses to deformity can be assimilated to nature in this way (a major assumption), gross deformities and allergies would be maladies while short stature and skin color would not.

While critics of the Clouser-Culver-Gert theory have attacked as counterintuitive or sexist its designation of pregnancy (which involves significant risks of various maladies) and menstruation (which involves pain and discomfort) as maladies (14), the features outlined above also have problematic implications. The problems follow from the effort to distinguish between maladies and normal human variation using the notions of a normal range and absence of a distinct sustaining cause, but without the notions of species design and references classes found in Boorse's strong naturalism. For this theory, procedures aimed at the postponement or reversal of menopause (perhaps no longer an outrageously unlikely prospect) would constitute treatment of a malady, but curing or relieving a mild asthmatic condition that prevents a marathon runner from competing at her previous level would not. A near-universal risk of succumbing to a virus, being statistically normal, apparently would not count as a malady. Gross deformities would count as maladies only if it could be proved that negative responses to them are natural to human beings.

In fairness to this theory, Clouser, Culver, and Gert do not recommend discarding all of the terms for which malady constitutes the commonality. It may be possible, then, to regard the marathon runner's asthma as a disease though not a malady, and its relief or cure as a therapy and not an enhancement. However, this simply raises the questions of how the theory *would* distinguish treating a disease and enhancing a trait and, more generally, what advantage the concept of malady has if it cannot cover admitted instances of diseases with significant effects.

Clouser, Culver, and Gert claim that the concept of malady does have several advantages over rival theories, but this is questionable. Two of the advantages claimed for it, namely the recognition that abnormality is neither a necessary nor a sufficient definition of disease (or related terms) and that malady applies equally to mental and physical conditions, are shared by all of the theories examined here. A third alleged advantage, that the concept of malady could be useful for setting precendents and negotiating borderline cases in determining insurance coverage, is doubtful: It would be an unusual insurance plan that would cover reversal of menopause but not asthmatic conditions of athletes. The fourth alleged advantage would occur if genetic interventions make it possible to enhance properties such as height, intelligence, memory, and strength: They argue that the concept of malady may help determine whether unenhanced (but still statistically normal) levels of these properties constitute deficiencies or not. But any theory that makes use of a normal range will offer this benefit. The more difficult

question is whether these properties are candidates for disease or malady or simply represent human diversity. This question turns not on the differences between malady and disease but on whether it is more plausible to link the normal range to a concept of species design (as Boorse would) or to the absence of distinct sustaining causes.

Finally, Clouser, Culver, and Gert quite plausibly deny that any concept of malady or disease can establish the proper limits of medicalization by determining what medicine should and should not choose to provide. In sum, this theory does not seem to offer, in comparison with other theories, any advantages that would offset its disadvantages.

ARGUMENTS FOR THE NORMATIVE SIGNIFICANCE OF THE DISTINCTION

The foregoing discussion indicates the unlikelihood that a theory of health, disease, and enhancement will succeed in determining which health care services should be covered or what limits to medicalization should be observed. Other arguments begin with a normative principle determined independently of such theories and then attempt to show how the principle requires a distinction between therapy and enhancement, at least in some contexts.

Risk–Benefit, Just Allocation, Discrimination

Two common arguments for distinguishing between therapy and enhancement and for excluding the latter from the practice of medicine refer to the risks of enhancement technologies and the problems they pose for just allocation and discrimination. W. French Anderson (15) and Juan Manuel Torres (16) invoke both arguments to assert the normative force of the line between therapy and enhancement in gene transfer technology, though the arguments and their shortcomings are applicable in other areas of medicine as well. Anderson focuses on the different levels of medical risk involved, respectively, in adding a normal gene to overcome the effects of a nonfunctioning gene and adding a normal gene to increase the productivity of a gene functioning at a normal level, and on the relation of these risks to the likely benefits of each kind of intervention, concluding that only in cases of very serious disease do the benefits outweigh the risks. However, this is only a temporary rationale for a line between therapy and enhancement.

It is reasonable to expect that the enhancement of normally functioning genes will gradually become safe enough that the risk–benefit ratio will increasingly favor at least some uses of gene transfer for non-health-related enhancements, while some uses of gene transfer for the treatment of disease will have unfavorable risk–benefit ratios. In this regard gene therapy will likely resemble every other branch of medicine where lines drawn by risk–benefit ratios cut across the line between therapy and enhancement. Anderson also points to the important and difficult questions of who should benefit from enhancements and how to avoid potentially discriminatory uses of them, arguing that until we resolve these questions we should limit our genetic interventions to the treatment of serious diseases. However, while these are urgent questions in nearly all areas of medicine, our inability to resolve them has not prevented us from intervening into less serious nongenetic conditions without a consensus on who should benefit from such interventions, or from carrying out other interventions and practices that could be accompanied by the discriminatory effects Anderson and Torres cite (i.e., pressure to undergo treatment or to adhere to eugenic goals, exacerbation of the gap between haves and have nots).

Genetic enhancement is neither a necessary nor a sufficient condition for moral problems of these kinds; such problems have occured in the past and present without genetic enhancements, and they need not occur, or even be exacerbated, because of genetic enhancements. It is true that gene transfer technologies could provide new and potentially dangerous occasions for such problems, resulting in discrimination of a much greater magnitude than at present. However, coerced treatment and eugenics would require extensive use of gene transfer in the population, which Torres himself considers highly unlikely, while the odds that gene transfer will expand the gap between haves and have nots more than a nonmedical enhancement such as education already does (with the enthusiastic complicity of many of those who argue against genetic enhancements on these grounds) is just as unlikely.

The Threat to Athletic Competition

Another context in which the distinction between therapy and enhancement is considered relevant is sports medicine. The issue, noted by Dan Brock (17), Eric Juengst (18), Thomas Murray (19), and others, is whether the use of sports medicine for purposes beyond the treatment of injuries destroys the very definition of the activity itself or the significance of the athlete's achievement. However, sports medicine appears to have developed to the point that the line is no longer drawn between therapy and enhancement but between different types and different levels of enhancement. For example, many nutritional supplements aimed at improving the performance of one or another somatic capacity are currently permitted in a number of sports. If effective, these supplements clearly enhance performance; they do not treat an injury or related condition. Of course, other nutritional supplements are banned. But usually this is either because they are deemed to be unsafe, or to have an unfavorable risk–benefit ratio, or because the level of advantage they offer is deemed to be unfair to other competitors or inconsistent with the meaning of the activity, not because they go beyond the treatment of injuries and related conditions. Clearly, then, the distinction between permitted and banned nutritional supplements occurs within the category of enhancements; few among those responsible for the oversight of these sports are advocating a ban on nutritional supplements altogether (if indeed the latter could even be consistently distinguished from ordinary dietary measures).

Fair Equality of Opportunity

Finally, the therapy-enhancement distinction is invoked to determine what medical services a just system of health

care allocation is obligated to provide. Norman Daniels argues that because disease and disability significantly affect the opportunities open to individuals, health care has a central function for justice (20,21). He defines justice in Rawlsian terms of fair equality of opportunity, which requires that opportunity be equal for persons of similar skills and talents (although the resulting inequalities are mitigated by the difference principle in Rawls's theory, which holds that these inequalities must work to the advantage of the least well off) (22). Justice therefore must protect individuals against factors, such as race or sex discrimination and disease or disability, that restrict the range of opportunities that would otherwise be open to persons of similar skills and talents. Justice requires access to health care services that prevent, cure, palliate, or compensate for diseases and disabilities. However, fair equality of opportunity does not require provision of health care services for all conditions that create inequalities of opportunity but only for conditions of disability or disease (i.e., for pathologies).

Daniels's position turns on two points: the principle of fair equality of opportunity itself, and a rationale for including pathologies but not disadvantageous normal human traits in the set of conditions for which society is obligated to provide health care services. Daniels recognizes that his principle of fair equality of opportunity is vulnerable to those who point out that because both pathologies and skills and talents are due significantly, though not entirely, to the natural lottery, they should be treated identically so that either society is obligated to remedy restrictions of opportunity in both cases (social welfarists) or in neither case (libertarians). However, Daniels's arguments against these opponents are largely circular: he repeatedly appeals to certain actual beliefs and practices that reflect his intuitions to criticize other actual beliefs and practices that do not. Moreover, his theory lacks a convincing rationale for excluding disadvantageous traits from fair equality of opportunity. If justifiable, such an exclusion would serve two important purposes: it would keep health care expenses in check (as Daniels notes), and it would help preserve human diversity against the leveling effects of the quest for competitive advantage.

In presenting his case for this exclusion, Daniels quite plausibly draws on Boorse to distinguish pathologies from traits. However, it is not clear why this distinction should matter to fair equality of opportunity. Considering Allen and Fost's example, Daniels observes that short stature has roughly the same effect on equality of opportunity whether it is caused by human growth hormone deficiency (a pathology) or simply reflects human genetic variation. He notes further that in both cases it is equally the result of the natural lottery and equally the object of social prejudice.

Why then should this distinction count for purposes of justice? In defense of the distinction, Daniels simply reasserts his principle of fair equality of opportunity, which (he claims) recognizes from the outset that skills and talents and (he now adds) "other capabilities" are unequally distributed. Styling this view "the standard model" which (he alleges) reflects "our actual concerns" and "our consensus," Daniels in effect tries to salvage the normative significance of the distinction between therapy and enhancement in these cases by pleading that the matter has already been settled in favor of the normativity of this distinction. However, even if this were in fact settled (which it is not), whether the settlement is defensible is precisely what the example calls into question.

Having appealed to an allegedly prevailing adherence to his rule of exclusion from basic coverage, Daniels goes on to undermine the credibility of his own adherence to it by arguing that if an inexpensive treatment for improving the cognitive capacities of children becomes available, there would be compelling reasons — he mentions the enhancement of education, the narrowing of the gap between children at the low end of the normal range and others, and the increase of social productivity — to seek enhancement in this way. Even if one ignores the unlikelihood that such a treatment would in fact realize the second benefit, one must wonder about a theory of fair equality of opportunity that invokes the therapy–enhancement distinction to treat as unequals two individuals who are equally the victims of the natural lottery and social discrimination while it violates that distinction to serve socioeconomic ends and advance the status of persons already in the normal range.

ARGUMENTS AGAINST THE NORMATIVE SIGNIFICANCE OF THE DISTINCTION

Critics such as Kathy Davis (23) and David Frankford (24) seek to show how the therapy–enhancement distinction could, if adopted in the formulation of health care policy, be used by insurers and bureaucrats to deny genuine health care needs on the grounds that they are enhancements rather than therapies. Neither Davis nor Frankford wishes to abandon the distinction, only to ensure that it is applied with context-sensitive judgment. Frankford questions whether such context sensitivity is possible in the policy arena, while the examples Davis draws from the Dutch health care system point to the same conclusion despite her hopes to the contrary.

CONCLUSION

The result of this critical survey is that distinguishing between therapy and enhancement is easier than it is often assumed to be while articulating and defending any normative force for this distinction is more difficult than it is often assumed to be. The implication is that the therapy–enhancement distinction has little if any relevance for determining basic health insurance coverage or the limits of medicalization. With regard to insurance coverage, some procedures that are technically enhancements according to current medical science (e.g., treatment of osteoporosis for postmenopausal women) are almost certainly of higher priority in many cases than some procedures that treat diseases. Fortunately, while the therapy–enhancement distinction does not help in this or in many other cases, it is possible to arrive at a rough agreement on which conditions seriously inhibit almost any way of life, and thus should, in principle, be covered in a basic plan while permitting trade-offs

when there are reasonable disagreements even at this level. If one's theory of justice does not permit all of these conditions to be covered, one may be forced to prioritize these conditions, after the example of the Oregon Medicaid program. If one's theory of justice requires funding basic coverage beyond these conditions, vouchers would enable individuals or groups to determine the composition of this additional coverage in accordance with their views about the relation of various somatic conditions to valued activities or ways of life. Medicalization and the normalization that accompanies it present a more difficult problem. To determine the appropriate limits of existing and emerging technologies will require a view of the ethical significance of the body and its capacities and limitations, and of the place (if any) of the discourses and practices of biomedicine in realizing these capacities and responding to these limitations (25).

BIBLIOGRAPHY

1. E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998.
2. C. Boorse, in D. VanDeVeer and T. Regan, eds., *Health Care Ethics: An Introduction*, Temple University Press, Philadelphia, 1987, pp. 359–393.
3. C. Boorse, in J.M. Humber and R.F. Almeder, eds., *What Is Disease?* Humana Press, Totowa, NJ, 1997, pp. 1–134.
4. D. Allen and N. Fost, *J. Pediatrics* **117**, 16–21 (1990).
5. J.E. Sabin and N. Daniels, *Hastings Center Rep.* **24**(6), 5–13 (1994).
6. J.M. Torres, *J. Med. Philos.* **22**, 43–53 (1997).
7. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997.
8. C. Boorse, *Philos. Sci.* **44**, 542–573 (1977).
9. L. Kass, *Toward a More Natural Science*, Free Press, New York, 1985.
10. H.T. Engelhardt, Jr., *Foundations of Bioethics*, 2nd ed., Oxford University Press, New York, 1996.
11. C. Boorse, in J.M. Humber and R.F. Almeder, eds., *What Is Disease?* Humana Press, Totowa, NJ, 1997, pp. 1–134.
12. K.D. Clouser, C.M. Culver, and B. Gert, in J.M. Humber and R.F. Almeder, eds., *What Is Disease?* Humana Press, Totowa, NJ, 1997, pp. 173–217.
13. B. Gert, *Morality: A New Justification of the Moral Rules*, Oxford University Press, New York, 1988.
14. M. Martin, *J. Med. Philos.* **10**, 329–337 (1985).
15. W.F. Anderson, *J. Med. Philos.* **14**, 681–693 (1989).
16. J.M. Torres, *J. Med. Philos.* **22**, 43–53 (1997).
17. D. Brock, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998, pp. 48–69.
18. E. Juengst, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998, pp. 48–69.
19. T.H. Murray, in T.H. Murray, W. Gaylin, and R. Macklin, eds., *Feeling Good and Doing Better: Ethics and Nontherapeutic Drug Use*, Humana Press, Clifton, NJ, 1984.
20. N. Daniels, *Just Health Care*, Cambridge University Press, New York, 1985.
21. N. Daniels, in T.F. Murphy and M.A. Lappe, eds., *Justice and the Human Genome Project*, University of California Press, Berkeley, CA, 1994, pp. 110–132.
22. J. Rawls, *A Theory of Justice*, Harvard University Press, Cambridge, MA, 1971.
23. K. Davis, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998, pp. 124–134.
24. D. Frankford, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998, pp. 70–94.
25. G.P. McKenny, *To Relieve the Human Condition: Bioethics, Technology, and the Body*, State University of New York Press, Albany, NY, 1997.

See other entries BEHAVIORAL GENETICS, HUMAN; see also HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY entries.

HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY, LAW, GENETIC ENHANCEMENT, AND THE REGULATION OF ACQUIRED GENETIC ADVANTAGES

MAXWELL J. MEHLMAN
Case Western Reserve University
Cleveland, Ohio

OUTLINE

INTRODUCTION

Genetic enhancement, whether in the form of somatic enhancement for adults or for children, genetic selection for enhancement, or germ cell enhancement, may give the recipients significant social advantages. It is impossible at this time to be certain which traits will prove susceptible to genetic enhancement, but they may include physical traits, such as beauty, stature, strength, and stamina,

personality characteristics such as charm, cheerfulness, charisma, confidence, and energy, and mental capabilities, including memory, intelligence, and creativity.

These improvements obviously will be in great demand. But how widely available will the technologies be that make them possible? Some genetically engineered drugs that produce somatic enhancements may be relatively affordable. Others may not. Genetic selection for enhancement in which fetuses were tested in utero might not add much to the cost of performing genetic tests to detect abnormalities or disease (1). This might make it, and the accompanying abortions, the "poor person's" genetic enhancement technique. But any enhancements performed on embryos would be expensive, since they would include the costs of in vitro fertilization (IVF). Currently IVF costs average $25,000, and there would be added costs of the genetic manipulations, which are likely to be substantially greater, particularly when the technology is first introduced.

Genetic enhancement is not likely to be paid for by any public or private health insurance program or policy. This is evident from the lack of third-party payment for cosmetic medicine, which is perhaps the most analogous biomedical technology currently available. The legislation governing the Medicare program contains a general prohibition against paying for "items or services ... which are not reasonable and necessary for the treatment of illness, or to improve the functioning of a malformed body part" (2), and includes a specific exclusion for "cosmetic surgery" (3). States have adopted the same coverage exceptions under their Medicaid programs (4). Private health insurance plans also do not cover cosmetic medicine; the language in the author's high-option Blue Cross policy is typical: "Coverage is not provided for services and supplies ... for surgery and other services primarily to improve appearance or to treat a mental or emotional condition through a change in body form...."

Even if the government wanted to provide general access to genetic enhancement, the cost would prove prohibitive. Widespread access to enhancements such as preimplantation selection or manipulation that depends on access to IVF currently would amount to $120 billion per year for the IVF services alone (5). Somatic enhancement would not be cheaper. A single somatic enhancement in the form of a substance like human growth hormone, which currently costs about $30,000 per child per year (5), would amount to $22 billion a year just to provide to the 1.7 million children who were in the lowest 3 percent of the population in terms of height (6). The figure for multiple somatic enhancements over a number of years for the entire population would be astronomical.

The high cost and the lack of coverage by third-party payment plans, of course, does not mean that no one will access genetic enhancement but only that it will be limited to persons who can purchase enhancement with their own assets. This gives rise to two related problems. The first is inequality, which is also discussed in another entry; the second is unfairness. The issue of fairness would arise at the micro level if genetically enhanced individuals compete for scarce resources against, or find themselves in a conflict of interest with, those who are unenhanced. Genetic enhancement could confer a decisive advantage in competitive circumstances. How should society respond to the potential unfairness?

SPECIAL NATURE OF GENETIC ENHANCEMENTS

In a society whose members believe in the possibility of upward social mobility, people seek to better themselves and their children. They educate themselves and try to obtain the best education they can for their children. They may try to marry "upward," hoping for a mate who will increase their opportunities, social standing, and wealth. They push themselves and their children to cultivate and make the utmost use of their talents.

Many of these efforts take the form of medical or pharmaceutical interventions. People take drugs to improve their athletic and cognitive performance. They subject themselves to surgery to improve their appearance. Some of these activities, such as selecting one's mate, have at least an indirect influence on the genetic makeup of succeeding generations.

Against this background of current enhancement practices, what is so exceptional about genetic enhancement? Society has had plenty of experience coping with the social implications of efforts at self-improvement. While society's response has not always been adequate or successful — witness the difficulties in trying to control the use of performance-enhancing drugs in sports — will the problems created by genetic enhancement be so different that they require special attention?

Of course, even if we felt that wealth-based access to genetic enhancement did not constitute a new kind of threat to social equality, society might still need to respond to it in a vigorous fashion. The additional inequality arising from genetic enhancement, when added to existing sources of inequality, might tip the scales in favor of social unrest. At least, we might well want to monitor the situation closely, and stand prepared to respond if necessary.

Yet in a number of important ways, genetic enhancement does differ from previous sources of inequality and unfairness. Taken together, these differences justify a significantly heightened level of concern, if not outright alarm:

1. The probable high cost of genetic enhancement means that fewer individuals will gain access to it compared to those who can avail themselves of other forms of self-improvement. Twenty million Americans are members of commercial health and sports clubs (7). In 1997, 480,588 persons obtained cosmetic surgery. In contrast, only 39,390 per year obtain IVF (8), and even fewer would be able to afford the additional cost of preimplantation genetic enhancement. Somatic enhancement might be cheaper, but it might still be beyond the reach of many who wished to enhance multiple traits or to produce long-term results.
2. The effects achievable with genetic enhancement are likely to affect some traits that are not highly susceptible to current forms of self-improvement, including

some that are fundamental to personal success. Current self-improvements are limited in scope. One can change one's weight (although usually not permanently); employ cosmetics and cosmetic medicine to improve appearance within certain limits; somewhat increase the ability to cope with loss, failure, and stress; build muscles; develop greater physical, mental, and social skills; and increase reading speed. Genetic enhancement, however, may improve intelligence, cognition, charisma, creativity, energy, cheerfulness, sense of humor, and other characteristics that are arguably central to success and well-being.

3. Persons who are fortunate enough to be able to gain access to genetic enhancements are likely to obtain a much greater and long-lasting advantage than those who employ more traditional forms of self-improvement. Performance-enhancing drugs in sports produce their effects on the basic human phenotype, and the effects, while perhaps enough to win competitions, are relatively modest (9). Although cosmetic interventions change the appearance, they rarely stray from "normal" ranges for physical traits, and with the exception of cosmetic surgery, are often transitory, as any dieter knows. To date, techniques for improving memory and other cognitive functioning do not appear able to significantly increase intelligence or to have a particular profound or permanent effect (10,11). But there is no telling how powerful genetic enhancement can be. It could stretch the limits of desirable human traits considerably, perhaps even indefinitely. For example, there may be no such thing as being "too intelligent." Moreover, enhanced persons can still employ traditional forms of self-improvement on top of their genetically enhanced starting point.
4. Current self-improvement techniques tend to affect at most only a few aspects of performance or appearance at one time. In most cases, people work on one trait — for example, their facial appearance, their weight, their ability to solve puzzles, or memorize facts. Cosmetic polymedicine, while not unknown, is rare (12). Even a professional athlete in full training mode can do no more than exercise and take performance-enhancing drugs to increase strength and stamina, hire a famous coach and perhaps a sports psychologist, and repeatedly practice a skill or routine. Genetic enhancement, on the other hand, may permit wholesale changes in characteristics. Parents with sufficient resources may engineer numerous improvements in their children, and they may purchase multiple somatic enhancements for themselves or their dependents (13).
5. As the result of the ability of genetic enhancement to alter multiple traits in significant ways, genetic enhancements may give people decisive advantages or major success not just in one or two spheres of social activity but in a broad range of social endeavors. This may enable them to cross what Michael Walzer calls "spheres of distributive justice." Imagine the following individual, he says:

> Here is a person whom we have freely chosen (without reference to his family ties or his wealth) as our political representative. He is also a bold and inventive entrepreneur. When he was younger, he studied science, scored amazingly high grades in every exam, and made important discoveries. In war, he is surpassingly brave and wins the highest honors. Himself compassionate and compelling, he is loved by all who know him (14).

   If genetic enhancement made such a person possible, he and his kind would be likely to dominate the rest of society.
6. Finally, unlike most advantages derived from self-improvement, some genetic enhancements — those achieved through genetic selection for enhancement or germ-line engineering — will be incorporated into the genetic makeup of future generations (15). Both the genetic enhancements and the societal advantages that they confer will be inherited, and those who obtain them will comprise a special class within society. Although initially defined by its wealth, this class eventually will come to be characterized by its superior genetic endowment.

In short, genetic enhancement possesses a number of characteristics that raise special concerns for society (16). Some of the objections, like playing God, are metaphysical. Others concern the serious practical consequences for the individual and for society. The rest of this article will concentrate on these two consequences, beginning with the implications for social equality.

THREAT TO EQUALITY FROM GENETIC ENHANCEMENT

Is it fair for some people to have greater genetic advantages than others? This is a question that forms the crux of the age-old problem of "natural inequality" which has plagued philosophers and social theorists at least since the ancient Greeks. If genetic enhancement is unfair, then presumably society should do what it can to rectify the situation, and this engenders the secondary, but equally vexing, problem of what form societal intervention should take and how feasible it would be.

Some philosophers tolerate natural inequality more than others. Meritocrats, for example, welcome substantial inequalities resulting from the distribution of natural talents, pointing to the benefits that accrue to society from the accomplishments of the gifted. John Gardner, for example, objects to what he calls "extreme equalitarianism," which, he states, "ignores differences in native capacity and achievement and eliminates incentives to individuals." In Gardner's opinion, this signifies "the end of that striving for excellence that has produced history's greatest achievements" (17). Others relish excellence as much for its own sake as for what it can achieve. According to Thomas Nagel, "[a] society should try to foster the creation and preservation of what is best, or as good as it possibly can be. . . .Such an aim can be pursued only by recognizing and exploiting the natural inequalities between persons, encouraging specialization and distinction of levels in education, and accepting the variation in

accomplishment which results" (18). Robert Nozick even disputes the idea that the fact that natural assets are arbitrarily distributed in society means that they are not deserved (19).

Philosophers who are morally troubled by inequality, on the other hand, tend to regard it as unjust for some individuals to benefit by virtue of their genetic endowment compared to others who do not fare as well in the genetic lottery (20). These philosophers generally agree that unchosen and unearned advantages and disadvantages must be minimized in order to achieve a more just society. As Rawls states: "It seems to be one of the fixed points of our considered judgments that no one deserves his place in the distribution of native endowments, any more than one deserves one's initial starting place in society" (21; also see Ref. 22).

Although liberal philosophers agree on the goal of rectifying the injustices of the natural lottery, they disagree substantially on how this should be achieved. A basic dispute, for example, concerns just what is to be equalized: "welfare" — that is, some subjective measure of well-being — or "resources" (23,24). Another contentious issue is how much inequality society should tolerate, whether of welfare or resources, in order to assure the production of desired goods. For example, meritocratic, libertarian, and free-market theorists all justify their tolerance for inequality at least in part on the ground that permitting people to profit from the exercise of their natural talents is necessary to induce them to increase the total sum of societal goods. In reaction at least in part to these and other difficulties, many philosophers abandon the quest for absolute equality, whether of resources or welfare, in favor of providing everyone with a minimum level of assets or of well-being, or with "equality of opportunity" (18,25,26).

Despite the gaps and imprecisions in the theory of equality and its application, it does provide at least one clear imperative in regard to wealth-based genetic enhancement: If genetic enhancements are viewed as natural assets, distributed largely by chance, then the principle of equality requires that society attempt to rectify the advantages that they confer, except to the extent that it may be necessary to allow individuals to profit to some degree from their natural talents in order to secure benefits for society as a whole. On the other hand, if genetic enhancements are viewed as "earned" advantages, obtained through diligence and effort, then at least some theories of equality would permit the enhanced individual to retain the additional value created by enhancement.

The easiest case for saying that genetic enhancements were unearned is when they were obtained by children from their parents or other family members. The children have done nothing to entitle them to such advantages; from a moral standpoint, their enhancement is no more than the luck of the draw (27). The interest of equality therefore would seem to justify depriving enhanced individuals of the benefit of enhancements that were installed by their parents or that were purchased with inherited or unearned wealth. Another easy case would be when persons acquired the money to purchase enhancements by immoral means; they too have no moral claim to the benefits.

But what about the person who obtains the necessary funds by dint of the sweat of her brow, without exploiting others or behaving otherwise immorally? This person has a strong claim to be allowed to retain the benefit from her genetic enhancements as morally deserved. Similarly a parent who earned the wherewithal to purchase genetic enhancements for her children in morally acceptable ways may contend that her children ought to be entitled to enjoy the benefits (28).

If we feel obliged to level the genetic playing field even when the advantages of genetic enhancements have been acquired in morally deserving ways, we might base our action on the view that equality is a moral imperative that overrides desert. This rationale no doubt to some extent lies behind schemes that redistribute earned wealth, like progressive income taxation. But the techniques that would be required to level the genetic playing field, as we will see, are far more intrusive than progressive taxation — even when accompanied by aggressive government enforcement. If these methods are to be politically acceptable, they must be premised on more than an abstract belief in the value of equality. They must be based as well on the conviction that genetic enhancement, if left unchecked, would be a grave threat to society. Can this view be sustained? I believe it can.

One of the most important of our societal goals is maintaining a liberal democratic form of government. This goal is directly threatened by wealth-based genetic enhancement: The inequality of social opportunity that results may be so great that a liberal democratic form of government becomes unsustainable, and our political system instead becomes autocratic or oligarchic. This follows from the assumption that a minimum degree of equality is necessary for modern liberal democracy to exist (29). If social inequality becomes too pronounced, liberal democratic political systems become unstable. As one sociologist states:

> Inequality in the distribution of rewards is always a potential source of political and social instability. Because upper, relatively advantaged strata are generally fewer in number than disadvantaged lower strata, the former are faced with crucial problems of social control over the latter. One way of approaching this issue is to ask not why the disprivileged rebel against the privileged but why they do not rebel more often than they do (30).

The characteristics of genetic enhancement that threaten to destabilize liberal democratic government are the features mentioned earlier that distinguish genetic enhancement from other forms of self-improvement: its high cost which may place it beyond the reach of all but the very wealthy, the broad and fundamental nature of the traits that it could enhance, the magnitude of its effects; their multiplicity, the resulting ability to gain advantages in multiple spheres of social activity, and the possibility created by germ-line enhancement that these advantages would be passed on to successive generations.

These characteristics not only give rise to social inequality; more insidiously, they undermine the belief in equality of opportunity. A widespread belief in equality of opportunity is the principal manner in which liberal

democracies accommodate the reality of inequality — that everyone is not equally endowed with natural assets, nor with the same luck or disposition to work hard. In the United States, although most people will tolerate large in equalities in the distribution of resources, the main principle is that everyone have an equal opportunity to these resources (31). As John Shaar notes, the belief in equal opportunity is instrumental in maintaining the prevailing social order:

> No policy formula is better designed to fortify the dominant institutions, values, and ends of the American social order than the formula of equality of opportunity, for it offers *everyone* a fair and equal chance to find a place within that order (32).

Genetic enhancement would create such profound, true differences in ability that they would endow the wealthy with opportunities that are irrevocably beyond the reach of the less fortunate. We have from history other societies with population characteristics similar to those that would be created by wealth-based genetic enhancement. In medieval Europe, individuals were born into their respective classes. Only in rare exceptions were peasants able to obtain education in religious institutions or become apprenticed to a trade, so, they remained found to their station in life (33). In slave-owning societies, people were born into bondage and could be freed only by escape (self-exile) or at the pleasure of their masters. In contempory India the caste system is an example of such a society, and it is a constant threat to the nation's democratic institutions (34).

In short, wealth-based access to genetic enhancement creates not only a moral challenge but a political threat. From a moral standpoint, those who gain enhancement may not have done anything to deserve it. Adults may have come by the means necessary to purchase enhancement in objectionable or morally irrelevant ways — through exploitation or the brute luck of inheritance. Children who are enhanced by their parents are unlikely to have done anything to earn it; this is patent in the case of more remote generations enhanced through prior germ cell manipulations. Yet genetic enhancement poses more than an ethical quandary. Even if the price of genetic enhancement had been earned in a moral sense, the social impact of wealth-based enhancement is likely to be severe. Somatic enhancement alone could so dramatically widen the gulf between have's and have-not's that class warfare would ensue, and the conflict could topple democratic government. Germ-line enhancement could create, quite literally, a master race. The question is whether there is any practical way to prevent this.

PROMOTING GENETIC EQUALITY

In the face of the serious threats to equality represented by wealth-based genetic enhancement, what options do we have to promote equality? One approach would be to "level up." An obvious example is to give everyone access to genetic enhancement regardless of wealth. As pointed out earlier, however, this would be prohibitively expensive.

Even if we decided to divert some enormous portion of the gross national product (GNP) to finance a massive enhancement entitlement program, what enhancement services would such an entitlement program provide? This raises the old argument over resource versus welfare equality. If the objective were to give everyone an equal amount of enhancement resources, those who started out with a more favorable distribution of natural assets would end up better off. If instead we attempted to give everyone an equal or minimum share of enhancement, or an equal or minimum share of enhancement-created opportunity, how would we measure equivalence? Would an extra inch of height be equal to an extra ten points of IQ? The problem would not be solved if we gave everyone an equal share of money and allowed them to purchase whatever enhancements they desired. Logically, unless all enhancements cost the same, those who desired expensive enhancements would be less advantaged than those who were content with cheaper ones — again, the problem of expensive tastes (32). Moreover, since we could not afford to provide everyone with access to the same enhancements that the wealthy could purchase, the wealthy always could stay ahead of the rest of the population. Now we could solve the problem if we gave everyone the maximum amount of enhancements available, but then again we would run into the problem of prohibitive cost. Such problems would plague any attempt to give the unenhanced some countervailing benefit other than genetic enhancements, like money, information, or political power which would level the playing field.

The fact that some people start out with a more favorable distribution of natural assets suggests another approach: subsidize access to enhancements, not for everyone, but for those who were genetically disadvantaged. In other words, bring everyone up to the same level of genetic well-being. This would comport with Rawls's difference principle by improving the fortunes of the least well-off (21). But it runs into the same problems that were just described in attempting to equalize access to enhancements. In addition, allocating enhancements to the genetically disadvantaged would necessitate identifying genetically disadvantaged individuals or groups within the population, and measuring their degree of disadvantage. This would raise serious practical, moral, and political objections. Determining what counts as a genetic disadvantage is similar to trying to identify whether or not someone has a disability — a determination that is controversial, and often appears arbitrary (35–38). Measuring the extent of disability is an even thornier enterprise: Witness the morass that the state of Oregon found itself in when it tried to ration Medicaid services on the basis in part of how much they alleviated disability. Even if we could identify and quantify genetic disadvantage, we would need to establish a "normal" degree of genetic well-being state that the disadvantaged could attain, so we could give them the correct amount of enhancements or money with which to purchase enhancements. But "normalcy," as noted earlier, is highly arbitrary, value laden, and subjective (11). Furthermore it can become a constantly moving target as the distribution of advantages and disadvantages within the population shifts and as

the average level of advantage increases with the number of people becoming enhanced. Finally, any attempt by the government to identify and rectify genetic disadvantage smacks of eugenics, which is politically suspect if not unthinkable.

If leveling up is not a feasible response to genetic enhancement, the alternative is to level down (14). The most straightforward approach would be to prevent anyone from obtaining genetic enhancement. A ban on genetic enhancements could be aimed at a variety of targets. For example, purchasing or possessing enhancements could be made illegal, similar to laws punishing illegal drug use or rules prohibiting the use of performance-enhancing drugs in sports. Another target would be health care professionals and institutional providers such as hospitals and IVF clinics. Congress or state legislators could make it a crime for health care professionals to provide genetic enhancements. Violators would face disciplinary actions by state medical boards, including loss or suspension of their licenses (39). Hospitals and other facilities like IVF clinics that continue to offer enhancement services would lose their licenses, their accreditation, or their ability to receive reimbursements under Medicare and Medicaid. Finally, if genetic enhancements were proprietary products such as drugs or medical devices, the Food and Drug Administration (FDA) could deny marketing approval.

But why wait until enhancements are available before banning their use? Why not prohibit research aimed at developing enhancement technologies in the first place? An analogy is the federal government's ban on federal funding of research on embryos and fetuses (40). Privately funded research could be restricted by penalizing institutions such as hospitals that participate in clinical trials, and by FDA denying permission to ship experimental enhancement products across state lines for purposes of human testing (41).

All of these restrictive approaches have limitations, however. Penalizing people who genetically enhanced their children would trigger intense constitutional debate. Particularly in the case of passive enhancements involving traditional "coital" methods of reproduction, the Supreme Court is likely to apply a strict scrutiny standard under which the right to decide what type of child to conceive or bring to term can be overridden only by a compelling state interest, and then only if the state uses the least intrusive means of regulation. Genetic enhancement accompanying less traditional methods of reproduction, such as IVF, may be entitled to less constitutional protection. However, even then, the courts are likely to take a hard look at overly intrusive government regulation. Somatic self-enhancement, while not raising issues of reproductive freedom, would set the state's interest in promoting equality against the individual's constitutionally protected interest in personal liberty and autonomy, including the right to make life-style decisions that do not harm others.

Those FDA restrictions on the sale of enhancement drug products, biologics, or devices able to survive constitutional challenge as an appropriate regulation of interstate commerce would be hampered by the way in which they are likely to become commercially available: as unapproved or "off-label" uses of products approved for therapeutic rather than enhancement purposes. A genetically engineered drug that enhances cognition, for example, could be approved to treat cognitive impairment, such as the effects of Alzheimer's disease. After it is approved for a therapeutic purpose, people might begin to seek it for unapproved enhancement purposes. The experience with human growth hormone mentioned earlier is a prime example. This genetically engineered drug is approved for use in children with "a lack of adequate endogenous growth hormone secretion," causing short stature (42). Yet parents are reported to be asking doctors to prescribe it for children who are merely short, and there are anecdotal accounts of parents requesting the drug for children who are already tall, in order to enhance their chances of playing competitive basketball (43).

FDA does not effectively regulate off-label uses of unapproved drugs. It merely limits the ways in which the manufacture may promote the drug for unapproved uses. Even if the FDA attempted to prohibit manufacturers altogether from promoting drugs for an unapproved use, enhancement uses would become public knowledge through media reports, the Internet, and word of mouth. Targeting health professionals who provided enhancement products to their patients would present similar obstacles. The FDA presently has no authority to control the prescribing behavior of physicians, who are free to prescribe products for uses which are not approved (44). There is nothing unlawful about a physician prescribing human growth hormone for children for an enhancement purpose which is not indicated on the product labeling. The only effective action the agency can take now is to ban or limit sales of the product altogether — for both therapeutic and enhancement uses. Yet in the case of products approved to treat serious and especially popularized diseases, this would carry an intolerably high political price.

The same problem would beset efforts to prevent research on genetic enhancements from taking place. Consider a ban on research on genetically engineered drugs to enhance cognitive function. Such a ban would be justified, it might be argued, on the ground that developing such a product would give those who used them an unfair advantage in competitions for scarce resources like college acceptances or aptitude-based job slots. But these same products most likely would be useful in treating diseases of cognitive deficiency, such as Alzheimer's and dementia. It is extremely difficult to curtail research on a specific use of a product. In any event there is little point, since, as stated above, once the product is developed for therapeutic use, it can easily migrate to enhancement uses.

Moreover an effective ban on access to genetic enhancements, whether aimed at individuals obtaining them for themselves or their children, or at providers and manufacturers, would require an elaborate enforcement regime. The analogies that best describe what would be necessary are programs to control the use of performance enhancing drugs in sports and the use of illicit recreational drugs. Indeed, the most appropriate government agency for regulating genetic enhancements may not be FDA but rather the Drug Enforcement Agency (DEA). After all, DEA, pursuant to the Controlled Substances Act,

is responsible for enforcing restrictions on access to physiologically active products stemming from societal objections to their use.

Like the war on drugs and the effort to ban drugs in sports, restricting access to genetic enhancements to promote equality is bound to be extremely intrusive and expensive. These precedents were not completely effective. Somatic enhancements in the form of drugs, although perhaps complicated to manufacture, may be easy to conceal. Even enhancements that depended on sophisticated medical procedures such as IVF might be procured if one "knew the right person," the way "back-alley" abortions could be obtained prior to *Roe v. Wade* (45). The overwhelming consumer demand for genetic enhancements is certain to spawn a robust black market. As the experience with abortions indicates, people who are prevented from obtaining genetic enhancements domestically simply will procure them abroad (46).

The most troublesome aspect of enforcing a ban on genetic enhancements is likely to be the difficulty of determining that someone has been illegally enhanced so that they, and/or the person who enhanced them, can be punished. In part, this is a technical problem of detecting the presence of enhancement products or enhanced DNA in the human body. A similar problem plagues attempts to ban performance-enhancing drugs in sports. Athletes and their coaches are becoming increasingly adept at deceiving drug-screening tests. The athletes may use substances such as erythropoeitin that are naturally occurring in the body so that the exogenous enhancement cannot be chemically distinguished (47). Furthermore the athlete may be able to use an enhancement substance to produce a benefit, such as increased muscle mass, and then stop taking the substance sufficiently in advance of a screening test so that its use cannot be detected.

In the case of genetic enhancements, the enforcement problem would be compounded by the difficulty of distinguishing between therapeutic and enhancement uses. As noted earlier, the difference between the two often is not clear. Someone could claim, for example, that an improvement in appearance was necessary to treat feelings of inadequacy, or that an increase in strength or dexterity was preventive therapy in that it enabled them to avoid injury. Furthermore, as has been noted, many genetic enhancements have lawful medical uses. Someone could take human growth hormone in an attempt to become tall enough to play professional basketball, but someone else could take the exact same substance to combat pituitary dwarfism. A ban on enhancements would require a complicated system for distinguishing between legitimate and prohibited activity involving the same products. In addition, banning genetic enhancement in conjunction with assisted reproductive technologies such as those delivered in IVF clinics would require a far more effective scheme of regulatory regime than is currently in place (48).

Yet the strongest objection to banning genetic enhancements has not been mentioned so far: Enhanced individuals not only may personally benefit from their advantages, but they may confer advantages on society. For example, a person whose science ability was enhanced (assuming that this collection of traits was amenable to genetic manipulation) might make important discoveries that would be impossible, or take much longer, for an unenhanced scientist. As noted earlier, even proponents of equality recognize the need to permit a certain degree of inequality in order to increase social benefit. In short, we might want to permit an individual to be enhanced if we expected the ratio of societal to personal benefit to be favorable enough.

Together with the practical limits on the effectiveness of a complete ban on genetic enhancements discussed earlier, the social value of certain kinds of enhanced performance make the goal of a total ban both unrealistic and undesirable. Some people will manage to enhance themselves no matter what it takes, and in some cases we will want people to do so. This leads to several policy suggestions: (*1*) enhancement licensing, (*2*) establising an enhancement lottery, and (*3*) regulating germ-line enhancement.

Enhancement Licensing

In order to permit genetic enhancements to produce desirable social gains, as well as to take some of the pressure off of a regulatory embargo that attempted to prevent the wealthy from purchasing enhancements, we should institute a system for licensing individuals to obtain genetic enhancements on the condition that they employ their enhanced abilities in some predefined manner to benefit society. By reducing the number of people who were enhanced, a licensing program would reduce the degree of social inequality, and the threat that genetic inequality poses to democratic institutions.

The system would be similar to legally enforced professional licensing schemes that give their holders powers and privileges denied ordinary citizens in return for agreements to abide by rules designed to promote social goals and to refrain from behaving in socially undesirable ways. The system also would bear some resemblance to licensing ownership or use of dangerous products such as handguns or automobiles. The administrative costs could be financed by licensing fees.

Such a licensing requirement could be enforced in the first instance against providers of genetic enhancement products or services. They themselves would be required to be licensed as a supplier, which would carry with it restrictions and reporting requirements. (A similar program operates under the Controlled Substances Act to keep track of the prescribing of narcotics and other dangerous drugs.) Individuals who seek to purchase enhancements would apply to a licensing board and would be required to propose the socially desirable purposes for which they seek to be enhanced. Those whose applications are approved would report to the board periodically to provide assurance of satisfactory performance, and their reports would be carefully audited. Licensed enhancements that involved manipulation of DNA would be genetically "tagged" so that lawfully enhanced individuals could be distinguished from those who obtained enhancements on the black market (49,50). Failure to fulfill the terms of the license would be penalized by loss of access to the enhancement or to its benefits. Depending on the nature of the enhancement, the penalty could take the form of being deprived of supplies of

the enhancement product, actual biological reversal of the enhancement, various forms of social handicapping, surtaxes or monetary penalties, and perhaps in cases of egregious violations, such as the use of enhancements to cause serious harm to others, imprisonment. Similar penalties would be imposed on persons who were discovered to have supplied or obtained enhancements without being licensed.

Establishing an Enhancement Lottery

The licensing scheme so far described would be open only to the wealthy, since they would be the only persons who could afford to purchase genetic enhancements. This would perpetuate the inequalities described earlier that would result from wealth-based access to enhancements described earlier. The solution would seem to be to provide some people with access to enhancements even if they could not afford it. One approach would be a government program that subsidizes enhancements for certain persons, perhaps those who, in return for their license, promise to provide the most desirable set of social benefits. But this would embroil the government in an enhancement-rationing program in which it was required to judge the relative merit of different proposals, a task that would raise objections similar to those that have been lodged in the past against health care rationing programs in general (51). On the other hand, such a licensing plan does not raise similar objections because the licensing authority would not compare individuals seeking enhancement but would allow anyone to purchase enhancements so long as they agreed to meet certain minimum social objectives.

A better solution than a rationing program would be to establish a national lottery for genetic enhancements (46). Everyone would be given one chance in each drawing. The winner or winners would be entitled to resources sufficient to enable them to purchase the maximum package of enhancements lawfully available in the private market, although in order to "cash in" their winnings, they would have to become licensed like everyone else who was enhanced. Like the licensing program itself, the lottery could be financed by license fees paid by those who purchased enhancements. Among the advantages of a lottery approach is that its randomness would give continued vitality to the concept of equality of opportunity.

Regulating Germ Line Enhancement

The greatest threat to social equality posed by genetic enhancements is the formation of a genobility — a class of related individuals who achieve and maintain an unassailable grip on wealth, power, and social privilege and who pass their advantages on to successive generations. As discussed earlier, a genetic aristocracy of this sort is antithetical to liberal democratic political systems. If genetic enhancements are obtainable at all, then to some extent the formation of an enhanced class cannot be prevented; persons who were wealthy enough to purchase enhancements presumably would be able to provide their children with greater material advantages than persons who were not enhanced, thereby making it more likely that these children would be able to purchase genetic enhancements in their turn. Yet the formation of such a genobility is far more likely to occur if individuals were permitted to make enhancement changes in their germ lines. Their offspring would inherit these genetic advantages, which they would be able to supplement with additional germ-line enhancements that they purchased, which in turn would be passed on to their children, and so on (52).

The social threat created by the inequality that would result from germ-line enhancement may not readily be mitigated by the licensing requirement that would accompany the lawful acquisition of somatic enhancements. It is difficult to imagine how to ensure that a person's children would abide by the licensing conditions agreed to by their parents. The children could be required to become licensed in their turn (e.g., when they reached the age of majority) on penalty of forfeiting their enhancement advantages, but despite the stipulation that the enhanced individual devote some degree of his or her enhanced capabilities to the public good, being licensed at the age of majority may not be sufficient to counteract the inequality that germ-line enhancements would produce.

The solution then would seem to be to prohibit germ-line genetic enhancement altogether. Conceivably the threat in social equality could be met by banning only those forms of germ-line enhancement involving gene transfer, and not the passive sorts of germ-line enhancement that would occur with genetic selection for enhancement, selective abortion for enhancement, or preconception enhancement. Moreover laws that make it illegal for individuals for enhancement reasons to discover their genetic endowment, to select embryos for implantation, or to abort a fetus, might be more realistic politically than laws that prohibit the alteration of germ cells for enhancement purposes.

A ban on germ-line engineering would raise a host of problems. It might be challenged as an unconstitutional interference with procreative liberty, although the justification that it was necessary in order to preserve democratic liberties from being engulfed by a genetic aristocracy might be deemed a compelling state interest. Detecting when someone had altered germ cells would be difficult and intrusive (53). Nevertheless, a ban may be necessary to promote a minimum level of genetic equality.

UNFAIRNESS

Regardless of the manner in which we attempt to reduce the inequalities that may be created by wealth-based genetic enhancement, some people invariably will become enhanced. A licensing scheme that is vigorously and effectively enforced will go some distance toward offsetting the advantages enjoyed by enhanced persons but not far enough. Enhanced individuals still will be in a superior position compared to unenhanced persons. This raises the question of whether and in what ways society should respond in order to reduce the resulting unfairness.

This unfairness will be felt most acutely when the unenhanced compete with the enhanced for scarce societal resources or when an enhanced individual exerts power over an unenhanced person in a relationship in

which their interests conflict. These circumstances can occur in a large number of settings: between rivals for someone's affection or in interpersonal relationships such as those between boyfriend and girlfriend (and similar relationships between members of the same sex); in contests, including sports, games, beauty pageants, and talent shows; in competition for access to limited privileges, such as admission to academic institutions; in fiduciary relationships, such as those between patients and health care professionals, trustees and beneficiaries, directors and shareholders, attorneys and clients, and insurers and insureds; and in ordinary business relationships, such as those between seller and buyer, landlord and tenant, realtor and purchaser, lender and debtor, manufacturer and consumer, and employer and employee. The object of the competition may be any desirable good: money, jobs, status, affection, sexual favors, political influence, or market power. The relative advantage conferred by genetic enhancement would depend on both the context and the nature of the enhancement: In a test of strength, for example, enhanced intelligence may be of little value. Unfairness could arise either in a zero-sum situation in which the enhanced person obtains benefit at the unenhanced person's expense, or in non-zero-sum situations in which, although both the enhanced and the unenhanced person gain, the share gained by the enhanced person is greater and the share gained by unenhanced person smaller than would be the case if the parties were equivalently advantaged.

All these situations are subject in some fashion to external rules of behavior. They may be formal public laws; legally enforceable private law, such as the bylaws and other governing principles adopted by corporations, partnerships and unincorporated associations; or social norms or customs. How should these systems of rules respond to the potential unfairness created by genetic enhancement? Should the rules treat these differences as if they do not exist or do not bear on the activity? Or should the rules attempt, in some fashion, to level the playing field? If one person possesses an advantage over another, should the rules permit her to profit from it at the other's expense? Although it would be fascinating to consider nonlegal responses based on social norms and customs, the focus of the rest of this article will be on legally enforceable rules, that is, on public and private law.

If we attempt to level the genetic playing field in response to genetic enhancement, what would our options be? Basically there are the same two approaches that we examined in the previous section in discussing how to reduce genetic inequality: Either we decrease the advantages of those persons who were genetically enhanced, or we improve the lot of those who were not. In short, once again, we can level "up" or "down."

Leveling up would entail giving those who were not genetically enhanced some countervailing benefit. This could be money, professional advice, information that was hard to come by, or any other desirable resource that would help level the playing field. It could be a preference in access to a scarce resource, such as an affirmative action program. Yet, it is difficult to conceive of how this approach would work in the context of personal interactions. Would an unadvantaged person be permitted to draw on some public store of resources to place her on the same level as the enhanced person? Obviously this could not be a store of genetic enhancements, since that would contradict the basic assumption that we cannot afford to provide genetic enhancement to everybody. Yet, the same problem of scarce resources would plague any other subsidy, monetary or otherwise: It would cost too much to put the unadvantaged on the same level as the enhanced.

A less expensive alternative might be to level up only those who were the most disadvantaged relative to the enhanced. This would resemble laws prohibiting employment discrimination against persons with disabilities (54). The effect of these laws is to require employers to subsidize persons with disabilities so that in competitions for employment they can match persons who are not disabled. Only disadvantaged employees or applicants for employment receive this benefit, thus leveling the employment playing field.

This approach is intuitively appealing. By focusing on improving the lot of the worst off, it moves in the same basic direction as Rawls's difference principle. Yet it would produce odd results if it were applied to a more realistically complex society in which some people are enhanced, some (the unenhanced) are merely "normal," and some are disadvantaged: If through access to countervailing benefits, the disadvantaged are truly brought up to the level of the enhanced, they would pass those who previously had been neither advantaged nor disadvantaged. The formerly unadvantaged now would constitute the disadvantaged. In short, unless everyone is made equal, or the distribution of countervailing benefits is a once-only event, a policy of benefiting the worst-off would create an infinite regression. There will always be a group that is disadvantaged or unadvantaged and that riskes being treated unfairly by the enhanced — and also, under a genetic "affirmative action program," by the formerly disadvantaged who have been leveled up.

This leaves the other option of "leveling down." Since we cannot prevent some people from obtaining genetic enhancements for themselves or their children, unfairness might be avoided by preventing them from taking advantage of their enhancements when competing with the unenhanced or exerting power over them.

Some idea of the ways in which we might level down the genetic playing field can be obtained by reviewing how rules currently respond to the potential unfairness inherent in interactions between advantaged and unadvantaged individuals. Here, instead of advantages conferred by genetic enhancement, society is concerned with natural or acquired advantages such as youth, beauty, size, strength, endurance, intelligence, memory, creativity, information and knowledge, experience, social status, money, and personal power. If we examine current public and private law rules, we see a number of ways in which they attempt to level the playing field by leveling down these sorts of advantages:

1. Competition that is arguably unfair is sometimes prohibited. A private law example is the ban on the use of performance-enhancing drugs in sports

competitions (55). Another sports example is weight classes in certain competitions such as rowing and wrestling. In these competitions, athletes who have an advantage in weight are precluded from competing with those who weigh less.

Banning competitions between advantaged and unadvantaged individuals is not confined to sports. A public law example is the prohibition against insider trading in securities. Here the advantage is information that is not available to the public about a corporation whose stock is publicly traded. The law attempts to deny those who possess this information any financial gain from it. The advantaged individual is given the choice of either disclosing the information or not trading stock in the company.

2. The rules permit a transaction to take place only if the person with the advantage forfeits it by sharing it with the unadvantaged. An obvious example is information possessed by one party in certain business transactions, such as when the advantaged person knows that "disclosure of the fact would correct a mistake of the other party as to a basic assumption on which that party is making the contract" and nondisclosure would be a failure to act in good faith and "with reasonable standards of fair dealing" (56). Presumably such transactions are not prohibited altogether because it is sufficiently inexpensive to enforce the forfeiture rule and there is a sufficiently high possibility that, given adequate enforcement of the rule, the result will be fair.

   Similarly in some cases the person with the advantage is handicapped so that the advantage is removed. This occurs, for example, in horse racing where jockeys who weigh relatively little are deprived of their advantage literally by having to carry weights. Better golfers are also deprived of their advantage by removing strokes from the score of other golfers.

3. The rules do not prohibit the competition but allow the unadvantaged to avoid the outcome if it seems too unfair. The doctrine of unconscionability in contracts is such a rule, which applies to advantages in the form of information or market power (57). Another example is the fiduciary rules that permit a court to a void a transaction by a trustee of a trust if the result would be unfair to the beneficiaries (58).

4. The rules sometimes level the playing field by eliminating the arm's-length nature of the transaction. The advantaged person is permitted to engage in the transaction but not allowed to employ the advantage in such a way as to take advantage of the other party. This is the result, for example, of fiduciary rules that mandate the fiduciary's undivided loyalty towards the entrustor and prevent the fiduciary from serving an interest other than the beneficiary's (59).

On the other hand, the rules could make no effort to level the playing field, and legislators could ignore or even to celebrate the advantages that some people have over others. With the exception of affirmative action programs, for example, admissions criteria at selective educational institutions do not adjust applicants' accomplishments in light of their background or abilities. A person applying to Harvard with an IQ of 120 competes with applicants with IQ's of 160; the fact that an A in AP Calculus or a high score on the Scholastic Aptitude Test (SAT) achieved by the person with the 120 IQ is a far greater accomplishment than the same grade achieved by the person with the 160 IQ is irrelevant. Many athletic competitions force athletes to compete with those who are advantaged by being younger: older baseball and basketball players must compete with those considerably younger, some straight out of high school. Shorter basketball players are not allowed to shoot from stepladders, and there are no professional leagues for players of "normal" height. In football the slight take the field at their own peril.

This raises the question of whether the unfairness problem raised by wealth-based genetic enhancement simply should be ignored, as it seems to be in the case of college entrance criteria and in certain sports settings. What would justify ignoring the problem?

In many cases the fact that the rules ignore advantages, or certain advantages, is probably arbitrary, coincidental, or an historical artifact of no theoretical significance. In horse racing, for example, jockeys' weights are equalized on the premise that it is the quality of the horse and the jockey's horsemanship that should matter. There are no weight categories in football because that is just not how the game was conceived. Organized chess competition does not prohibit the use of cognitive enhancers such as nicotine or stimulants because the organizers simply never thought of it (E.C. Johnson, Assistant Director, U.S. Chess Federation, personal communication, October 4, 1995).

Nevertheless, we can posit several principled reasons why it may be inappropriate to deprive genetically enhanced individuals of their advantage in specific transactions or relationships: (*1*) loss of societal benefit from the enhancement, (*2*) difficulty in detecting enhancement, (*3*) difficulty in distinguishing between enhancement and effort, (*4*) nonenhancement advantages, (*5*) public intrusion into private affaires, and (*6*) transaction costs.

Preventing the Loss of Societal Benefit from the Transaction

An enhanced scientist, presumably enjoys personal advantages by virtue of being enhanced; she otherwise might not have been admitted to MIT, for example, or be able to earn a fortune from her patents. But despite the unfairness to unenhanced persons who applied to MIT or tried to develop patentable inventions, we might forgo trying to strip her of her personal benefits. By allowing her to benefit personally, we encourage people like her to purchase scientific enhancements so that society could reap the benefits. (This might well be the justification for not leveling the playing field in terms of intelligence in the case of admissions to institutions of higher learning.)

An example of a societal benefit that might be a sufficient reason to permit enhanced individuals to retain personal benefit are reductions in the costs of accidents. A naturally talented automobile mechanic might be expected to make safer repairs than someone with less talent,

and therefore might be entitled to a hiring preference over someone who lacked her natural talents. The same might be said for an enhanced automobile mechanic. The argument becomes even more compelling in the case of persons responsible for the safety of large numbers of people: airline pilots, railroad engineers, operators of nuclear power plants, and the like.

Difficulty of Detecting Enhancement

Systems of rules might have no choice but to ignore the unfairness created by competitions involving enhanced and unenhanced individuals if enhancements cannot be detected. This is a severe problem in attempting to prohibit the use of performance-enhancing drugs in sports. The earlier discussion of licensing catalogued the difficulties of detecting enhancement and the potential solutions.

One further approach to the detection problem might be to permit unenhanced individuals to assert a rebuttable presumption against persons they interacted or competed with whom they believed were enhanced. Unless persons against whom the presumption was asserted could establish that they were not in fact enhanced, the rules would proceed to level the playing field (e.g., by having courts undo the deal, or penalizing the person presumed to be enhanced for participating in a prohibited competition). Inability to produce proof of lack of enhancement would satisfy the burden of proof that the person was enhanced. This would encourage the enhancement industry itself to develop a workable tracking and record-keeping system.

Distinguishing Between Enhancement and Effort

Arguably society should focus its leveling efforts on advantages derived from genetic enhancement, rather than on advantages obtained through personal effort. Otherwise, the effect will be to discourage effort, leading to sloth and loss of social benefit. As noted earlier, however, it may be difficult if not impossible to distinguish between achievements that are earned and achievements that result from enhancement. Accordingly, it might be argued, the rules ought to ignore genetic enhancement.

Yet genetic enhancement may be a sufficient social threat that it is appropriate to level playing fields regardless of this risk. Indeed, society often deprives people of plainly earned advantages in order to promote equality or fairness. Though not without its critics, progressive income taxation transfers earned wealth to achieve a more just distribution of resources. Weight categories in sports are enforced regardless of whether an athlete's size is the product of diet and exercise or steroids. Fiduciary law requires individuals with superior information to disclose it to beneficiaries, clients, and patients, even though the information may have been obtained through great effort. Similarly an enhanced person automatically might be required to disgorge her advantage (of information, market power, etc.), even though she obtained some or all of it through her own efforts. Alternatively, as mentioned before, the fact that someone was enhanced could establish a rebuttable presumption that any advantage related to the enhancement was due to the enhancement rather than to effort.

Nonenhancement Advantages

Just because someone is not genetically enhanced does not mean that they lack sufficient resources or talents to compete fairly with someone who is enhanced. The unenhanced person may possess great wealth, or have some special store of knowledge, or have access to the best advisors. Leveling the genetic playing field may exacerbate unfairness if it focuses on genetic enhancements to the exclusion of these other types of advantages. If society attempted to correct for all differences between people, however, there would be no end to societal interference.

On the other hand, as discussed earlier, the advantages conferred by genetic enhancement could be so great that it would be appropriate to single them out for remediation. Moreover, where enhancement merely creates a rebuttable presumption of unfairness, the enhanced individual would be free to prove that her enhancement advantages are equaled or outweighed by nonenhancement advantages possessed by the complaining party.

Intrusiveness

Given the difficulties of detection and differentiation described above, any attempt to level the playing field would invite public intrusion into highly personal affairs. To rebut an inference of enhancement, for example, people would have to reveal their personal and medical history, including particularly sensitive information relating to their genetic makeup and their reproductive activities.

If the stakes are high enough, however, we seem to be willing to require people to compromise their privacy rights. For example, athletes must submit to physical examination and to yield samples of bodily fluids for testing, often under nearly public conditions. Given sufficient concern for maintaining privacy and the confidentiality of sensitive personal information, and so long as the least intrusive means were employed to decide if someone were enhanced, the cost may be justifiable.

Transaction Costs

Leveling the genetic playing field is liable to be costly. Forums, advocates, and referees would be required to resolve fairness disputes. Black markets, both domestic and foreign, would need to be policed. The specter of a "war on genes" is not an attractive one. Yet again, the threats posed by genetic enhancement might well be worth the cost of leveling.

In short, there seems to be no obvious reason why we would ignore the unfairness created by wealth-based genetic enhancement, except in situations like preventing accidents or achieving scientific breakthroughs in which the ratio of social to personal benefit clearly demonstrated a substantial net benefit to society, or in situations in which the costs of leveling were deemed to be greater than the costs of unfairness. In all other cases, one or more of the leveling techniques listed earlier would be appropriate, depending on the circumstances.

LEVELING THE GENETIC PLAYING FIELD

Although our overall objective is to minimize individual unfairness caused by interactions between enhanced and unenhanced individuals, at same time, we want to maximize the societal benefit from individual enhancements. This raises the question of how to respond when the two objectives are incompatible, namely when individual unfairness can be prevented only by sacrificing societal benefit.

The answer depends on the nature and magnitude of the unfairness and of the forgone societal benefit. Ultimately public policy should favor preventing unfairness if the cost of unfairness is deemed to exceed the expected societal benefit. Conversely, a substantial amount of societal benefit should be sought at the expense of a relatively small amount of individual unfairness. The more substantial the unfairness, and the more equal the costs of unfairness and societal benefits tend to be, the more emphasis should be given to correcting the unfairness of the transaction.

For example, suppose that we are reviewing applicants for scarce medical research funding. Successful research is expected to yield significant societal benefits. Applications are submitted by both unenhanced individuals and individuals enhanced in ways that significantly increase their chances of research success. All other things being equal, we ought to award funding to the enhanced individuals. If the impact on the unenhanced individuals' careers is deemed significant enough, however, consideration might be given to mitigating the unfairness, such as by allocating a certain amount of funding for them alone (a sort of "unenhanced persons' affirmative action program"), or by favoring applications involving both enhanced and unenhanced investigators.

As suggested by the approaches described above to leveling the playing fields, there are a number of techniques for mitigating genetic unfairness. Some of these techniques are more costly than others, both in terms of implementation costs and in terms of forgone social benefit. The enhanced individual might be required to share with the unadvantaged person the advantage created by enhancement. At a minimum, enhanced individuals would have to disclose that they were enhanced. In a business transaction an enhanced party who by virtue of their enhancement has obtained superior information could be required to disclose that information to the unenhanced party. Sharing might be a preferred mitigation technique where a transaction is expected to yield societal benefit because it encourages the enhanced party to engage in the transaction by allowing that party some degree of personal benefit.

If sharing is impractical, such as in zero-sum situations, or if the implementation costs of sharing are too great compared with the expected societal gain, unfairness might be mitigated instead by handicapping the enhanced party. For instance, in contests, including athletic competitions, the enhanced individual could be put at a disadvantage, such as being given a longer distance or a harder question.

Another technique worth considering is allowing the interaction to take place but permitting the unadvantaged party to apply to a court or an administrative agency to challenge and overturn or adjust a result if it is too unfair. This flexible, posthoc approach might be appropriate where the unfairness costs and societal benefits of a transaction were difficult to predict in advance. Business deals might be candidates for this approach, for example, particularly those in which the advantages enjoyed by the enhanced party, such as market power, could not, like information, be shared, and in which the particularities of transactions made the application of a priori handicapping rules too inexact. Making outcomes voidable also saves the costs of intervening in every transaction; only those results that seem too unfair will be reviewed.

An interesting option is to eliminate the arm's length nature of transactions between enhanced and unenhanced individuals. Like fiduciaries, the enhanced would be made responsible for the welfare of the unenhanced, a sort of genetic noblesse oblige. As in true fiduciary relationships, this could decrease the costs of monitoring the behavior of the enhanced by substituting a system of sanctioned trust for a regime of direct external controls. It also would encourage the unenhanced to interact for their benefit with the enhanced, facilitating resulting societal benefits (60).

Finally, if no significant social benefit were expected from an interaction, it could be prohibited. An alternative to handicapping enhanced athletes, for example, would be to forbid them from competing against athletes who were unenhanced. Such a competition might be allowed only if the costs of enforcing such a prohibition were great compared to the unfairness.

BIBLIOGRAPHY

1. C. Tucker, *The Weekly Standard*, December 2, p. 20 (1996); J. Ketelsen, *Forbes*, May 9, p. 184 (1994).
2. 42 USC §1395y(a)(1)(A).
3. 42 USC §1395y(9)(10).
4. 42 C.F.R. §440.230(d).
5. U.S. Department of Health and Human Services, *Health United States 1995*, Table 3, USDHHS, Washington, DC, 1995, p. 82.
6. L. Cuttler et al., *JAMA, J. Am. Med. Assoc.* **276**, 531, 532 (1996).
7. *U.S. News & World Rep.*, May 13, p. 20 (1996).
8. R.J. Paulson, *West. J. Med.*, December 1, p. 377 (1996).
9. D.R. Lamb, *Am. J. Sports Med.* **12**, 31–34 (1984); N.A. Ghaphery, *Sports Med.* **26**, 433 (1995); D.A. Smith and P.J. Perry, *Ann. Pharmocother.* **26**, 653–658 (1992).
10. G. Cowley and A. Underwood, *Newsweek*, June 15, p. 54 (1998).
11. P.J. Whitehouse, E. Juengst, M.J. Melhman, and T.H. Murray, *Hastings Cent. Rep.* **27**(3), 14–22 (1997).
12. *Orlando Sentinel*, October 12, p. E2 (1991).
13. E. Parens, *Hastings Cent. Rep.*, January, February, p. S8 (1998).
14. M. Walzer, in M. Robertson, ed., *Spheres of Justice*, Oxford University press, Oxford, UK, 1983, p. 11.
15. D.S. Diekema, *Perspect. Biol. Med.* **34**(1), 109, 112, 113 (1990); T.H. Murray, *The Worth of a Child*, University of California Press, Berkeley, 1996, pp. 90–91.

16. E. Parens, *Hastings Cent. Rep.* **28**(1), S1, S11 (1998); G. McGee, *ibid.* **27**, 16 (1997).
17. J.W. Gardner, *Excellence: Can We Be Equal and Excellent Too?* Norton, New York, 1984, p. 30.
18. T. Nagel, *Equality and Partiality*, Oxford University Press, New York, 1991, p. 135.
19. R. Nozick, *Anarchy State and Utopia*, Basic Books, New York, 1974, pp. 223–224.
20. Id. at 347, n.41.
21. J. Rawls, *A Theory of Justice*, Belknap Press of Harvard University Press, Cambridge, MA, 1971, p. 104.
22. R. Dworkin, *A Matter of Principle*, Harvard University Press, Cambridge, MA, 1985, p. 207.
23. R. Dworkin, *Philos. Public Affairs* **10**, 185, 283 (1981).
24. J.E. Roemer, *Theories of Distributive Justice*, Harvard University Press, Cambridge, MA, 1996, pp. 237–252.
25. H. Frankfurt, *Ethics* **98**, 21 (1987).
26. R. Arneson, *Philos. Stud.* **56**, 83, 84 (1989); M. Rosenfeld, *Cal. K. Rev.* **74**, 1687–1699 (1986).
27. E. Rakowski, *Equal Justice*, Oxford University Press, New York, 1991, p. 159.
28. E. Rakowski, *Tax Law Rev.* **51**, 419 (1996).
29. J. Riedinger, *Cap. Law. Rev.* **22**, 893, 895–897 (1993).
30. F. Parkin, *Class Inequality and Political Order: Social Stratification in Capitalist and Communist Societies*, Praeger, New York, 1971.
31. D.B. Grusky and A.A. Takata, in E.F. Borgatta and M.L. Borgatta, eds., *Encyclopedia of Sociology*, Macmillan, New York, 1992.
32. J.H. Schaar, *Legitimacy in the Modern State*, Transaction Books, New Brunswick, NJ, 1981, p. 195.
33. G.A. Hedger, ed., *An Introduction to Western Civilization*, Odyssey Press, New York, 1949, p. 217.
34. A. Bonner et al., *Democracy in India: A Hollow Shell*, American University Press, Washington, DC, 1994.
35. M. Mehlman, M. Durchslag, and D. Neuhauser, *J. Health Politics, Policy and Law* **22**, 1385–1391 (1997).
36. A.M. Capron, *Hastings Cent. Rep.*, November 22, pp. 18–20 (1992).
37. D.C. Hadorn, *JAMA, J. Am. Med. Assoc.* **268**, 1454–1459.
38. D. Orentlicher, *JAMA, J. Am. Med. Assoc.* **271**, 308–314.
39. AMA Council on Judicial and Ethical Affairs, *Arch. Fam. Med.* **3**, 633 (1994).
40. R.M. Cook-Deegan, *Cloning Human Beings*, vol. II, National Bioethics Advisory Commission, Washington, DC, 1997, p. H8.
41. R.A. Merrill, *Va. Law. Rev.* **82**, 1753, 1777–1782, 1821 (1996).
42. *Physician's Desk Reference*, 1998, p. 993.
43. R. Rubin, *Dallas Morning News*, July 7, p. A1 (1986).
44. 21 U.S.C.A. §396 (Supp. 1998).
45. Z. Leavy and J. Kummer, *S. Calif. Law. Rev.* **35**, 123 (1962).
46. M.J. Mehlman and J.R. Botkin, *Access to the Genome: The Challenge to Equality*, Georgetown University Press, Washington, DC, 1998, p. 119.
47. J. Maher, *Austin Am.-Statesman*, July 26, p. E3 (1992).
48. Fertility Clinic Success Rate and Certification Act of 1992, 42 U.S.C. §263a-1 to 263a-7 (1994).
49. M. Pollan, *N.Y. Times Mag.* October 25 (1998).
50. R.F. Service, *Science* **474** (1997).
51. M.J. Mehlman, *Wis. Law. Rev.* **239**, 256–260 (1985).
52. D.S. Karjala, *Jurimetrics* **32**, 121, 136 (1992).
53. R. Cooke, *Newsday*, August 30, p. A06 (1998).
54. Rehabilitation Act of 1973, 29 U.S.C. §706. Americans with Disabilities Act, 42 U.S.C. §12102.
55. U.S. Olympic Committee, *Drug Control Education*. Available at: <*http://www.olympic-usa.org/inside/in_1_3_7_1.html*>
56. Restatement (Second) of Contracts §161(d).
57. Restatement (Second) of Contracts §153 (1981), §208.
58. A. Scott and W. Fratcher, *The Law of Trusts*, 4th ed., Little, Brown, Boston, MA, 1987.
59. S.P. Shapiro, *Am. J. Soc.* **93**, 623 (1987); M.J. Mehlman, *Conn. Law. Rev.* **25**, 349, 368–371 (1993).

See other entries BEHAVIORAL GENETICS, HUMAN; see also HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY entries.

HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY, POLICY, TECHNOLOGICAL ENHANCEMENT, AND HUMAN EQUALITY

MICHAEL H. SHAPIRO
University of Southern California Law School
Los Angeles, California

OUTLINE

INTRODUCTION

This article outlines some of the moral, legal, and general policy difficulties societies and individuals will face if technological enhancements via germ-line and somatic mechanisms become possible (1). It identifies and analyzes some of the conceptual structures necessary to explain the nature of these difficulties, suggests some alternative basic scenarios — such as greater or lesser scarcity of technological enhancement resources, impacts on how we perceive each other, different remediation patterns — and then maps and reverse-maps the projected technological developments against the value and legal structures. It also describes and comments on what many see as the most critical threats and promises, from our present value standpoints, of the anticipated changes, as well as on what might be the fate of these very standpoints themselves. The idea of enhancement is compared to other processes of human change, principally to the familiar forms of self-progress and the practices of treating disorders, injuries, and the like. Questions are raised about the very significance of these distinctions as rational *authorizing* and *limiting* tools that might guide us in distinguishing among permissible and impermissible interventions, and among obligatory and nonobligatory ones.

The moral and legal issues are explored primarily by way of the concept of equality. Some of the classic and (possibly) novel difficulties in the equality analyses include matters of access to and distribution of technological resources; the possibility of increased socioeconomic and political stratification that may be irreversible; the effects of a technological enhancement regime on the ways in which we view each other (as planned and assembled objects or as persons? some blend of these or other attitudes?) and on the viability of present views about equality and its relationship to justice, fairness, autonomy, utility, and ideas of merit, virtue, and desert. Considered in the discussion of our notions of merit, virtue, and desert is whether they are to be reconstructed or abandoned. Distributional criteria and, more generally, different egalitarian arguments based on different visions of equality are sorted, and there is a brief exploration of how the structure of our democratic institutions might be altered by responses to particular distribution patterns of merit attribute enhancements — in particular, by adopting a plural voting system of the sort envisioned by John Stuart Mill. Different forms of remediation or prevention of inequality and of affirmative promotion of equality are briefly touched on. At the end there is a brief review of some issues arising under the United States Constitution: If technology changes as many anticipate, the acute moral and policy issues will eventually be vetted and disputed within the legal system.

MEANINGS OF ENHANCEMENT

Mapping Meanings of "Equality" onto Meanings of "Enhancement," and Vice Versa

To explain how equality may be affected by technological enhancement requires some account of the meaning of "enhancement"; a review of the intimidating complexities of the equality analysis, including equality's relationship to other basic values; and an examination of how varying understandings of the concepts of enhancement and equality affect each other.

This article outlines a bi-directional mapping of differing versions of equality and projected forms of enhancement against each other. Different ideas of equality may lead to different valuations of enhancement, and the reverse. An initial task is to distinguish different equality arguments, and this rests, in part, on asking the now-familiar question, *What* is supposed to be equal to *what?* An obvious example of variant meanings of equality is suggested by the tension between equality of opportunity in its several forms (2) (certain ex ante positions are to be set equal) against equal-outcome standards (certain ex post positions are to be set equal). These opposing pulls are especially vivid when considering the possibility of, say, major enhancement of intellectual abilities. Equal opportunity understood as rights against interference by others with access to enhancement resources may yield unequal outcomes that track and intensify existing inequalities in wealth, income, status, and power. Diminishing returns in the value of increments in ability may set in slowly, thus prolonging the incentive to continue adding increments to intellectual talents, further deepening inequalities in power and social status. As a given form of enhancement becomes widespread, its value to a particular individual may shift from enabling her to tower over others to enabling her to avoid being towered over. For some traits, then, the more widespread the enhancement, the more urgently the less able need it in order to avoid losing more and more ground to more and more persons. At some later stage, relative interpersonal positions may be unchanged, although "absolute" performance capacities are amplified. An equality of outcome standard, on the other hand, would require major centralized intervention either to narrow the ability gaps among persons, or at least to preserve their relative standing. In the latter case, equality of outcome would encompass — not flat-out equal abilities — but preserving the status quo ante concerning the relative "distance" between persons. Of course, egalitarian maneuvers might involve redistribution of traditional goods and services, either in addition to or instead of enhancement opportunities (3).

Technological Expectations; Germ-Line and Somatic Enhancements

Current directions in technology clearly justify assuming for argument's sake that we will be able to influence significantly the development of our targeted traits as compared with what life's lottery (genetic or environmental) might otherwise have presented. The apparently successful germ-line alteration that resulted in superior learning

ability in mice illustrates the point nicely (4). In earlier experiments, mouse embryos assimilated rat genes coding for growth hormone, producing some large mice that themselves bred several hefty offspring. The important point to take from these results is that even complex polygenic and multifactorial traits such as intellectual ability or size may be heavily influenced by a given gene (5): Not all genes and environmental factors are equal — some may have outsize or disproportionate effects (6). The accompanying point, of course, is that similar outcomes in human beings may be quite far off, if they are possible at all.

One distinction requires immediate attention — that between *germ-line* and *non-germ-line* techniques for altering traits in a specific possible or existing person. (Selective breeding would alter the distribution of traits in a population but not in a particular individual.) The latter include genetic alteration of somatic (body) cells ("gene therapy"); such alteration does not affect gametes (mishaps aside) and so does not affect one's descendants. Nevertheless, because "gene therapy" and the development of substances directly affecting gene operations require extensive knowledge of genetic mechanisms, all these modes of enhancement will be mentioned here.

The Idea of Enhancement

In General. For convenience, "enhancement" and "augmentation" are used interchangeably, although the latter might also suggest "extension" or "supplementation" (e.g., a springier vaulting pole). Standing alone, the term "enhancement" will refer to technological enhancement, not to socially and legally accepted processes of self-improvement — say, gradually increasing one's strength by lifting weights, or improving analytical skills through study. Many observers view the results of such accepted measures as "internally" rather than "externally" or "unnaturally" generated changes that compromise claims of personal, meritorious achievement (7). In fact these varying paths toward superior traits will generally be intertwined, further complicating our analysis.

Here are three basic families of overlapping questions concerning the meaning of "enhancement." Who is to be enhanced and why? What is to be enhanced and why? What counts as enhancement and why?

What Is the "Unit of Enhancement"? We need to ask first *who* or *what* is to be enhanced. A possible person presently in the form of an early embryo, or as-yet-unjoined gametes? A particular living person or group? The present or future human race?

What Is Enhanced? Traits, attributes, and characteristics are the targets for enhancement, but to what do these terms refer? Behavior patterns? Physical appearance? Tissue structure? Molecular arrangements — such as genomic structure — and biochemical processes? Competitive performance? Predispositions to develop particular physical or mentational conditions? And which of these targets should be selected for improvement? What role does culture play in characterizing and valuing traits, and how might this track genetics? A culturally valued trait may rarely have a clear genomic correlate.

What Counts as Enhancement? The two most discussed issues are, first, whether and how to distinguish permissible forms of enhancement (practice, pumping iron) with impermissible forms (ingesting memory-aiding substances, altering the germ line of one's children — technological fixes or shortcuts generally); and, second, whether and how to distinguish between enhancement, on the one hand, and repair of "defects" or injuries or control of disorder, on the other. As to the latter, a major reason for insisting on the treatment/augmentation distinction is the belief that it provides limits on an enhancement imperative that might seem to be utterly unbounded. This is not without cost, however: We risk devaluing and stigmatizing those with "imperfections" of various sorts. In any case, resource constraints may impose severe limits on the extent to which the "treatment model" and the "enhancement model" will be "lumped."

Enhancement of Merit and Wealth-Attracting Attributes (Resource Attractors). Traits plainly vary in importance. Those strongly favored for whatever reason — special abilities, health, appearance, personality, culturally preferred predispositions — are critical variables affecting the distribution of life's rewards, including social and political status, income and wealth, praise, mating opportunities, and prizes. The moral and conceptual foundations of merit and desert judgments are complex (8) and cannot be plumbed here, but it is vital to see that "merit attributes" are often (but not necessarily) "resource attractors" — they are distributional criteria of sorts, whether in market, centrally directed, or other economic systems. The close coincidence of these ideas permits some interchangeable use. Suppose, now, that we can enhance these distributional criteria through technological alteration. Those persons already in a position to draw substantial resources may sharply augment their resource-attractiveness — possibly in a self-accelerating cycle that draws increasing wealth and power to the enhanced persons. In at least a metaphoric sense, then, one's very "merit" is increasing — one's "merit basis" is "stepped up" — thus amplifying one's claim for still more of everything, including still more merit. The resulting risks of increased and more inflexible social stratification are obvious (9). (The risks seem lessened when enhancement is temporary and must be repeatedly renewed — a point to retain throughout this discussion.) One existing parallel is the distribution of educational resources, particularly advanced and specialized higher education. Another is wealth itself: One needs it to get more of it, and even to keep what one has. Here the intersection of equality with justice and fairness considerations is obvious. (A close comparison and ranking of these values is not possible here.)

"Repair of Disorder" as Distinguished from "Enhancement"; Disorder versus Enhancement Models for Justifying Trait Alteration. If a pathological condition is successfully treated, we are unlikely to describe the restorative process or its result as "enhancement" unless the intervention appears to go beyond "canceling out" the disorder and induces a "nonnatural" condition that masks or displaces the impairment rather than restoring the patient's

ex ante personal baseline. If the improvement is justified by a supposed medical need, however, complaints of unequal distribution of enhancement resources may be blunted, although complaints about unequal access to medical resources may continue. In many cases the two models are partially "merged" — for example, measures to increase immunity or other resistance to certain disorders or conditions.

More generally, equality issues concerning enhancement will not just vanish when a disorder model is invoked. For one thing, reliance on a disorder model will not go unopposed. One can question the moral relevance of the distinction between disorder- and enhancement-based justifications for distribution: If an egalitarian imperative requires remedial measures, then — resource scarcity aside — what difference does the therapy-enhancement contrast make if one's relative position will be improved either way? One can also note the difficulties in distinguishing between the two in various other cases. For example, repair of a fracture might make one less vulnerable to future fractures, leaving one in effect stronger than before the injury. And, as we saw, germ-line or somatic manipulations that make one less vulnerable to, say, infectious diseases straddle the treatment/repair versus enhancement distinction. An oft-mentioned example of these difficulties is the use of human growth hormone. Its administration as treatment for short stature caused by pituitary or other disease is more readily accepted than its use on short persons not suffering from a height-impairing disorder — although from the short person's viewpoint, it may make little difference what accounts for his fate (10). The latter use is often seen as technological enhancement — although of a socially handicapping trait (11) — rather than treatment. Of course, whether something is a treatment at all rests on whether it is directed at a disorder, disease, injury and the like, and this may again depend in part on cultural habits and existing environmental conditions: What are socially (un)acceptable moods, or prevailing attitudes toward persons of very short/tall stature? Does the society's current "physical plant" (e.g., lots of stairs, few elevators) contribute to the limitations of persons with particular conditions? Still more, the treatment/enhancement distinction might be viewed as immaterial to the more general goal of "normalization" (12) — a concept overlapping, but distinct from, that of enhancement. But the standard of "normalization" may rise with enhancement or treatment, and the notion of relative handicap — or even disorder — may thus expand, with unfortunate consequences for the "handicapped."

Despite problems with the systematically murky treatment/enhancement distinction, however, it is far from meaningless (12).

Enhancement, Illicit Transformations, and Compromise of Identity, Merit, and Desert; the Paradox of Perfectionism; Effort and Merit. The traits a given culture most values at a particular time — say, intelligence (13), strength, and the capacity for diligent effort — are (1) the main targets of traditional improvement efforts such as training and practice and (2) arguably the most sacrosanct against technological tampering (14).

So the very traits selected for improvement are precisely those whose "artificial" ("nonnatural," "identity-compromising," "externally induced") augmentation is the most suspect. Indeed, if the wrong paths are taken, we may not even *count* the result as "improvement" or the accomplishment as (fully) merited. The very status of merit attributes may be impaired when they are technologically refashioned to extend beyond one's preexisting natural baseline, as augmented by traditional effort. But traditional baseline methods of self-alteration are not only *not* banned, they are *required* by perfectionist/progress ideals. (Whether such personal obligations accompany social obligations to assist individual perfectionism, of course, depends on the content of the ideal, and in turn on underlying basic value conceptions.) The upshot is that bettering ourselves in *inappropriate* ways does not "perfect" us — it lessens us. Thus, more is less. A possibly connected idea is that dispensing with effort as a critical component of achievements and improvements might "cheapen their value and cheat the social practices in which they play a role" (15).

As a rough intuitive matter, enhancement also raises troubling images of compromised personal identity, and thus of assignments of credit or rewards. If technology threatens identity, it also threatens the moral and political relevance of merit and desert. In turn, where merit/desert ascriptions are undercut, equality constraints become more muddled than usual. To assign greater rewards to those of greater merit than to others does not — on *some* views of equality — breach equality standards, and indeed may be required by them. If we cannot say who won the race, we cannot fully justify our assignment of prizes.

The capacity for effort at self-improvement or anything else seems to be a merit attribute, so this deserves some additional comments. We often prize *trying*, which we commonly view as under our control, as much or more than native endowments, many of which seem arbitrarily fixed. The results of traditional forms of striving are thought to be consistent with a stable identity. Now, the capacity to try is thought to be influenced by genetics and noncontrollable aspects of environment, as are other merit attributes. The talent for struggle is itself subject both to technological and nontechnological improvement (an infinite regress of trying?). How should we morally rate an increased capacity to exert diligent effort when the capacity is itself altered technologically? Isn't this as questionable as alterations of supposedly "fixed" traits such as intelligence, and of incrementally improvable traits such as strength? (One thinks of steroids in athletics here.)

It is also possible, in context, to view technological enhancement of endowments — including the capacity for effort — as itself reflecting a kind of praiseworthy effort. An increment in powers of memory, for example, may be unearned but nevertheless possess intrinsic value and instrumental value, as where it aids air traffic controllers in keeping up with the ever-increasing flood of data.

Demand for Enhancement Resources; Economics. The scale of demand for such resources depends on many

variables that cannot now be clearly identified and measured. These variables include the nature of the enhancement, its monetary costs and perceived medical risks, the deterrence or incentive effects of gatekeepers and their standards (physicians as well as bureaucrats may keep the gates because of medical risks), cultural variables (whether technological enhancement is (dis)favored in general or for specific traits will obviously affect its level of use), interpersonal pressures (also influenced by culture), links and interactions among different traits, and personal preferences. Different forms of technological enhancement may of course fare quite differently in the market (16).

Any stable economic unit requires supply and demand equilibrium for any commodity, and this presupposes diminishing returns for incremental distributions. Diminishing returns will no doubt set in at some point for distributions of "increments in merit," but as noted, this onset may be quite late in the game (as with many medical resources generally). You may not value yet another hotdog, but you can always stand to be smarter. There may even be expanding returns at various distributional stages. To make modeling and prediction still more difficult, different assemblies of merit traits may interact in unpredictable ways in the market; some traits will reinforce or "potentiate" each other, some will impair each other, some increments can substitute for others, and so on. Finally, recall the impact of extent of social use: Being highly intelligent is less valuable when all are highly intelligent, yet there may be as much or more pressure to consume enhancement resources even if solely to maintain one's position.

COMPETING VERSIONS OF (IN)EQUALITY

A Thought Experiment Not Far Removed from Reality

This exercise is meant to illustrate differences in specifying what is to be "equalized" through distribution (2,17). If the egalitarian goal is to attain $X = Y$, what might X and Y be and what does "=" mean?

Suppose we have a mechanism (e.g., drugs, somatic gene "therapy," or germ-line alteration) that can significantly enhance one's mental abilities. There are many possible distributional schemes for the enhancing techniques. The distributees might be individuals or possible individuals in early embryonic or even dissociated gametic form. The distributive mechanism might be central direction by government, markets, or kinship and other interpersonal relations. The effects of the technology, for good or ill, are likely to vary among persons. Sophisticated models will take account of the variability, but simplifying assumptions concerning uniform efficacy are appropriate for now. For example, one might assume that we will see the same per-dose linear increments in a given ability for all persons, or that effectiveness is a direct or inverse function of preexisting ability — the abler one is, the greater or lesser the increment. It is also helpful to leave aside the fact that mental abilities come in many varieties as well as strengths and that their recognition and status may vary among cultures. The particular regulatory or licensing schemes that would implement the distributive plan are not discussed here. Whatever the schemes are, for oneself, one's children, or some group, the licensing procedure must embrace either a substantive criterion or some objective mechanism such as a lottery or queuing (neither of which is entirely "objective").

Market Distribution. Free market exchange implements a sort of equality of opportunity based on ability to pay (including insurance and borrowing power). There is an extensive literature on the moral foundations of markets, including commentaries on the role of characterizing existing distributions of goods and bads as just or unjust, and on the impact of "preinstitutional desert" and fairness in promoting realization of legitimate expectations, but this is left aside here (18). "Enhancement" might be financed through health insurance mechanisms, particularly where the procedure can be viewed as treatment for disorder, defect, or injury, or at least as ambiguous (recall the example of immunological augmentation). (19) Here, a *ratio* is equalized across persons: *dose/financial resources-economic power*. As for nonmarket distribution via central direction, we might consider —

Centrally Directed Distribution of Equal "Doses" to Everyone. This is a simple, ham-fisted sort of equality. It suppresses individual variations and thus bypasses questions of need, merit, and utility. The ratio of dose to threshold status as a person is the same for everyone.

Distribution in Proportion to Need. What is equalized is the ratio of *dose to need*. "Need" itself is a disputed concept for several reasons, including the fact that many asserted needs are based on one's *relative* status within a population, that need exists in degrees, that it is afflicted with the difficulties within the treatment/enhancement distinction, that it may be linked to nothing more than one's preferences or goals, and that it may be unclear what follows from an ascription of need: Are there duties not to interfere with anyone's trying to meet their needs, or government duties to provide assistance? If need rests on having a recognized disorder or injury, then only the afflicted receive doses — e.g., the demented or persons with Down syndrome. It is unclear how to apply a "need" standard to statistical "outliers" who are not disordered but nevertheless are handicapped by their distance from the median. And need may, as suggested, be task related: Did Einstein "need" enhancement to make progress on a unified field theory?

The fact that one's needs may be based on being relatively worse off, whether in natural endowments or in environmental circumstances, requires special attention (20). Well-known political and moral theories call for measures of "redress" because many of the worse off are seriously disadvantaged. Rawls's difference principle, for example, suggests distribution of resource-attractors to equalize the dose/need ratio — where need is linked to relatively low status (21). The difference principle, however, can also be viewed as threatening other visions of equality, as well as values of autonomy, justice and fairness. Redistribution entails interference with

autonomy, and is arguably unfair in allowing some persons (the worse off) to reap the full benefits of their native abilities while preventing others (the better off) from doing so. Of course, a central question is "redistribution of *what?* The existing stock of wealth? Opportunities for enhancement?

Distribution in Proportion to "Social Utility." The ratio equalized here is dose to social utility. The social utility of some distribution pattern might be inversely *or* directly related to the distributee's relative ability, without regard to whether disorder underlies his or her low-end status. Distribution to those handicapped by low intelligence might reduce the need for social services. As for the very talented, think of encryption specialists trying to break an enemy's code (suppose that the British "Enigma" program hadn't cracked the German code in World War II?). Only those on the "edges" of human ability would become licensees. One might even expect pressures on government to require certain workforce groups, if not everyone, to use enhancement resources, though enforcement might be quite unpleasant — not to mention immoral and unconstitutional (e.g., violations of the Thirteenth Amendment).

Distribution in Proportion to Preexisting Merit. There are "native" ("endowed") merit attributes such as mental abilities, physical agility, and the capacity for diligent effort. There is also "acquired" merit based on accomplishment, good works, and developed skills and aptitudes. [The division between the two is hazy (8,53–68,109–131).] Here the ratio equalized is *dose to merit*. On this standard, the answer to "Who merits (more) merit?" is simply: those who *already* are highly meritorious. This sharply contrasts with a view of equality that sees natural variation in aptitudes as something to be overcome rather than presupposed as a suitable basis for distributing life's rewards (21).

Distribution in Proportion to Intensity of Personal Preference. The criterion here is how badly one wants something — including enhancement itself as a major facilitator for success generally. Extremely (pathologically?) intense preferences might be viewed as needs. Preferences and their intensities may also be regarded as a form of merit. We admire persons whose "desire to win" is strong enough to overcome serious odds.

Distribution to Achieve Equality of Outcome. Here, doses are distributed so that all have equal intelligence, however this level is chosen. There is no unitary concept of intelligence, and the task of equalizing all recognized forms of intelligence (22) seems far-fetched, but this is a thought experiment, after all. This rather open-ended outcome standard may entail that everyone be as intelligent as the previously most intelligent, or that the more intelligent are affirmatively impaired (they receive "negative doses") to reduce their status (23) while the less intelligent are upgraded. If the desired uniform ability level is Φ, the ratio equalized is the absolute value of *dose effectiveness to distance from* Φ (including positive and negative doses). The driving force here might be the alarming idea that equality requires or is aided by making persons as identical as possible. On the other hand, "equality of outcome" may refer to identical proportional increases, leaving everyone's position in the "pecking order" the same. Of course, these different forms of equal outcome are in general quite different, and they bear only an uncertain relationship to more comprehensive forms of equality of outcome — income or wealth, social standing, political power, and so on.

Distribution by Lottery. In an effort to bypass the immense difficulties in applying the ideas of equality, fairness, and justice, some have recommended distribution of scarce resources via lottery (24). Perhaps in this sense lotteries represent a form of being "unprincipled on principle" (25). In any case, the suppression of interpersonal differences entailed by lotteries (once the lottery's constituency is defined) is both the *point* of resorting to them and the chief *objection* to them.

Randomization schemes cannot be properly denounced on egalitarian grounds without a theory of equality that explains why equalizing over one field rather than others (e.g., doses, dose/merit, or other ratios) reflects or produces "true equality," or at least a preferable form of equality. If a satisfactory equality theory is unavailable, values other than equality must be invoked. As it stands, lotteries serve some visions of equality and rationality and contravene others. On one view, lotteries promote equality because all who qualify for the lottery (qualifying itself raises serious equality issues) have an equal chance of winning it, *despite* their varying personal characteristics. Indeed, it is precisely the attention to these varying individual traits that constitutes for lottery supporters a violation of equality: These interpersonal variations — rather than basic personhood itself — are to be *suppressed*. On the other hand, some will receive the resource and others will not, *without a "substantive" reason.* This situation is arguably irrational and thus a violation of equality standards, possibly under prevailing views of personhood and its entailments.

Equality Wars: Conflicting and Concurring Versions of Equality and Inequality; Remedies for Inequality; Equality, Enhancement, and Respect for Persons

In General: Equality of Whom or What and with Respect to What? What *do* we assert in saying that $X = Y$? 'X' and 'Y' might designate persons, groups, opportunities or prospects held by persons or groups; means for taking advantage of opportunities to achieve one's goals; specific outcomes (wealth, victories, etc.); social or moral status; political power; rights as persons (26), without regard to differing traits; traits characterizing different persons; ideas, conceptual systems and philosophies; and overall ("net") personal or group merit or social worth despite differing traits.

Each possibility rests on concepts that are themselves difficult to penetrate, and are likely to reflect serious political and philosophical differences. To assert equality of noninterference rights — such as free speech, free exercise of religion — is far from asserting equality of

means, opportunity, or prospect in securing audiences or places of worship (2). And these differing equalities may hold drastically different positions of respect and commitment among different persons and groups within a society, and from society to society. "Fundamental" noninterference rights are protected under the U.S. Constitution. Affirmative ("welfare") rights generally are not, even when directed toward increasing or preserving equality. The rejection seems to be founded partly on autonomy grounds, and partly on rival views of equality: Redistribution entails that some receive unearned rewards and others do not, and different persons will be allowed to keep different proportions of their wealth or income. The flags of equality, fairness, and justice are carried by all sides here — a point that is retained throughout this discussion.

One could raise parallel questions by asking about the meaning of the equality operator "=." Is it an assertion of fact (Arnold's strength is equal to Sylvester's), and if so of what sort? Is it a moral or political claim about equal rights or entitlements, and if so to what? Does it reflect an ideal both of the *threshold* equality of persons without regard to their differences and of how they should be treated? If we say persons are equal because they are all equally persons, why is undifferentiated personhood the right level of abstraction rather than personhood qualified by particular (dis)favored traits? If we say that the political power of (person)(group) X "equals" that of (person)(group) Y, we may mean they have equal numbers of votes, or equal power to elect candidates of their choice, or equal power to influence government policies (a particularly obscure claim), or any of several other options. It is not clear that we can justify our choice of "meaning" here via reference to equality alone, without reference to justice, fairness, autonomy, and utility — even if equality is not fully "reducible" to any of these other values or to some subset of them.

Equality and the Special Status of Merit Attributes. Judgments about an individual's merit are often relied on as a fundamental ground for sorting people — specifying certain (in)equalities among them — and for acting on these characterizations. The governing moral intuition (perhaps not in all cultures) is that outside the domain where only threshold personhood counts, persons are to be judged on their relative merits, and not on "arbitrary" personal characteristics or relationships. It is difficult to formulate a coherent theory for sound application of the epithet "arbitrary." For example, is it arbitrary — and thus perhaps morally improper — for individuals to search for a mate solely within their principal social group(s), which may be defined in part on the basis of ethnicity, race, religion, or national origin?

But comparative merit judgments are also criticized because, among other things, they produce unequal outcomes and may rest on unjust features of the status quo. More distribution of resources that strengthens one's measure of merit may expand and reify existing inequalities — even if it remains unclear whether artificial enhancements would be recognized as merit claims.

Valuation of Equality. As Temkin asks: "Is equality really desirable? And what kind of equality should we seek — that is, insofar as we are egalitarians, should we want equality of opportunity, primary goods, need satisfaction, welfare, or what? ... When is one situation *worse* than another regarding inequality?" (17, p.3).

The question of what "equality" means is distinct from the question whether it is desirable or valuable — although the two inquiries are linked in complex ways: Assignment or recognition of meaning often involves value analysis. There is, to be sure, an oddity in asking about the value of equality or of any "basic" value. How can one "value" basic values when these basic values represent the very terms in which value is defined and assessed? However paradoxical this may seem, we characteristically rank-order our values and assess them with respect to each other. But this is a matter for a comprehensive enterprise in moral theory.

Rectifying Inequalities. Plainly a major issue in genetic enhancement is whether it should be used to rectify inequalities by affirmatively creating equalities — and if so, how. Suppose that we reach a rough consensus on the preferred meaning of "equality" in various situations. There nevertheless may remain significant differences over appropriate measures to rectify or prevent inequalities. For example, if A's cache of goods is v but B has more, holding w, is A intrinsically worse off when B acquires still more but A continues to hold v? (17) From an equality standpoint, is it better to achieve equality by raising A's holdings from some outside source or transferring some of B's holdings to A? Or to enhance A's aptitudes and let him, on his own, try to overtake B? Should we worry more about inequality between certain groups than inequality within those groups (17)? Rectifying existing or past inequalities may implicate procedures that *themselves* may violate specific conceptions of equality — such as transfer payments. Such redistributions arguably impair the right to reap the benefits of one's natural gifts, as amplified by skills acquired through effort. As mentioned, they entail that some persons — those less well off — can acquire additional resources earned by others, and perhaps can keep a larger proportion of what they earn than can others.

"Rectifying" Differences. A population consisting of a single human clone (in the collective sense) might have equality problems, but the problems would surely be rather different from (but not necessarily lesser than) ours. A rather drastic (and perhaps technologically impossible maneuver) would be to make as many persons as identical in major respects as possible. The costs (from our present framework) in reduction of cultural and physical diversity and the loss of multiple perspectives in human endeavors seem very difficult to bear, although the radical transfiguration of human life makes them hard to assess. And, of course, the resource costs might be prohibitive. The point here is simply to observe that "difference" does not entail "inequality" in any sense relevant here, and few are on the stump for technological erasure of human variation (27).

Equality and the Morality of Inclusion and Exclusion. One might select among competing ideas of equality by appealing to a preferred morality of inclusion — of lumping by appealing to commonalities. The point of this brief reference is to emphasize two observations. Some may see the "technology of perfection" as allowing displacement of chance variations by planned similarities: We will all be perfectly equal because we will all be equally perfect. However, the very emphasis on perfection may impose serious burdens on those viewed as disabled or handicapped (3).

Is Equality "Empty"? Perhaps the indeterminacies (a term left undefined here) or conflicts *within* the idea of equality cannot be eased by further analysis of equality. There seems to be no overarching notion of equality to appeal to in all contested cases. The tensions may be irresolvable (17), though occasional consensus on certain matters may be attainable.

This is the central idea behind the claim that equality, at least in many important circumstances, is "empty" — a vacuous concept (28). The emptiness claim is roughly that the egalitarian maxim, "treat persons (dis)similarly situated in (dis)similar ways," cannot be understood and followed without a substantive moral/political theory of (dis)similarity that cannot itself depend on equality. We need, on this view, a theory with normative content to tell us what difference a difference ought to make. Equality alone does not tell what characteristics or actions (or anything) to "lump" as relevantly similar, nor what to "split" or suppress as relevantly different. For example, if government action permits some speech and restricts other speech on the basis of content, we cannot tell whether a constitutional or moral equality principle has been breached without a substantive free speech theory.

Equality and Other Values: Conflicts and Connections

It is often said that in many circumstances, equality, autonomy, fairness, justice, and utility (or any subgrouping) conflict. The nature of the conflict of course depends on the versions of equality and other values under review (29). A standard example is affirmative action. Distributing benefits on the basis of racial, ethnic, or gender criteria entails reduction of opportunities — a form of reduction of liberty and autonomy — for those without the relevant characteristics, and imposes forms of personal association on unwilling persons. These processes and outcomes conflict not only with particular views of equality, but with those of fairness, justice, autonomy, and utility (29,31). On any given set of views, justice may dictate what egalitarian maneuvers to prefer — say, that equality of opportunity, in justice, requires some degree of access to enhancement resources either via noninterference rights or via positive entitlements. Or some states of affairs or actions may be viewed as unjust because of a violation of some equality standard. (There are level-of-category problems in listing "basic" values. Not everyone would place these values on the same plane of moral reality or discourse. "Justice as fairness," for example, presupposes that at least certain versions of fairness are criteria for a higher-order concept, justice. Such difficulties cannot be further addressed here.)

Distributional Equality Generally; Distribution That Transforms the Distributees; Distributional and Nondistributional Equalities

Distribution and Personal Transformation. Distributional equality concerns who gets what, why, when, and how under any given system for distributing scarce resources. It addresses matters both ex ante (e.g., who gets the "merit-enhancing" commodity) and ex post (e.g., who gets what rewards — including still "more merit" — after the distribution and, at least in part, as a result of it). This ex ante, ex post distinction is particularly important given the possible "transformative" effects (left undefined here) of the distribution of enhancement resources. Augmentation may change the structure of the distributional game by disproportionately enlarging the distributee's resource-drawing power — ratcheting it up so that it is hard to undo. One might argue that all distributions "transform" the recipients and that there is no sharp distinction between the transformative effects of education or training, on the one hand, and of technologically augmented intellectual or physical functions, on the other. This is obviously true, but the absence of clear borders marking a distinction does not of itself trash the distinction.

Equality and Reduction, Mere Use of Persons, and Objectification: Some (Largely) Nondistributional Problems. Suppose that we believe a practice of enhancement reflects and generates excessive concern with the measures of specific traits and thus "reduces" persons to the (often commercial) value of these traits. This reduction is intimately connected with the processes of "mere use" and "objectification" of persons — their devaluation or descent from persons to "objects." A person who is (at least partially) objectified, reduced, and subject to mere use has thus suffered an egalitarian loss. (If everyone were reduced or objectified, however, there might be equality among the objects.)

ENHANCEMENT AND ITS EFFECTS ON (IN)EQUALITY; (NON)DISTRIBUTION OF ENHANCEMENT RESOURCES; REGULATORY CHOICES

Nondistribution Options: Nonallocation at the Macro Level; Restrictions on Manufacture, Distribution, and Use; Black Markets; Paternalism and Community Self-Protection

In General. There are many commodities that we think, for whatever reason, should not be distributed widely, if at all. To limit distribution, we can avoid allocating resources to the creation of such evils. If this fails, we can enact prohibitions or lesser regulations concerning distribution and ultimate use, although this may risk greater loss of control because of the rise of black markets. For example, prohibitions or severe restraints on use of enhancement resources may compound their risks by inhibiting safety controls such as physician guidance.

Whatever we decide about use of the commodity, the selection of the best regulatory mechanisms to implement our preferences remains open (31). Resolving this may

require empirical inquiries to inform the moral, legal and policy options. Certain arguments frequently offered to justify nondistribution require attention because they bear on equality.

Nondistribution to Protect Those Who Prefer Non-use: The Perceived Risk of Greater Inequality as a "Coercive" Factor; Technology-Driven Demand for Greater Skills and Thus for Enhancement. People often do things they would rather avoid because, if they don't, they fear others will gain advantages over them. One might say that doing so reflects a "straitened preference"—it is what they want only under adverse, dispreferred circumstances. This is not necessarily bad: A child averse to learning to read may find a "second-best" reason to do so when advised that "all the other kids are doing it." Similarly the risk of falling behind in athletic activities may drive some otherwise unwilling competitors to steroids or other supposed enhancers. Although one can dispute the claim that this is rightly called "coercion" via excessively strong incentives, a strong sense of pressure in some form on the unwilling is likely. In some contexts this matter is easily put aside: Many athletes and students do not wish to practice, study, train hard, or diet, but few speak of coercion in these contexts. One must compare such standard efforts with, say, ingesting memory-or muscle-enhancers to shorten the task. Some pressure to do the former is widely considered desirable—possibly obligatory under some circumstances; not so with the other. Few complain of coercion in these contexts, partly because the general endorsement of such self-improvement counts against applying pejoratives such as "coercion" or even "undue influence." Still the wide use of these characterizations suggests that something is believed to be amiss in the choice situations in question.

More generally, pressures favoring enhancement of living persons, fetuses, and possible persons are likely to grow. There is some evidence of technology-driven increases in the demand for "human capital" as reflected in investment in education and training (32).

Paternalistic Nondistribution to Protect Persons against Physical or Psychological Harm; Autonomy versus Autonomy. Enhancement measures, whatever their efficacy, carry risks of adverse effects. Of course, the nature, incidence, and seriousness of such effects—perhaps even whether they are in context adversities at all—is largely unknown. One justification for nondistribution may thus be pure paternalism (a term left undefined here) (33). Another justification, suggested above, reveals autonomy's internal tensions: promoting autonomy by reducing "coercive incentives" (34) to use disfavored commodities. But a broad interpretation of "coercive" may impair the autonomy of those wanting to use the suspect resources and who knowingly assume their risks.

An Equality Argument against the Preceding Nondistribution Arguments. Those seeking access to enhancement resources may of course also offer an important equality argument: They are denied equality of opportunity to better themselves and are relegated to an inferior status as against their superiors, whose natural gifts are as arbitrarily deemed meritorious as are the enhanced attributes.

Nondistribution to Reinforce Equality Values and to Avoid Devaluation of Life and of Effort; Ambiguities of Identity.

The Lombardi Effect: Winning Isn't Everything—it's the Only Thing (35); Paradoxes of Reduction and Valuation; Reduction as Compromising Equality. We learn in part from observing and interpreting social practices, and the societal risk here lies in the "lesson" or "message" that athletic victory is worth serious bodily or mentational harm. [A 1984 poll of uncertain rigor reported that Olympic athletes would accept death at an early age in a Faustian exchange for guaranteed gold (36).] If the practice is banned, getting caught and sanctioned is also perceived as a risk. However, it is the ban itself that is in question here. Note also that to characterize the risk as "mere" (as in "risking life and limb for mere athletic victory or for show") pressupposes certain value premises. One might urge, for example, that in a nation besieged by enemies on all sides, the supposedly adverse lesson about winning at all costs is useful for national self-defense by reinforcing a warrior state ideal.

To the extent the victory-is-all lesson is learned (by observers and the competitors themselves), it may reflect and constitute a reduction of human value to a single function or goal—athletic or other competitive success. Such devaluation bears on equality. Those persons who are reduced (a concept strongly linked to "mere use" and "objectification") are seen as less worthy than others, and thus unequal to full-valued persons. The enhanced—if that appellation survives—may be viewed as (partial) artifacts of *lesser* merit who leapfrogged over their associates and competitors. If so, it may be the *enhanced* who require protection from the unenhanced (37). (In that event, resulting inequalities of distribution would not be perceived as reflecting true differences in personal worth, and artificially enhanced talents would not be seen as lessening or devaluing the natural talents or acquired skills of others.) The value of enhanced persons will have collapsed into a narrow range of traits based on their prospects of victory. This supposed reductive risk carries us to the next difficulty.

Troublesome Links Among Enhancement, Value Reduction, and Positive Valuations; Role of Risk-Taking; "Person Perception". Value reduction by focus on specific traits or accomplishments is intricately linked to valuing persons positively—viewing them as meritorious and deserving of our high regard. We value persons not only because of the traits that define their threshold value as persons, but because the strength of those traits distinguishes them from others. How does this differ from "reducing" them to their traits? If we cannot say, the contrast between reduction and positive valuation seems empty. Suppose that an athlete says "winning is the only thing and is worth my life." Does this reflect "reduction" and lesser status—in her own eyes or the eyes of others? Or, on the contrary, supervaluation and greater status? Or some combination? The more general question concerns how we indeed perceive each other as particular persons—an issue not only for moral and political theory and practice, but for continuing work in cognitive psychology.

The answers, if any, may depend partly on cultural baselines. The United States for example, is not about to ban football or boxing because they risk severe permanent injuries or death. But *incremental* risks beyond a traditional baseline may be rejected because they reflect inappropriate trade-offs: All risks are justified if winning is everything or the only thing. (Think of escalating boxing to fights to the death.) Such a view arguably reflects a debasement of the value of life in the eyes of the audience and in the competitors' own eyes. (This particular argument would of course not directly apply to "magic bullets" — zero risk but effective enhancers. But there may nevertheless be possible adverse effects of extreme focus on the traits needed to win.)

As for risks undertaken for intellectual enhancement — such as a dangerous, possibly fatal drug that greatly augments memory — the situation is unclear and may depend on cultural circumstances affecting the comparative valuation of traits (e.g., intellectual as against athletic abilities). In Frank Herbert's *Dune*, rival feudal houses far in the future relied on resident "Mentats" — "wizards" of a sort — who amplified their preexisting exceptional intelligence with an addicting "spice." They evidently were judged *both* on their initial baseline abilities and the level of their enhanced abilities — the latter probably being a partial function of the former (38). Perhaps they were also judged on the skill with which they enhanced themselves and on their courage in facing risks. In this sense, at least a sliver of their "endowed" merit endured. Given their culturally imposed duties, what they did can be characterized in context as praiseworthy effort rather than merely as deriving an unearned benefit. Moreover their work would continue to reflect effort and struggle if, partly as a result of their enhancement, the complexity of their tasks increased (14); those who can do more complex work are likely to have it presented to them. Still, as suggested, artificial enhancement may not be viewed as meritorious enhancement at all. Indeed, from this standpoint, the more revered an attribute is, the more it is corrupted through technology's manipulations, and the more degraded is the "enhanced" person.

Avoiding the Social Devaluation of Effort. This category embraces both paternalistic and nonpaternalistic reasons for nondistribution of "elite-creating resources." As suggested, a possible (perhaps inaccurate) impression conveyed by a visible practice of enhancement is that the enhanced are getting a free ride (or at least a reduced-cost one). So, it is argued, enhancement entails getting too much bang for the buck — cheating of sorts — even if done in the open. If effort is devalued, merit judgments are distorted, and so are our judgments concerning the (in)equalities holding among competitors. (The claim that merit evaluations are "distorted" presupposes some preferred baseline from which to measure distortion.) What are we to make of the fact, for example, that some students, relying on their own gifts as aided by work, are disadvantaged in college-entrance examinations as against "artificially" or "unnaturally" able persons? To avoid this inegalitarian disadvantage and the disruption of prevailing norms, would a student need a license reflecting that she had not been enhanced before the exam? (Or might it be the other way around — "nonenhanced need not apply"?) Of course, one may question the sanctity of prevailing norms, but within reason, communities have some defeasible moral and legal right to maintain the major features of their normative systems (39).

Devaluation of effort might also (paradoxically?) arise from enhancing the very capacity or inclination to make efforts: Some may view the result as external to one's character — an outside supplement that carries no merit with it. Some authorities, for example, say that steroids may expand one's capacity to exert efforts before reaching exhaustion (40).

More on Threats to Identity; Confused Attributions; Effects on Equality Judgments.

1. *Altering living persons.* Recall the distinction between germ-line alteration and gene therapy or other somatic treatment on living persons, fetuses, and embryos. In dealing with living persons, some aspects of equality concern assessing the fairness of returns on effort and of rewards for one's native endowments — particularly when effort and endowment are combined. Such judgments obviously require identifying and comparing persons. Enhancement may confuse notions of personal identity in at least two ways. It may create ambiguity as to the "source" of one's performance — whether it is "internally" generated and thus causally attributable to that person, or the result of "external" artificial augmentation and thus "attributable" to an outside source. If the latter, then the person in question "didn't do it." The judgments of merit and desert underlying the distribution of rewards will also be confused.

   Moreover, enhancement may, by altering personal traits, appear to interrupt the continuity of human identity, which in other circumstances endures despite gradual, historically acceptable change. In both situations equality appraisals — at least in extreme cases (not yet at hand) — may be distorted or even meaningless. Who or what is equal to whom or what? Who won the fight? Should an enhanced Mentat transplanted from a Dune world be eligible for a chaired professorship or the Nobel Prize? (Data, an android in *Star Trek: The Next Generation*, was ultimately awarded the Lucasian Chair — the same Cambridge chair held by Sir Isaac Newton.) Should he be rewarded only if all other candidates had similar enhancement opportunities?

   Do major trait changes truly compromise personal identity — say, a sudden escalation of intelligence from average to extraordinary (41). After all, one's "baseline" identity doesn't really disappear: The new one is built on it, and all new identities will continue to differ sharply from each other. One can well imagine, as suggested, educational institutions debating whether their admissions criteria should exclude augmented persons as not "truly meritorious" unless their natural abilities ex ante — their "enduring merit" — would have secured their admission. Should we compare and rate

persons only on the basis of endowments pre-enhancement? Or is this irrelevant history? Even if pre-enhancement merit remains relevant, the "locked-in" resource-accumulations made possible by accelerating returns might be viewed as going far beyond fair rewards for ability — assuming that *that* distributional criterion is believed sound.

2. *Altering germ lines.* Identity problems may take on a different form when one considers the genomically enhanced — through alteration of early embryos or of gametes — as well as those who were enhanced somatically. The latter can be further divided into persons enhanced in utero, as embryos in vitro, or as young children, or adults. Our sense of self-identity, autonomy, and personal worth may differ sharply depending on our knowledge of the nature and timing of our enhancement, and on the reasons for it. Knowing that one's genome was altered will not necessarily have the same impact as knowing that one's physiology was altered. There will also be differences depending on the timing of the somatic changes — principally whether pre-memory (where the person would know herself only in her enhanced — and, to her, "native" form) or within memory. In any of these variations, think of children asking their parents just why their "natural identities" were tampered with or even changed, whether genomically or somatically. Was a potential person (the *un*enhanced entity) adversely affected because its existence was blocked and replaced by "another" person deriving from the altered entity (42)?

Equality Impacts of Technological Enhancement: More on Distributional Options

In General: Enhancement That Alters the Bases for Distributing Benefits and Burdens. As we saw earlier, the "equality impacts" of distributing anything depend not only on matters of fact but on what notions of equality are used. In turn this may determine what "units" or entities are being compared and targeted for what forms of "equalizing" with respect to what distributable entities. We may be addressing existing or future persons, families, groups, and so on, with respect to equalizing income, wealth, social status, legal and political rights, and opportunities of many sorts, and so on (2).

We now need also to distinguish between distribution of the resources needed for enhancement, and distribution of all other commodities. Distribution of enhancement resources changes the game by altering the criteria for resolving distributive claims; this "feedback" may be far more striking than the distribution of education and wealth.

Do Disorder/Treatment Models Blunt Equality-Based Objections to Enhancement Distributions? — Treatment as Restoring Equality Rather Than Distorting It through Enhancement. The connection between disorder models and equality was suggested earlier. If one's relative incapacity is disorder based, health care may restore normality. (It may also create it for those congenitally disordered by raising them to a "normality baseline.") This is less likely to be viewed as a suspect form of enhancement, at least where the disorder and matched treatment are well-recognized as such. Indeed, such medical intervention may be thought to *promote equality* by reinstating equality of opportunity, undistorted by adverse medical conditions (3,15). In other respects, it may worsen equality conditions through pressures to move affected persons to normality. This may downgrade "alternative lifestyles," with adverse effects on various groups — those within the deaf culture, for example. There will, nevertheless, be strong pressures to expand the boundaries of the disorder/treatment model in order to secure insurance or other forms of payment. Of course, the greater the expansion of coverage of various "medical conditions" and of persons, the higher the price of insurance, and the greater the exclusion of lower income groups.

But enhancement not justified within a disorder model is likely to be seen as impairing equality in several ways. One is by distorting "nature-based" equality of opportunity resting on native endowments as elevated by customary forms of self-improvement. Another is by interfering with the unequal but arguably justified outcomes of competitive pursuits (43). (This is of course heavily dependent on the reigning political philosophy.) On the other hand, suppose that enhancement becomes legal, its use disclosed, its price relatively low, and its efficacy roughly the same for all. Then, whatever other objections would remain, inequality concerns, though remaining important because of existing and enduring positional differences, would be partially muted. Of course, other moral issues about enhancement would endure.

However, if variations in natural endowments were thought irrelevant to merit and desert, the point of distinguishing treatment from augmentation would largely be lost, except for clear medical need. As Daniels puts it, if one rejects the "standard model" that takes the distribution of abilities as given, then "the distinction between treatment and enhancement has no point, at least where enhancement is aimed at equalizing capabilities (43, pp. 124–125)."

Enhancement and the Demise of Merit; Person Perception Again; Interpersonal Comparisons of Merit, Desert, and Equality; Entrenchment of Elite Blocs (New or Old); Racial, Ethnic and Gender Dimensions of Enhancement.

In General. Perhaps the very idea of merit would largely "drop out" if the use of enhancement resources were widespread and comprehensive, surviving, if at all, only when applied to judging skill in arranging for and using such techniques. (If we also do not deserve the traits we each received From Above, what meaning does "merit" have other than a thin estimate of economic worth?) Assuming such "no-merit" assessments are made, they are likeliest when living persons are augmented by medical/surgical means, including somatic cell gene therapy or "genetic pharmacology." But our more immediate target is to trace possible effects of such enhancements on different forms of equality: social equality; political equality; equality of opportunity (broken

down into matters of means, prospect, and so on (2,44)); group equality; and the roles of merit and need in making equality judgments. Although these broad and rather clumsy categorizations of equality can carry us only so far, they are useful starting points.

"Social equality" rests partly on the differing frameworks for "person-perception" (44) we use to appraise each other — and ourselves. Perhaps the genetically enhanced would perceive themselves — and be perceived by others — as "superior" in any of several senses: possessing greater intrinsic merit (even if artificially elevated by humans) and hence desert for various rewards; being more useful to society — and so more worthy, in both moral and nonmoral senses; and belonging to an elite group holding substantial political power. Perhaps this elite group would be the successor to an established powerful group; or it might constitute a new kind of elite based on genetic or other augmentation.

Another concern relates to the formation of blocs defined by the particular nature of the enhancement. People regularly sort themselves into groups defined roughly by the strength of particular traits: the more intelligent, the more physically fit, the more nerdy and so on. Enhancement of these traits might solidify these groups and strengthen their political and economic power. Their continued existence as discrete and enduring entities is suggested by how they differ from, say, political parties in the United States.

More generally, if distribution of expensive enhancement resources followed a market or preexisting merit path, existing socioeconomic distances would be enlarged and less bridgeable. This might reinforce adverse views about various ethnic and racial characteristics. Because of the self-reinforcing nature of distributions of "merit" — of the very grounds for distribution generally — the creation of entrenched elites may be hard to reverse. (If traits involving regard for others were so distributed, it might seem odd to speak of unbridgeable distances.)

Threats to Equal Respect for Everyone's Common Personhood: Enhancement as Intensifying Concern with the Strength of Specific Traits, Leading to Reductionism Generally and to Devaluation Resulting from One's Reduced Standing. Some equality judgments may involve a sort of "suspended belief" concerning the extent of interpersonal differences. One reason technological enhancement seems more unsettling than familiar forms of self-improvement is that it calls attention more specifically to the enhancing agent's target traits. Why such enhancement is more salient than long-term, gradual advances is an issue for cognitive psychology, particularly as it bears on person-perception. In any event, our common personhood may be overshadowed by a more intense focus on interpersonal differences. Our moral value *as persons* may be partially displaced by our increased value *as bearers of certain traits in certain measures* (14). To *plan* a person's traits suggests that those traits, as augmented, reflect his primary or even his only value — "value" here meaning social utility. This reduction, as suggested, is affiliated with the ideas of mere use of persons (in violation of the second formulation of Kant's categorical imperative) (45), objectification (46), and related processes. (The nature of this affiliation will not be investigated; it is not necessary to specify which of these notions are criteria for the others or inferences from the others, and they are taken as more or less substitutable here.)

Threats to the Valuation of Persons with Conditions Generally Viewed as Disabling, Particularly Those Who Decline Measures — "Therapeutic" or "Enhancing" — to Eliminate (or Improve) the Condition or Prevent It in Others. There are conditions that, in some cases, are not viewed as disabilities but as enablers. The best-known example is that of deaf persons, at least within what has come to be known as the deaf culture. If they decline measures to enable them to hear — assuming effective measures became available — they might be severely called to task, particularly if they continue to press for special social services. If parents, deaf or hearing, decline these measures for their children, they may be criticized even more severely. These measures — the limiting case of "normalization" (as compared to "assimilating" or "mainstreaming") may, as Silvers puts it, "devalu[e] alternative or adaptive modes of functioning" (12, p. 112). Still more, there is controversy concerning the moral propriety of terminating pregnancies when the developing fetus is believed *not* to be affected by a form of hereditary deafness. Moves to make this more difficult for deaf couples may also be viewed as devaluing deaf culture.

Forms of Regulation; Markets and Other Procedures; More on Equality's Internal Conflicts.

Natural Differences. There are obvious natural and acquired differences among persons, although their significance and even their recognition may rest both on competing moral frameworks and on cultural variables (18). Not all differences — assuming they are perceived at all — count as inequalities. In some cases, cultural variation may be of modest significance. Persons born without limbs, for example, will have difficulty in moving independently from place to place, and this is obviously a crucial ability for most persons in most cultures. But just how well such persons fare depends heavily on variations in familial and general social practices affecting those impaired in this and other ways.

Where differences are recognized as significant, however, one must inquire into their moral status how they should be dealt with. The differences might be taken as given and their effects left to the workings of decentralized market, kinship, or other private arrangements. Or, communities might try to improve matters from an egalitarian perspective, viewing the fact of major interpersonal differences as "natural wrongs" or injustices (17,18). Any effort to displace the market would of course take us into a different phase of moral and policy analysis: establishing criteria for distributing resources, including enhancement resources, and specifying procedures to verify that the criteria have been satisfied by prospective recipients (native endowments, accomplishments, prospects, interpersonal connections, etc.). Any choice of distributional regime — market, nonmarket or mixed — necessarily involves contested moral issues. The "genetic supermarket" (Nozick) may be efficient (21,47) in some sense, but it does not bypass foundational problems.

Recall that the developing technologies involved here are "reflexive" in the sense that they are meant to alter and enhance its consumers, and access to such resources is likely to roughly track prevailing distributions of economic or political power. This changes the normative terrain considerably because each distribution changes any static models we have been dealing with by (possibly) sharply changing the criteria for each successive set of distributions. We can no longer rest on assumptions that major human traits change only gradually, if at all (48).

Janus-Faced Equality. Technological enhancement provides some opportunities to even out nature's hierarchical roughness. It also creates the possibility of worsening it, as emphasized above (1,37). If such leveling is indeed a community moral obligation on egalitarian (and other) grounds, a ban on enhancement — when used to promote rather than impair equality — might violate a principle (corollary to some forms of equality) mandating rectification of specified inequalities through certain mechanisms.

But such apparently egalitarian rectification efforts would require centralized intervention into distribution of enhancement resources. Enabling the have-lesses to move closer to the have-mores might thus *itself* violate some aspects of equality — and of autonomy, fairness, justice, and utility — through coercive redistribution. As we saw earlier, some would lose more of what they earn than others, and some would receive unearned benefits while others would not.

In any case, it is unlikely that any distributive scheme would "level out" human traits. And few — from current perspectives — would think it desirable, morally or otherwise.

A Review: Inequalities Compounded; the "Matthew Effect" and Terminal Social Stratification: The Problem of "Who Merits Merit?" Again. The rich get richer, the poor get poorer, ... and the smart get smarter? Why not? The well-educated already more easily qualify for still more education, often to the exclusion of the less-educated (49). "For unto every one that hath shall be given, and he shall have abundance: But from him that hath not shall be taken away even that which he hath" (Matthew, 25:14–30). (Merton coined the phrase "Matthew Effect" in referring to allocation of resources in scientific research.)

Distribution of scarce resources is of course a classic problem for economics, ethics, political theory, public policy, and just plain politics. But the distribution of enhancement resources, as we saw, raises special issues. Enhancement almost inevitably targets merit attributes — which are generally wealth-attracting resources.

The relevance of the Matthew Effect is obvious. The distribution of resources for enhancing "merit" claims for distribution, as we saw, involves a sort of feedback loop: It alters the very ground on which the initial distribution is made, generating a multiplier effect. Under such conditions Thomas Jefferson's "natural aristocracy" of "virtue and talents" is replaced by an artificial aristocracy of technologically enhanced abilities (50).

Of course, if the idea of merit does not survive the new age of technological alteration, "Who merits more 'merit'?" becomes doubly a nonsense question. Not only are we unable to increase our merit artificially, but merit itself is gone as a relevant moral category. Even if all or some characteristics lose all or some of their status as merit attributes, it seems likely that enhanced intellectual and physical powers, unevenly distributed, will continue to attract wealth and resources. Business, after all, is business: Intelligence counts for scientific research; heft counts for football; attractive faces and bodies draw attention and money — whether or not we talk about merit. The demise of merit, moreover, may not dispatch the view that we are nevertheless "entitled" to the fruits of our varying natural *or* enhanced abilities — whether we "deserve" any of them or not. In any case, the outcome of decentralized distribution of resource-attractors, as suggested, might ratchet up social, economic, and political stratification, and the hierarchical structure of community life generally. At least this is a potential outcome if distribution is based largely on decentralized mechanisms, such as markets, kinship, or old-boy/girl networks.

Enhancement and Interpersonal Desert: Time Scales, Life Plans, and Social Stability.

Enhancement of Living Persons within Their Respective Memories. We are accustomed to the gradual acquisition of merit earned by effort, resulting in gently escalating desert Sudden, major alterations in attributes, particularly merit attributes that help define one's identity, aren't associated with ordinary persons; Western culture links "shape-shifting" to mythological para-human creatures. But part of the very point of technological enhancement is to shorten the time span and reduce the effort needed to strengthen one's attributes beyond their endowed "maximums," and to gain the resulting incremental rewards. Such sudden changes in individual capacities may present major difficulties to a transformed person, to those around her, and to society generally (41). Our choices about life style and life plan have always depended strongly on presuppositions about our attributes — both assets and deficiencies — and their general stability (a stability consistent with their gradual elevation or deterioration). Suppose, however, that someone of modest talents and accomplishments rapidly becomes abler. Would she think that she *deserves* more of life's rewards because her abilities have sharply increased? How would she acquire these rewards? The newly intelligent or memorious (51) can't just saunter onto the grounds of Acme University and demand entry and possibly displacement of their (new) inferiors. Or can they?

Perhaps the spreading self-awareness of new powers — and the spreading fears of those stuck where they are — will provoke political and social instability. A somewhat distant analogy would be the sudden emancipation of large numbers of slaves or indentured servants who had been denied education and other resources needed to flourish as free persons. Think also of the comparatively rapid (if incomplete) change in the status of women in the United States and elsewhere. Virtually every aspect of equality would be challenged by technological enhancement. In particular, the nature of the contests between different forms of equality may also change. The perennial war between forms of equality of opportunity and forms

of equality of outcome may be intensified by the limited availability of enhancement resources that greatly enlarge one's prospects, and by the growing "distances" between the enhanced and the unenhanced.

There is thus a two-stage egalitarian problem: determining who gets "merit-enhancing" resources, and determining what collective or individual responses to make when faced with the escalating demands of the newly enhanced. These *nouveau intelligent* do not suddenly enter the fabled set of fully qualified rocket scientists. (Memory and skills transfer is not discussed here.) But they will argue that they have joined the set of persons immediately entitled to further education and training, and, within a short time, to appropriate forms of employment and their attendant rewards. It is too early to say whether it will make any difference whether they frame their claims by relying on merit and desert or on economic and social utility.

Other Enhancements. Questions parallel to those just raised arise with those whose genomes were altered, or possibly whose traits were revised during embryonic development (but without genome changes) or fetal development or in early childhood. Our responses, however, may be different. All of these persons are likely, in different ways, to look upon themselves as identified with traits they have had "wired in" for as long as they can remember. Moreover, for whatever it is worth, they cannot themselves be accused of having tried to evade or soften the struggle for self-improvement and to reap unearned benefits. Those enhanced as adults or older children, however, will be able to compare their attributes "before and after" enhancement, and those who affirmatively opted for enhancement might be blamed for such (partial) evasions.

Again, it is unclear how limited-access institutions such as educational facilities and desirable employment opportunities could quickly adjust, even over one or two generations, to a sharp escalation of merit claims for entry. A further complication is that some forms of labor may become even more disfavored than they are now among more educated groups — cleaning/sanitation, simple but hard labor, some forms of blue-collar work, and various low-skilled personal-service functions. Other things remaining fixed, however, the shortage of supply for such labor would raise its wage rate, which presumably would draw applicants willing to trade (temporary?) embarrassment for an enlarged income. (But then, their services might then be too expensive for many consumers, especially among the unenhanced.)

Both Groups in the Long Run. The questions just raised also apply to long-run considerations, and here matters become still more speculative. How would we forecast shifts in attitudes and beliefs about interpersonal comparative valuations? We do not know how we will value (or reduce) each other when merit traits are significantly malleable.

Still we are not entirely at sea and can make at least minimalist projections. One would think that with escalating demand for the (possibly) superior services produced through stronger merit attributes applied to increasingly more demanding tasks, investment would gradually yield institutional responses: more educational facilities, more complex mental and physical competitions, and new technologies enabling disfavored lines of work to be done more by machines and less by persons. This would generate greater incentives and pressures for still further personal enhancement, new stages of institutional response, and so on. It is hard to say where diminishing returns and equilibrium would set in. Recall also that one effect of the greater salience and strength of merit attributes might be to amplify the social, economic, and political importance of the enhanced traits: enhancement efforts would be likely to require major investments of all kinds, both financial and emotional — and people want returns on their investments. The result would be still greater emphasis on interpersonal differences.

One theoretical possibility should be kept in mind for analytical purposes, however unlikely it may be: With broad access to similarly effective agents, there might be little *relative* interpersonal change, even though everyone's individual performance level was raised. But this would also represent a major source of pressure for social and economic revision: Nearly everyone will be abler, more insistent on appropriate rewards for ability, and more concerned about the responsive formation of new institutions to satisfy their new levels of talent. The sluggishness of social and political responses to the claims of those with newly enhanced attributes may contribute heavily to various forms of social instability. And some instabilities might well arise because of negative shifts of attitude both toward those with clearly defined disabilities or handicaps and those who are unenhanced (or not successfully enhanced) and find themselves ever-lower in relative standing.

Equality of Groups and Blocs: More on Social Stability. One uncontroversial point is that groups and communities play major roles in social and political life and, partly as a result of this, in the formation of one's sense of identity and self-regard. Matters of interpersonal equality are thus conceptually linked both to intergroup and intragroup equality — and both realms of equality are affected by prevalent views on merit and desert.

Humanity has generally sorted itself into groups, and some existing groupings are defined by observed or supposed differences in merit traits and accomplishments. Indeed, in a distributional system based entirely on the purest notions of merit, with the arbitrariness of prejudice, stereotyping, corruption, fraud, and coercion largely absent, one could infer that resulting differences in attainments, rewards, and status are based entirely on differences in abilities or other merit or wealth-attracting resources. Perhaps this just replaces one set of "arbitrary" criteria (old-boy/girl networks, ability to pay, kinship preferences) with another (genetic and environmental lotteries), but this is another issue. In any case, this somewhat intimidating prospect has been addressed in several well-known (and controversial) works (52). If realized, we would in theory lose our excuses for failure (e.g., "politics did you in"). Our relative status would rest on "the merits" and unambiguously reflect our attributes, perhaps as in Mensa, whose membership is chosen (in

theory) on the basis of pure ability rather than interests or accomplishments. Once again, we face a raising of the borders between existing groups, and possibly the creation of new entrenched factions.

Still, we do not know whether any given pattern of enhancement would inspire social/political instability. Much may depend on whether, despite the greater gaps between individual and groups, the lot of the worst off is nevertheless improved (53). If the size of the gap between the better and worse off is great enough, the overall risks of instability may go up even if the resources of the less well off increase. Our notions of poverty seem to involve ordinal rankings as well as the cardinal value of one's holdings.

Political Equality Imperiled: In General.

Shifts in Political and Moral Ideals; Widespread Use and Low-Cost Access. It seems that many prospective parents now prefer the genetic lottery and are eager, or at least willing, to accept whatever they receive from it. We often eschew planning even where it is possible to plan, preferring vagueness and uncertainty to precision and predictability. Perhaps a partial explanation for this is fear of responsibility when things go awry or simply not as planned, or confusion over what to select. Beyond this, such preferences seem linked to what is perceived as a defining element of personhood: We envision persons as creative and autonomous, not bounded by fixed life plans imposed on them. Nonpersons have none of these attributes. Neither natural nor assembled objects possess them, and other living things seem too far off the mark to justify such characterizations (54). But competitive pressures may inspire many parents-to-be to seek greater precision of outcome in their reproductive plans, and (possibly) to rigorously enforce these plans on their offspring. Fear of an unenhanced child's eventual reaction to discovering that she is disadvantaged might well play a role here. (The parents in most cases cannot respond by telling the child that she had no alternative existence. The selfsame embryo from which she developed could have been isolated and altered, or her traits could have been changed after birth.)

Assume that the longer-term results of these pressures and of economies of scale are nearly universal low-cost successful efforts to enhance. Equality complications attributable to enhancement would be then be greatly attenuated, though not entirely removed (and standard equality problems would likely endure). From this particular egalitarian perspective, if not others, the more technology and the wider its use, the better.

However, where there are large-scale distributional inequalities, there is a risk of (irreversible?) erosion of equality's status. Equality could be adhered to (if at all) only in the sense of preserving the abstract idea of equality of opportunity: no affirmative blockade interfering with one's right to use her preexisting intelligence and wealth to secure more intelligence and wealth, for herself and for her existing or future offspring, and to reap the benefits of her enriched capacities — and so on down the dynastic generations.

How might this shift our ideals? Institutions and practices, by their very existence and visibility, "communicate" ideas and impressions, and these may have learning effects (55). Of course, what is "learned" depends on what is perceived or understood, and might be reshaped by responsive public debate.

Segmented Society. One feature of a world with both genetically and nongenetically enhanced persons might be a more rigorous division of labor, perhaps of the sort envisioned by Plato in his *Republic* (56). After all, if we take the trouble to (re)assemble our offspring with certain "engineered" traits, they had better do what we planned, right? Equality analysis here is of course beset with factual and normative/conceptual uncertainty. On the one hand, the escalation of technological complexity combined with enhancement might lead to greater division of labor and social stratification. On the other hand, enhanced persons might form a world with less rigorous division of labor because they become polymaths and jacks-of-more-than-one-profession.

Still, it is conceivable that regardless of how rigorous the division of labor is, political and social equality of a sort may hold. Different professions, trades, and occupations, and the varying aptitudes underlying them, might be viewed as equally worthy — an "equality of the enhanced." The "alphas" may be viewed as equal to the "betas," though their augmentations (via the germ line or the living body) and life work may be entirely different.

But this is nothing to count on. It is also plausible to expect that equality is largely "read out" where (from our present perspective) it is most applicable and most needed. The more entrenched the social stratification becomes, the greater will be the need for corrective notions of equality and of "remediation," but the less likely it is that there will be influential partisans for equality in any sense.

A More Equalized Society Instead? As suggested, enhancement resources might be distributed in ways that promote equality in several forms, consistently with whatever divisions of labor are implemented. Distribution might rest on need, where "need" is linked to enhancing equality of opportunity, perhaps vindicated by some degree of social assistance. Moreover every person might be considered to have a stronger claim to augmentation than his immediate "superiors" in preexisting attributes, giving him the right of first refusal for the next set of resources. And, as we saw earlier, where different traits are enhanced, the "net equality of the differently enhanced" may hold — as equality of "overall" merit. Finally, greater equality might be pursued by familiar redistributive or other social measures that are not directed toward trait alteration. So there is a slight possibility of a "more equalized" society. But this possibility should now be vetted through the lens of democratic theory.

More on Political Equality Imperiled: Democracy and Governance.

Enhancement and Democratic Theory: Millian Plural Voting and the Attenuation of Democracy

1. *Kinds of democracy: Is one-person, one-vote a defining characteristic of democracy?* What are and should be the effects of sharp differences in human

characteristics on matters of political governance? If we are not in fact equal to each other in deliberative ability, judgment, and drive, why do we all have equal voting power in the sense that when casting ballots in general elections, no one's vote counts for more than another's? We are not equal in our knowledge of the issues, our abilities to assess competing arguments, the nature and intensities of our preferences, our capacities to contribute to our social and economic system, our stakes in the outcomes of particular government policies, or even in our interest in participating in public affairs. And enhancement technologies may amplify these differences.

Yet for most of us, "democracy" seems to be all but definitionally connected with the maxim "one person, one vote." Unless this maxim holds, there is on this view no true democracy. Is this definitional link indeed appropriate given our vast interpersonal differences? Not all political thinkers have thought so. As Thompson summarizes John Stuart Mill's discussion of plural voting:

> The principle of competence expresses Mill's belief that a democracy should give as much weight as possible to superior intelligence and virtue in the political process (57).

Mill thus did not think that equal votes among electors was essential to democracy or for promoting the public good — quite the contrary. He endorsed plural voting (though perhaps with later reservations and possibly as a temporary measure) in which individual citizens had votes proportional to their "individual mental superiority" (58,60). The number of votes per elector would thus be a function of his or her revealed competence. Mill discussed occupational success, test results, and educational status as criteria for assigning more than one vote (59, pp. 475–476)

For Mill, plural voting is one method, among others, for furthering the principle of competence (57). His idea of competence is complex, however. It is not addressed simply to intelligence, but to skills (57), since highly intelligent persons might lack skills needed for sound governance. His vision of ideal competence also includes "moral competence." Education seems to be not just a proxy for competence, but partly constitutive of it. (One supposes that it might be a proxy for native ability.) Finally, Mill qualified his recommendations by recognizing that participation values were in tension with competence values (59,60).

The link between Mill's competence principle and human enhancement is clear. Genetic engineering, for example, has long stimulated fears that enhancement would threaten democracy — at least in forms demanding equal votes for all electors. Sinsheimer has asked, "Could ... deeper knowledge of the realities of human genetics affect our commitment to democracy (61)?" If mere knowledge of the physical bases (both genetic and nongenetic) of the differing endowments underlying Millian competence can threaten democracy, one might well expect that the vivid reality of the huge gulfs between the enhanced and the unenhanced would represent an even greater threat.

But the nature of the threat to democracy must be specified. This is no simple task, given the fact that "democracy" may take quite different forms, and that the status of the one-person, one-vote standard as the premier form of democracy is a question at issue. All forms of democracy are linked by the idea that the governed — or some portion of them — are to have a significant say in what affects them, and that this voice is to be broadcast by some form of majoritarian aggregating of votes on important matters. This voice is an obvious component of autonomy, which is in turn an essential ground of democracy, and it is not simply advisory or merely a request for redress coming From Above. It is to be decisive within significant domains, although it may be subject to principled constraints derived from constitutions or other sources of law. What constitutes an "important matter" is of course hugely uncertain, but clarifying it is unnecessary here.

Return now to the question that opened this section: Why is the political equality that is implemented by one-person, one-vote accepted in the face of individual differences? Dahl raises a parallel question:

> [I]f income, wealth, and economic position are also political resources, and if they are distributed unequally, then how can citizens be political equals? And if citizens cannot be political equals, how is democracy to exist (30, p. 326)?

(Dahl is describing conflicting theoretical perspectives, not necessarily endorsing any.)

2. *Applications to an age of enhancement; one-person, one-vote.* A system of equal votes at the ballot box is far from ensuring "equal" political influence or equality in anything. Think, for example, of the suppression of group preferences in at-large voting districts (62). Nevertheless, in its own way, one-person, one-vote implements equality both practically and symbolically. If effective enhancement is feasible, might equal-vote democracy (somewhat paradoxically) be the preferred form of political governance *because of* — rather than despite — greater interpersonal differences? After all, even though we are not equally *able*, we may be more or less equally *affected* by particular government policies, and, whatever the unequal impacts, they are not uniquely correlated to ability. To respond that impacts on the less able count for less than impacts on the more able is to presuppose a far different theory of the equality of persons *as* persons than is now held, at least in many quarters. Still, the idea that all persons are equally affected by a given kind and degree of adversity might itself be under siege in an enhancement age.

It seems unlikely that unequal allocation of votes would be seen as a realistic, efficient, and benign recognition of differences in ability, native or augmented, or of the varying impacts of political policies. It will probably be taken, correctly, as reflecting deep disrespect for those allotted fewer votes (63). And no doubt, many of those with more votes — and some with fewer — will believe that disrespect is justified given the substantial gulfs in resource-attractive or merit traits.

There are, of course, conceptual issues and "exceptions" to the one-person, one-vote standard. That standard is arguably attenuated by many institutions, such as the U.S. Senate, where states have equal votes whatever their respective populations, or special voting units such as water districts, where votes are allocated on the basis of varying rates of use or on other variables.

But in general elections, at the ballot box level, plural voting is excluded from most modern ideas of democracy. Mill himself did not necessarily endorse it over other techniques for enhancing the influence of competent elites (57,59). He seemed well aware of the substance of the Matthew Effect: Those with excess voting power may draw increasingly disproportionate shares of rewards (48,65) and possibly still more voting power, in an extended cycle. He did not endorse the ' "blind submission of dunces to men of knowledge." ' (57, p. 85) As mentioned, he also strongly emphasized participation values in democracy, which help to control government and to educate the participants, making them more competent (57). As Thompson describes Mill's resolution of the tension:

> Just as the educative benefits of participation partly justify the extension of participation, so the educative value of superior competence partly justifies the influence of a competent minority (57, p. 79).

But participation to promote competence will not necessarily save the day for one-person, one vote — particularly in an age of enhancement, with its increased and entrenched gulfs in ability, as discussed next.

Enhancement and Democratic Governance. Plural voting is a long way from dictatorship or other autocracy, but it is nevertheless likely to be taken as inconsistent with the idea of equality of persons *as* persons (60). One might thus question the seriousness of enhancement's challenge to democracy by recalling that we now maintain democratic ideals *notwithstanding* the *present* perception of *very* wide interpersonal differences. A major rationale for maintaining the one-person, one-vote regime is to prevent further consolidations of power that leave persons with inadequate access to basic commodities and opportunities.

But our commitment to democracy — and/or to various sociopolitical conditions that enable democrats to implement their commitment — might be fragile nonetheless

> Equality of control is an unstable equilibrium. Differences in knowledge, skill, opportunity and activity create inequalities of control; these in turn tend to generate further differences, which create further inequalities. [Note how this may be compounded in still further cycles by enhancement.] Hence the struggle to maintain a polyarchal organization ["[t]he main sociopolitical process for approximating (although not achieving) democracy . . ."] is never won; indeed, it is always on the verge of being lost (65, pp. 41, 282).

Enhancement and Participation. Representative democracy is not just a matter of voting rights and voting power, and elections do not confer unreviewable, irreversible delegations of authority to representatives or officials. Ideally it entails genuine opportunities for participation, in order to promote its underlying ideals of autonomy and equality, in one form or another. The sort of narrowly defined "efficiency" promoted by restricting voting and governing to superior elites is not part of the democratic canon (47). Participation is a troublesome concept to interpret: There is speaking one's piece before the appropriate representatives and government officials, there is *influencing* their exercises of power, there is having access to relevant information and ability to comprehend it, there is being a plausible candidate for office as a representative or for appointment to public office, and so on. All these aspects of democratic participation may be affected by enhancement, whatever mode of distribution of enhancement resources is selected. One's greater or lesser abilities may expand or contract one's audience, or ability to communicate with politically powerful persons and groups, or relative deliberative skills, or capacity to quickly grasp the issues of the day, or ultimate influence. Moreover, in republics we delegate responsibility for governing to others, partly from the sheer need for division of labor, partly because we want government to be run by persons capable of doing so soundly. To be nonenhanced — that is, to be relatively less capable — may be to risk exclusion from government office.

As suggested, however, the prospect of enhancement may not be fatal to egalitarian democracy, either in political theory or in fact.

First, the arguments about allocating votes as a function of competence may be somewhat misdirected. Democracy, again, is in part about having a say in what affects one. But as we saw, how much something affects you may have little or no connection to your varying competences. Moreover, to justify plural voting on our understanding of democracy, we need a moral premise concerning the proper relationship between one's political power and one's particular circumstances — including not only one's competence but one's vulnerabilities to harm under government policies. Within our present political framework, the premise is not confirmed. Representative democracy may contemplate an ideal of superbly qualified electors and even more superbly qualified representatives, but the ground for democracy is not the superior decision-making competence of the people and their delegates, as opposed to despotic rulers or elites. It rests *generally* on the unfairness and injustice of impairing autonomy by subjecting people to policies, conditions, and interactions that seriously affect them when they do not have a

voice in the matter, at *some* important level of choice. We cannot order the President to cease bombing the principality of Lower Paregoric, but we can select the President — via electors we vote for. And the ground for *equal-vote* democracy, as we saw earlier, rests partly on the unfairness of giving unequal power to persons whose vulnerabilities are likely to be quite similar, whatever their mental and physical aptitudes. Thus, the "equally affected" argument may overpower the "superior contribution" argument. Nevertheless, the possibility remains that our current notions of equal vulnerability and impact will change as enhancement technologies develop.

To turn matters around, one might urge that under given circumstances it is the *enhanced* whose participation is at risk — particularly if they are a numerical minority. But even if they are endangered by their "suspect" status, plural voting may not be the best mechanism for protecting them as compared with a strong regime of individual rights. Of course, that regime may also be impaired by a hostile majority of the nonenhanced.

Second, egalitarian democracy might survive even within the Millian framework because the available enhancements might not be seen as affecting competences relevant to democratic governance. It is not clear, for example, that moral competence can be affected in any but the most slapdash way by genetic engineering — although the possibility of doing so should not be entirely dismissed (66).

Third, as a matter of theory, it is unclear how a competence criterion for ballot power can be assessed entirely independently of certain background moral issues concerning, say, the fair/just/egalitarian distribution of goods and services. With enhancement, the "ability gulfs" between persons are themselves a partial function of preexisting wealth differences. Depending on what forms of distribution of enhancement resources were in place, these wealth differences would be unjustifiably ratified and reified by plural voting. Because Millian competence is empirically tied to wealth, which may be morally irrelevant, to defend plural voting on competence grounds thus begs some questions of moral evaluation concerning distribution and its underlying issues of equality and fairness.

Superior competence, in this context, thus remains a murky concept. Indeed, as Singer observes, "Mill himself said, later in life, that [plural voting] was a proposal which found favour with no one. The reason, I think, is not that it would obviously be unfair to give more votes to better qualified people, but rather that it would be impossible to get everyone to agree on who was to have the extra votes" (67).

Fourth, even if technological enhancement did affect relevant forms of competence, those who remain unenhanced are not "incompetent" in any sense, including Mill's. A loss of relative standing in ability or depth of learning does not entail deliberative incompetence. "Competence," at least for present purposes, arguably concerns attaining a certain threshold at least as much as it concerns the distance between oneself and others, though the two are connected. In this respect it is similar in structure to "personhood." Here a Millian might respond that enhancement could simply elevate the accepted competence threshold for qualifying as a voter, establishing a new minimal baseline for competence — but this still would not make the case for supernumerary votes.

Fifth, far from being inconsistent with equal-vote democracy, the increasing gaps between persons make it all the more desirable to retain that voting system, as suggested earlier. The less endowed and less enhanced are not likely to suspend pursuit of their own interests, despite their new relative dimness. Although the better endowed might be better able to protect themselves, given their superiority, a possible result of plural voting might be dangerous instabilities, partly because of the perceived risk of — and actual — aggrandizement of resources by the elites. The greater the fear of such risks, the more that departure from equal voting will be seen as sending us down a steep, greasy slope emptying into an abusive oligarchy — run either by the numerically inferior enhanced or by the unenhanced, each fearing domination by the other. In such a world, not only is equality compromised, but so also are all other basic values.

Turn now from equality to autonomy. (This in turn will shortly return us to equality.) What will become of *it* if enhancement is institutionalized to some degree? In parallel to the dismissal of the respect owed to the less gifted, one might urge that not only do they deserve fewer rewards, their autonomy is of lesser worth. From contemporary liberal perspectives, however, basic autonomy is not tied to one's measure of abilities, unless it falls below the general competence threshold, however defined. Yet, just as we make interpersonal comparisons of "worth" in various senses, we may in fact think that autonomy as exercised by different persons may decline in value with the declining *relative* competence of these actors. This view may have still greater pull where enhancement is practiced. Perhaps if autonomy for all is to be protected, some sort of equal-vote democracy is necessary to preserve it. As we saw, democracy might remain preferred partly because of the posited inequalities, not despite them. Still, defenders of plural voting or rule by an elite are likely to suggest that, precisely because of the elite group's superior competence, autonomy and even equality itself are better promoted by what seems like an inegalitarian system (29). It is thus hard to deny that participatory/autonomy values are at elevated risk in an enhancement context.

Sixth, perhaps the most obvious defense of equal-vote democracy is that it may be instrumental in promoting opportunities to obtain the very enhancement resources that inspired this debate about democracy's requirements — a continuation of enhancement's potential role as "remediation" of natural inequalities. There is certainly no assurance that the elites will look out for anyone's interests but their own, except on the doubtful assumption that they will also be moral elites with a strong egalitarian or altruistic bent. It bears mention at least once in this entry, despite the point's familiarity, that the result of superior competence may be greater and more successful evil.

Finally, plural voting defenders will, sooner or later, make the simple-sounding argument that there is *no*

threat to equality in an enhancement age. Equality, after all, concerns the similar treatment of similarly situated persons and the dissimilar treatment of dissimilarly situated persons. If the more able are relevantly different from the less able, treating them differently is not only not inconsistent with equality, it is required by it. The obvious response, which can only be summarily stated here, is that this claim presupposes a large set of unconfirmed and strongly contested moral propositions.

Social Changes in Attitudes Concerning Equality, Self-Regard, and Community: Symbols, Communication, and Learning. The operation and observation of our social institutions and practices generate learning effects. Present conceptions of equality and other values may eventually confront a world where long-standing assumptions about the relative stability of traits and character will be loosened. This emergence of a world in which human traits are far more controllable than now may, as suggested, drive changes in our attitudes about the demands of equality and fairness generally, and merit and desert in particular. These value shifts may occur for several reasons. For example, the consolidation of political power into hierarchies (whether or not reflected in plural voting) may result from the distribution of enhancement opportunities to those already holding wealth and power. Hierarchical institutions and practices may generate self-perpetuating learning effects through citizen participation or observation. People may come to perceive themselves and their social stations differently, perhaps as fully locked in. Enhancement may spur an increasingly intense focus on traits and their comparative measures, and magnify their apparent social and commercial value. True, we might still think that traditional enhancement enhances but that technological enhancement reduces. But whether the latter will indeed *reduce* persons to the social value of their enhanced traits or *elevate* them in a morally relevant sense is not now predictable.

CONSTITUTIONAL CONSIDERATIONS IN BRIEF

Constitutional Frameworks

In the United States, government regulation of use and distribution of enhancement technologies must be tested against claims of violating implied "fundamental liberty interests" under the due processes clauses of the Fifth and Fourteenth Amendments. (As for equal protection considerations, see the discussion below.) Federal action must also be tested against express and implied limitations on the powers of the federal government. Mention of some constitutional considerations is thus called for, and constitutional argument structures are in any event useful in discerning and addressing some of the most important issues generated by enhancement.

The right to procreate, as articulated in *Skinner v. Oklahoma* (68), might be taken to encompass at least certain forms of germ-line engineering or fetal manipulation, although the strength of such rights is open to serious question; the U.S. Supreme Court may be forced to consider a hierarchy of procreational liberty interests, each imposing a greater or lesser burden of justification for government regulatory maneuvers. The recently emerged practice of prenatal and preconception screening for disorders, which is likely to be protected as a major adjunct to procreational autonomy, might suggest parallel protection of affirmative intervention to forestall the disorders via germ-line or fetal alteration. Nevertheless, the issue is uncertain, partly because of the differences between "standard" (if technologized) reproduction and anticipated future forms: *having children at all* is not the same as *having children in certain ways, or of certain (arranged) sorts.* It is one thing to leave matters to unrevised sexual recombination, and another to affirmatively determine the traits of a specific individual.

Somatic trait augmentation — at least for competent adults — arguably ought to have greater constitutional protection than parental choice to manipulate the germ line or alter fetal development because the affected party is the decision maker. Nevertheless, it is more challenging to describe the constitutional terrain because there is no clear, recognized conceptual bin in which to place it. (Compare "procreational autonomy.") There is no general constitutional liberty or "privacy" interest embracing a right to do what you will with your body, although some commentaries, scholarly and nonscholarly, might suggest otherwise. There are recognized liberty interests of sorts in refusing various forms of medical treatment and in "personal security," which are likely to extend to forced administration of enhancement techniques, medical or nonmedical. But these doctrines do not settle matters of noninterference with voluntary use or positive assistance in securing access.

This is not to say, however, that a persuasive case cannot be made for protecting the decision whether to enhance one's basic merit attributes as an important feature of the liberty protected by the Fifth and Fourteenth Amendments; it is much too simple to assert that such textually unmentioned rights are impossible because no long-standing "tradition" protects it. There is in fact a tradition of substantially free choice in making use of changing methods of instruction, training and general pedagogy for the purpose, among others, of self-improvement. It might well be thought to extend presumptively to control of mental functions and of bodily physiology generally. [Because of the logical link between mental functioning and communication, a First Amendment argument for fair access to intellectual enhancement resources — as well as the right to refuse such resources — might also be crafted (69).] One might also urge that the liberty interest in shaping the nurture and education of one's children encompasses enhancement. The interpretive maneuvers underlying these constitutional arguments are complex and entertaining; they are described briefly below. If serious enhancement arrives on the scene, however, arguments of these sorts are certain to be offered in opposition to restricting access to augmentation services.

Under the logic of constitutional protection of liberty interests, if any of these characterizations of a right to

noninterference with enhancement decisions are successful, governments will have to justify their prohibitions and their regulatory systems generally. The weight assigned to the liberty interest will, in theory, determine how heavy these burdens of justification will be. Government will at a minimum have to identify serious interests that may be compromised by attempted or successful augmentation — for example, avoiding injuries to existing or possible persons. It will also, in theory, have to defend the precision of its means for protecting these interests. Imposing a major burden of justification on government action, for whatever reason, is the core component of a judicial decision path known as "heightened scrutiny." In recent years the U.S. Supreme Court has recognized liberty interests that apparently draw an "intermediate" level of scrutiny, rather than the maximum "strict scrutiny" standard (70), and with technological change it may well have to construct still more levels of calibrated protection.

Paths of Constitutional Interpretation

It is especially difficult to project future constitutional analysis when the transformative processes in question seem so far removed from traditional paradigms and historical understandings — assuming these matters remain constitutionally relevant. Tradition, history, original intent, and lexical understandings at the time of framing continue to be viewed as important and perhaps decisive interpretive criteria (separately or in some combination) for both explict and implicit liberty interests, although some cases have been offered as counterexamples (71). Those arguing that enhancement falls within a strongly protected liberty interest — whether as an aspect of procreational liberty, a right of personal development, or a right to control our mental and physical functions — will have varying difficulties making their case. If their characterization is rejected by the courts, then the government's burden of justification is very weak — a minimal rationality test that constrains far less than use of the term "rational" would suggest in everyday language. The difficulty in constitutional characterization of an interest is greatly compounded when the asserted interest reflects an innovation that does not seem to fit existing categories. Thus, determining whether reproductive ventures involving germ-line engineering are included within a strongly protected liberty interest will, as with any form of legal characterization, involve (inter alia) comparisons to exemplars of what is or is not protected. Partisans then characteristically state whether the interest proposed for special protection is "too far removed" or "distant" from the archetype. (This vastly oversimplifies huge interpretive issues.) The problem, however, is that the supposedly defining features of the models offered may be contested. Is the process of creating a person who didn't exist before a sufficient condition for calling the process procreation — either in common discourse or in constitutionalese? Or must the person have been created by human sexual recombination rather than asexually? Or must the person not only be the result of sexual recombination but of sexual recombination *simpliciter* (where we simply rely on the genetic lottery and avoid affirmative trait-changing, though possibly using prenatal and preconception screening, possibly followed by abortion or nonconception)? If we do not know what defines the standard example, we cannot tell how "far" we are from it. As things stand, prohibiting prenatal or preconception screening would seem to impermissibly burden procreational rights, but this does not show that forbidding germ line alteration — even for enhancement rather than disorder prevention — is also be impermissible. How "far" is affirmative genetic change in persons-to-be from prenatal or preconception testing in aid of deciding upon abortion or nonconception? All are forms of "genetic control," but germ-line alternation is vastly different, at least when viewed through prevailing constitutional frameworks. When biological technologies separate and rearrange life processes in ways not contemplated by our existing concepts, the interpretive difficulties we already face may be greatly amplified.

For now, it is enough to say that human procreation has come to vary along several overlapping axes. They concern technological facilitation of gamete union (e.g., in vitro fertilization); social arrangements (within or outside marriage; collaborative — e.g., surrogacy; use of gamete banks, whether or not for eugenic purposes); whether the efforts involve asexual methods (cloning); technological mechanisms for trait prediction; and technological mechanisms for positive control or influence over traits. It seems plausible to think that procreational autonomy extends as a *presumptive* protection — in full or near full strength to technological facilitation of gamete union (subject to limited health and kinship regulations), to procreation regardless of marital status (subject to certain protection-of-marriage limitations), and to prenatal or preconception screening for disorders, defects, or injuries. This presumptive protection might be overcome by compelling or important countervailing interests. Broader protection of collaborative procreation, and for which participants, is less certain, although strong coverage is likely for gamete donation or sale (at least for a modest price to avoid charges of "economic coercion") by persons within or outside the intended nuclear family. The constitutional fate of human cloning is seriously in doubt because of the perception by many that asexual reproduction is a truly radical departure from standard procreation, and does not belong within protected constitutional categories. Although germ line enhancement within sexual reproduction is a striking departure from standard reproduction because of its partial nullification of the genetic lottery, it seems likelier than cloning to be assigned some serious presumptive protection, at least within a disorder model. (Cloning may not go entirely unprotected, however.) As noted, however, the strength of "compelling" or "important" governmental interests can in theory override the individual rights claim, if the regulations are carefully tailored to further those interests so as to reduce intrusions on constitutionally protected interests.

Constitutional Equality Standards

There may also be questions concerning the status of the enhanced or the nonenhanced as members of discrete, identifiable groups at risk for discrimination and

exploitation. If so identified, classifications concerning the group may be treated as "suspect" to some degree under the Fourteenth Amendment's equal protection clause and the Fifth Amendment's implied parallel protection. This will again trigger heightened scrutiny and, in theory, impose a nontrivial burden of justification on government action. If this "suspectness" characterization fails, the government is likely (but not certain) to prevail (72). The point here is that many egalitarian claims find little or no purchase within the constitutional framework of equality, which offers strong protection against certain forms of discrimination (racial, ethnic, gender, etc.), modest protection against certain forms of classificational irrationality involving vulnerable groups (it is hard to predict which groups will be considered vulnerable), and for all practical purposes no protection for any other form of classification. In egregious cases of abuse or manipulative control over enhanced or unenhanced persons, however, one might claim violation of the Thirteenth Amendment (banning slavery) or the Nobility Clause (Article I, §9) (73). Both provisions are heavily inspired by considerations of equality.

Congressional Powers

Congress has an uncertain range of powers to promote constitutional rights under Section 5 of the Fourteenth Amendment (and parallel provisions in other amendments), subject to Supreme Court control. If any group — nonenhanced or enhanced — seems especially put upon, Congress may consider remedial legislation (perhaps as a form of "affirmative action") (37). Congress also retains considerable powers to protect or promote constitutional rights under the commerce clause, and the taxing and spending powers.

CONCLUSION

Dealing with technological enhancement and its impact on basic values is beset with the usual problems associated with value analysis — vagueness, ambiguity, "open texture," indeterminacy, and collision with other values. But these problems are aggravated because the new powers seem to undermine assumptions concerning our understanding of these values. Determining just how crucial these assumptions are to the tasks of moral and legal evaluation of enhancement technologies forms a major portion of the analytical work required.

The most obvious assumption being tested, of course, is that we are severely limited in altering native traits — including our most valued merit attributes and resource attractors — by the constraints of our individual genetic endowments and by the very nature of familiar and slow-working tools of self-improvement: study, training, practice, effort, self-discipline.

It now appears, however, that technological intervention via the germ-line and somatic mechanisms will eventually allow us to alter, at least in certain ways, the limits of what we now view as relatively fixed potentials for improvement. Today, with extended study and practice as he grows up, Forrest Gump can learn to make change, to balance checking accounts, and to do some algebra, but quantum gravity will forever elude him. Later, perhaps such limits will no longer hold: evidence of the possibility of serious and accelerated trait changes seems to be growing. Does one's merit, virtue, and ultimate desert rest only on traditional paths toward personal progress?

BIBLIOGRAPHY

1. M.H. Shapiro, *Wake Forest Law Rev.* **34**, 561 (1999).
   M.H. Shapiro, *S. Cal. Law Rev.* **65**, 11 (1991).
   M.H. Shapiro, *S. Cal. Law Rev.* **48**, 318 (1974).
2. D. Rae, *Equalities*, Harvard University Press, Cambridge, 1981, Chap. 4.
3. D.W. Brock, A. Buchanan, N. Daniels, and D. Wikler, *From Chance to Choice: Genes and the Just Society*, Cambridge University Press, Cambridge, 2000.
4. J. Travis, *Science News* **156**, 149 (1999).
   K. Howard, *Sci. Am.* **281**, 30 (1999).
5. Associated Press, *New York Times*, June 3, 1998, A19; T.H. Maugh II, Los Angels Times, February 13, 1997, B2; 1997 WL 2181814.
6. C. Holden, *Science* **280**, 681 (1998).
7. E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, 1998, p. 66.
8. G. Sher, *Desert*, Princeton University Press, Princeton, 1987.
9. L.M. Fleck, in T.F. Murphy and M.A. Lappé, eds., *Justice and the Human Genome Project*, University of California Press, Berkeley, 1994, pp. 133, 143.
10. D.W. Brock, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, 1998.
11. D.B. Allen and N.C. Fost, *J. Pediatrics* **117**, 16, 19 (1990).
12. A. Silvers, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, 1998.
13. Times Wire Reports, *Los Angeles Times*, September 9, 1999, A14; 1999 WL 26173732.
14. R. Cole-Turner, in E. Parens, ed., *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, 1998.
15. E.T. Juengst, in E. Parens, ed., *Enhancing Human Traits. Ethical and Social Implications*, Georgetown University Press, Washington, 1998, pp. 29, 38.
    D.M. Frankford, in E. Parens, ed., *Enhancing Human Traits*, Georgetown University Press, Washington, 1998.
16. P.H. Huang, *Wake Forest Law Rev.* **34**, 639 (1999).
17. L.S. Temkin, *Inequality*, Oxford University Press, New York, 1993.
18. N.H. Hsieh, *J. Polit. Phil.* **7**, 7–4–1.1 (1999).
19. M.J. Mehlman and J.R. Botkin, *Access to the Genome: The Challenge to Equality*, Georgetown University Press, Washington, 1998, pp. 62–85.
20. Americans With Disabilities Act of 1990, 42 U.S.C. §§12101–12213 (1994 & Supp. II 1996).
21. J. Rawls, *A Theory of Justice*, Harvard University Press, Cambridge, 1971.
    R. Nozick, *Anarchy, State, and Utopia*, Basic Books, New York, 1974.

22. R.J. Sternberg, *Beyond IQ: A Triarchic Theory of Human Intelligence*, Cambridge University Press, New York, 1985.
23. K. Vonnegut, *Welcome to the Monkey House*, Dell Publishing, New York, 1950, p. 7.
24. B. Goodwin, *Justice by Lottery*, University of Chicago Press, Chicago, 1992.
M. Waldholz, *Wall St. J.*, June 21, 1995 (1995 WL-WSJ 8730611).
G.J. Annas, *Am. J. Law Med.* **3**, 59–76 (1977).
25. A.M. Bickel, *Harv. Law Rev.* **75**, 40, 76 (1961).
26. S. Darwall, *Philosophical Ethics*, 1998.
27. E. Parens, *Kennedy Inst. Ethics. J.* **5**, 141, 146–147 (1995).
28. P. Westen, *Harv. Law Rev.* **95**, 537 (1982).
29. I. Berlin, in A. Quinton, ed., *Political Philosophy*, Oxford University Press, London, 1967.
G. Dworkin, *The Theory and Practice of Autonomy*, Cambridge University Press, Cambridge, 1988.
30. R.A. Dahl, *Democracy and Its Critics*, Yale University Press, New Haven, 1989.
31. M.J. Mehlman, *Wake Forest Law Rev.* **34**, 671 (1999).
32. J. Mincer, in C. Kerr and P.D. Staudohar, eds., *Labor Economics and Industrial Relations: Markets and Institutions*, Harvard University Press, Cambridge, 1994.
33. R. Sartorius, ed., *Paternalism*, University of Minnesota Press, Minneapolis, 1983.
34. A. Wertheimer, *Coercion*, Princeton University Press, Princeton, 1987.
35. J. Bartlett, *Familiar Quotations*, Little Brown, Boston, MA, 1980.
36. B. Edlund, *Reuters Library Report*, September 27, 1988, p. 1.
37. J.B. Attanasio, *U. Chi. Law Rev.* **53**, 1274 (1986).
38. F. Herbert, *Dune*, Chilton Books, Philadelphia, 1965.
39. S. Gardbaum, *Harv. Law Rev.* **104**, 1350 (1991).
40. E.J. Keenan, in J.A. Thomas, ed., *Drugs, Athletes, and Physical Performance*, Plenum Medical Book, New York, 1988.
41. D. Keyes, *Flowers for Algernon*, Harcourt Brace and World, New York, 1966.
P. Anderson, *Brain Wave*, Ballantine Books, New York, 1954.
42. R.M. Berry, *Wake Forest Law Rev.* **34**, 715, 729–731 (1999).
43. N. Daniels, in T.F. Murphy and M.A. Lappe, eds., *Justice and the Human Genome Project*, University of California Press, Berkeley, 1994.
44. D.J. Schneider, A.H. Hastorf, and P.C. Ellsworth, *Person Perception*, 2nd ed., McGraw-Hill, New York, 1979, pp. 166–169, 267–269.
45. T.E. Hill, Jr., *Dignity and Practical Reason in Kant's Moral Theory*, Cornell University Press, Ithaca, 1992, pp. 38–39.
46. M.J. Radin, *Harv. Law. Rev.* **100**, 1849, 1933–1936 (1987). Contested Commodities 102–114 (1996).
M.H. Shapiro, *Hastings Law. J.* **47**, 1081, 1180–1199 (1996).
47. M.H. Shapiro, *U. Pitt. Law. Rev.* **55**, 681, 686–687, 765–769 (1994).
48. H.L.A. Hart, *Harv. Law. Rev.* **71**, 593, 622 (1958).
C.R. Beitz, *Political Equality: An Essay in Democratic Theory*, 1989, p. 35.
49. R.K. Merton, *Science* **159**, 56 (1968).
50. L.J. Cappon, ed., *The Adams-Jefferson Letters*, University of North Carolina, Chapel Hill, 1959, p. 388.
51. J.L. Borges, *Ficciones*, Grove/Atlantic, New York, 1962, pp. 107, 115.
52. R.J. Herrnstein, *I.Q. in the Meritocracy*, Atlantic Monthly Press, Boston, 1973.
R. Herrnstein and C. Murray, *The Bell Curve: Intelligence and Class Structure in American Life*, Free Press, New York, 1974.
Richard Stone Science, ed., **267**, 779 (1995).
G. Leach, *The Biocrats*, rev. ed., Pelican, Harmondsworth, 1972, pp. 221–223.
M. Young, *The Rise of the Meritocracy*, Transaction, New Brunswick, 1994.
53. A. Posner, *J. Law Econ. Org.* **13**, 344 (1997).
54. T.L. Beauchamp and J.F. Childress, *Principles of Biomedical Ethics*, 4th ed., Oxford University Press, New York, 1994, pp. 120–188.
55. M.H. Shapiro, *U. Pitt. Law. Rev.* **55**, 681, 772–774 (1994).
56. Plato, *The Republic*, Benjamin Jowett, trans., Heritage Press, New York, 1944.
57. D.F. Thompson, *John Stuart Mill and Representative Government*, Princeton University Press, Princeton, 1976.
58. M.H. Morales, *Perfect Equality: John Stuart Mill on Well-Constituted Communities*, Rowman and Littlefield, Lanham, 1996, p. 86.
59. J. Stuart Mill, in J.M. Robson, ed., *Considerations on Representative Government*, University of Toronto Press, Toronto, Buffalo, London, 1977, pp. 371, 475.
60. J. Peters, *Colum. Law. Rev.* **97**, 312, 334–336 (1997).
J. Waldron, *Geo. Law. J.* **84**, 2185, 2211–2212 (1996).
R.J. Arneson, *J. Hist. Phil.* **20**, 43, 59–62 (1982).
61. R. Sinsheimer, in *Limits of Scientific Inquiry*, Daedalus, Spring, 1978, p. 34.
62. *Rogers v. Lodge*, 458 U.S. 613 (1982).
63. J. Cohen and C. Sabel, *Eur. Law. J.* **3**, 313 (1997).
64. N. Lemann, *New York Times*, April 26, 1998, §4, p. 15.
N. Lemann, *The Big Test: The Secret History of the American Meritocracy*, Farrar, Straus and Giroux, New York, 1999.
65. R.A. Dahl and C.E. Lindblom, *Politics, Economics, and Welfare*, University of Chicago Press, Chicago, 1976, p. 282.
66. H.T. Engelhardt, Jr., *Soc. Phil. Pol.* **8**, 180, 186–189 (1990).
67. P. Singer, *Democracy and Disobedience*, Oxford University Press, New York, 1973, pp. 34–35.
68. *Skinner v. Oklahoma*, 316 U.S. 535, 541 (1942).
69. M.H. Shapiro, *S. Cal. Law. Rev.* **47**, 237, 256–257 (1974).
70. E. Chemerinsky, *Constitutional Law: Principles and Policies*, Panel Publishers, New York, §6.5, 1997.
71. *Roe v. Wade*, 410 U.S. 113 (1973).
Brown v. Board of Education, 347 U.S. 483 (1954).
72. *Plyler v. Doe*, 457 U.S. 202 (1982).
73. F.C. Pizzulli, *S. Cal. Law. Rev.* **47**, 476 (1974).

See other entries BEHAVIORAL GENETICS, HUMAN; see also HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY entries.

ENCYCLOPEDIA OF

ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

VOLUME 2

WILEY BIOTECHNOLOGY ENCYCLOPEDIAS

Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation
Edited by Michael C. Flickinger and Stephen W. Drew

Encyclopedia of Molecular Biology
Edited by Thomas E. Creighton

Encyclopedia of Cell Technology
Edited by Raymond E. Spier

Encyclopedia of Ethical, Legal, and Policy Issues in Biotechnology
Edited by Thomas H. Murray and Maxwell J. Mehlman

ENCYCLOPEDIA OF ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

EDITORIAL BOARD

ENCYCLOPEDIA OF

ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

VOLUME 2

Thomas H. Murray
The Hastings Center
Garrison, New York

Maxwell J. Mehlman
Case Western Reserve University
Cleveland, Ohio

A Wiley-Interscience Publication
John Wiley & Sons, Inc.
New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

This book is printed on acid-free paper. ♾

Published simultaneously in Canada.

For ordering and customer service, call 1-800-CALL-WILEY.

Library of Congress Cataloging in Publication Data:

Murray, Thomas H. (Thomas Harold), 1926–

Encyclopedia of ethical, legal, and policy issues in biotechnology / Thomas Murray, Maxwell J. Mehlman.

p. cm.

Includes index.

ISBN 0-471-17612-5 (cloth : alk. paper)

1. Biotechnology — Moral and ethical aspects — Encyclopedias. 2. Biotechnology industries — Law and legislation — Encyclopedias. 3. Biotechnology — government policy — United States — Encyclopedias. I. Mehlman, Maxwell J. II. Title.

TP248.16.M87 2000

174′.96606 — dc21 00-021383

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

ENCYCLOPEDIA OF

ETHICAL, LEGAL, AND POLICY ISSUES IN BIOTECHNOLOGY

VOLUME 2

H

HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY: OVERVIEW

ROBERT WACHBROIT
University of Maryland
College Park, Maryland

OUTLINE

INTRODUCTION

There is little doubt that the most controversial issue regarding biotechnology is the prospect of employing it for the purpose of human enhancement. The following discussion is intended to serve as a roadmap to the various questions and topics raised by the prospect of enhancement, with special emphasis on the conceptual issues. A detailed examination of the ethical issues is the topic of another article.

We will begin by examining the so-called demarcation problem: What is enhancement and what is it being contrasted with? We will then survey some of the types of modifications that lead to enhancement. Although the primary modification people have in mind is genetic, it is worth looking at nongenetic modifications — biotechnological and nonbiotechnological — in order to place the concerns with genetic modification in a broader landscape. Finally, we will examine some general approaches for assessing genetic enhancement.

THE DEMARCATION PROBLEM

Enhancement modifications are typically defined alongside therapeutic modifications. A therapeutic modification is one that brings a trait that was below a recognizable, specieswide norm up to that norm. (The term "traits" is meant in its broadest sense, including physical attributes, mental or physical abilities, dispositions, and capabilities.) As a first approximation, we can characterize an enhancement modification in contrast as one that is a nontherapeutic improvement. The norm referred to here is the one that separates conditions of health from those of disease. The distinction between enhancement and therapy is therefore linked to the distinction between health and disease.

Two important points should be raised about this linkage. First, while the various proposed theories of health and disease yield corresponding accounts of what enhancement and therapeutic modification means, controversies and obscurities in the former will translate into the latter. Problems with particular theories of health will have counterparts in problems with the corresponding account of enhancement modification. Indeed, skepticism about there being an objective contrast between health and disease will translate into a corresponding skepticism about the distinction between enhancement and therapy. Consequently the health/disease distinction is of limited use in *explaining* the enhancement/therapy distinction. None of this, however, undermines the link. The first distinction will be as clear and useful as the second. Thus, while it is true to say that therapeutic modifications attempt to treat disease whereas enhancement modifications attempt to improve a trait that is not diseased, there can be considerable debate over whether a particular modification therefore constitutes an enhancement and why.

Perhaps the most debated issue regarding theories of health and disease is whether or not the distinction — what constitutes the norm — is value-free: Is the judgment that someone is diseased — that someone's condition falls below a norm — an objective discernment of a biological state or a value judgment? Is a particular condition a disease independent of whether we think it is bad or undesirable? Can a condition be a disease in one culture and not in another? Although this dispute is not particularly salient in discussions over enhancement, these discussions have typically proceeded with the idea of the norm being fixed and not relative to individuals or cultures. If what constitutes enhancement varies with individual or culture — if enhancement is in the eye of the beholder — then it is not clear that we can sensibly articulate an (ethical) issue about enhancement as such. Nonrelativistic conceptions of normality tend to favor objective theories of health and disease, though that still leaves considerable latitude over how to conceive of normality, from statistical conceptions (1) to biological conceptions (2). Nonetheless, objective conceptions are not the only kind of nonrelativistic conception of normality. Norms that are recognized to be arbitrary and conventional can still frame the issue, as discussions over the problems of enhancements in sports demonstrates. Indeed, even a normative conception of norms could be invoked, as long as the relevant values are themselves understood to be nonrelativistic.

The second point to note is that acknowledging the link between health/disease and enhancement/therapy can suggest that the latter is a medical matter. Ethical issues regarding enhancement modification should then be seen in terms of the ethics of medicine and the professional duties and responsibilities of health professionals. As plausible as this suggestion may be, we need to distinguish at least theoretically between questions regarding the ethics of enhancement modifications and questions regarding the ethics of physicians performing enhancements modifications — for example, whether a particular enhancement modification is ethically objectionable from whether it is unethical for a physicians to perform such a procedure. It may well be that the answer to the second determines the answer to the

first — for example, that the ethical questions regarding enhancements comes down to questions about the role morality and professional ethics of physicians, but a claim like that requires more of an argument than pointing to the connection between the enhancement/therapy distinction and the health/disease distinction (3, Introduction).

An immediate challenge to any account of the enhancement/therapy distinction is the presence of apparent borderline cases or exceptions to the classification scheme. These fall into two classes.

One class of cases is modifications that, strictly speaking are enhancements, but whose purpose is to respond to (the threat of) a disease. For example, a modification that improves people's resistance to particular diseases beyond the normal capacity would count as an enhancement but its purpose would be disease prevention and so arguably therapeutic.

A different class of borderline cases or exceptions arises from an ambiguity in the idea of "normal traits." It can mean a trait whose appearance and function is normal, but it could also mean a trait whose appearance, function, *and* development is normal. Moreover normality itself often refers to a range within a trait rather than to a sharp line. Thus, imagine two people, both of whose height is five feet. While the first person has short parents, the second has tall parents but suffers from a growth disorder. Both people have a height that falls within the normal range, but the second person's height is the result of a disease. A modification that brought the second person's height from five to six feet would be a modification within the normal range *and* a response to a disease (4).

Both types of cases indicates that a classification scheme generated by outcomes *and* by purposes is a scheme driven by one too many criteria. Since there is no direct line between outcomes and purposes, we should not be surprised if there are cases that fit each criterion differently. Which criterion we should use will depend on what we are trying to classify. If we are trying to classify modifications, then outcomes would be the better choice; if we are instead trying to classify practices or aims, then purposes might be a better criterion. Using both criteria is inevitably confusing in that not only can the same procedure and outcome be associated with different purposes but the same event often is shaped by multiple purposes. Ambiguities in classifying particular cases will inevitably arise. (See Juengst's article in Ref. 3 for a discussion of various other alternatives and their problems.)

The point about normality not being a sharp line raises an important distinction within the category of enhancement modifications: There could be modifications that raise a trait above the norm and there could be modifications that raise a trait from one point within the normal range of that trait to a higher point in that range. This suggests that the classification of modifications should be tripartite: therapeutic, intranormal, and (proper) enhancement. Nevertheless, we should note that many commentators understand enhancement to mean any improvement of a normal trait, thereby collapsing the second and third categories. Cosmetic surgeries, which can often be regarded as intranormal modifications, are thus placed in the same category as genetic modifications to create superpeople. Whether fewer distinctions or categories is better will depend on how the issues are analyze and whether one classification clarifies matters more than the other. As we will suggest below, it is better to keep intranormal modifications, which are differences of degree, distinct from enhancements proper, which are differences of kind.

TYPES OF MODIFICATIONS

Biotechnology covers a range of technologies and procedures, many of which could conceivably be employed for enhancement. Drugs could be designed to interact with the body's chemistry in such a way as to alter behavior, biological functioning, structure, or affect. Even without introducing drugs, special procedures — such as transfusing a person with their own blood or "blood doping" — can affect traits or behavior. But the most discussed enhancement technology is one in which a person's genome is altered.

It is an empirical question which traits can be enhanced by modifying an individual's genes. And it may turn out that enhancing certain traits requires not only genetic modifications but also certain alterations in the individual's environment. That is to say, a particular genetic modification might not by itself bring about an enhanced trait; it might give the person a capacity to developed the enhanced trait whose realization demands a special exercise regime, diet, or other efforts. The idea of genetic enhancement technologies therefore does not rest on an assumption of genetic determinism — that a genetic alteration alone is sufficient to bring about a particular trait. While a popular image of genetic enhancements is that of some magic-wand transformation in which the person is a passive recipient, the matter can be more complex. Realizing a genetic enhancement might involve hard work. This point will become important later when we consider assessments of enhancements.

Genetic modifications are often separated into two kinds — somatic and germ line. The difference is whether the particular genetic modification affects the individual's gametes so that the modification can be passed on to the individual's offspring. The object of a somatic modification is a modified individual, but the object of a germ-line modification is a modification that becomes part of the individual's legacy or inheritance. In saying that there are these two kinds of genetic modifications, we are *not* claiming that of any particular genetic enhancement there is a somatic version and a germ-line version. That is entirely an empirical matter. It may well be that certain kinds of enhancements can only be done as somatic while others can only be done as germ line. For example, a modification may only be somatic because it interferes with the individual's ability to reproduce. A modification may only be germ line because the only feasible way of delivering the modification to all the relevant cells requires inserting the modification in the few cells of the embryo stage, which would then likely affect the individual's germ cells. Although the distinction between somatic and germ-line modification is conceptually clear, it may not be applicable everywhere.

Nevertheless, many commentators find the distinction useful. It is reasonable to assume that somatic

enhancements are simpler as far as ethics and public policy is concerned. Germ-line enhancements appear to raise all the issues of somatic enhancement and then some. And so it would seem that we should first examine the acceptability of somatic enhancements and only after settling that should we proceed to an examination of germline enhancements. This strategy however is difficult to sustain if we allow for the possibility that enhancements might not stand or fall as a group. If we are open to the possibility that some enhancements might be acceptable and others not and we acknowledge that some enhancements may have only a somatic or a germ-line version, the strategy of considering first the somatic case and then the germ-line case may not always be applicable. Some enhancements might as a matter of technology not have a somatic version.

(Of course there is one way of ensuring a genetic enhancement is not in effect germ line — combining the modification with one that also renders the individual infertile. But this possibility is probably not worth dwelling on: It is difficult to conceive of a case where, as a matter of ethics or public policy, an enhancement would be acceptable but only if the individual agrees to sterilization.)

ASSESSING ENHANCEMENTS

There are two broad approaches to determining the acceptability of an enhancement: the assessment can be based on what the enhancement is — the product — or on how the enhancement is achieved — the process. We will briefly provide an overview of the questions each of these raise in turn, leaving a more detailed discussion of the ethics for a separate article. For this entire discussion, we will be assuming that the modification is safe and effective so as to target our inquiry on the acceptability of the enhancement modification itself rather than on side issues regarding the acceptability of risky biotechnological procedures.

Product Assessment

Recall the earlier contrast between enhancements, properly called, and intranormal modifications. An assessment that focuses only on products will pass on intranormal modifications, since there is presumably nothing wrong with the result of an intranormal modification, at least at the level of the individual. For example, there is nothing wrong in itself with being six feet tall, so an intranormal modification that renders a person six feet tall cannot be unacceptable because of the result.

Confining ourselves therefore to only proper enhancements, the first kind of product assessment is directed at trade-offs the modification allegedly imposes. For example, suppose that a modification enhances a people's memory capacity but with the result that the speed in accessing memory is considerably slower. Or suppose that an enhanced memory capacity results in greater irritability. How should these trade-offs be assessed? Should it be a matter of individual choice or public policy?

A second kind of product assessment is directed at the "humanity" of the modification. According to this approach there is something wrong in itself and not because of alleged trade-offs in having a particular trait enhanced beyond what is (normally) human. Indeed, because of the enhanced trait, the individual might not be regarded as human. The problems arising from the various racial and ethnic divisions of humanity might well carry over to this new kind of division. In addition some people might regard enhanced individuals as an insult to the integrity of the species or, seen religiously, an insult to God's creation. How should these concerns be addressed in a pluralistic society?

A third kind of product assessment is directed at the widespread use of enhancements. Even if there is nothing wrong with any particular enhancement use, problems arise when many people or certain sectors of the population primarily make use of this enhancement. For example, suppose genetic enhancement of memory were possible and it resulted in memory-enhanced individuals being significantly more successful in several aspects of life. If only the wealthy had access to this technology, genetic enhancements would create or exacerbate troublesome inequalities. But even if the technology were made available to everyone, problems could arise. The desirability of some traits arguably rests on their not being common or widespread; if everyone is a blonde, then blondes will not have more fun. Furthermore people who do not want to be enhanced might nevertheless feel under some considerable pressure to avail themselves of it because many other people are doing so.

Process Assessments

When the issue becomes the process, then intranormal modifications are as much a subject for examination as enhancements, properly called. Indeed, it might be argued that we only need to examine intranormal modifications: If an intranormal modification is unacceptable from the standpoint of process, it would seem that extending that modification to the point of enhancement would also be unacceptable from the standpoint of process. The obverse would also seem to be true, though we should emphasize that acceptability from the standpoint of process does not entail acceptability from the standpoint of product.

The first kind of process assessment is directed at the suggestion that using biotechnology to effect an improvement is wrong because it is artificial. This concern need not be one that crudely equates natural with good and nonnatural with bad, raising concerns even about ordinary medical interventions. The worry here is a "commodification" of certain traits, and the people who have them, because of their being made to order, so to speak. This type of assessment is often linked with the concern about the humanity of the modification mentioned above.

The second kind of process assessment arises from a concern that using biotechnology in order to effect an improvement undermines the value of the improvement. The value we place on certain achievements may depend on the struggle and effort required to achieve them. If they could be made effortless — at least on the part of the individual — and common, we might well cease to value them. As we noted earlier, some (genetic) enhancements may only result in enhanced capacities; realizing them may still require effort, discipline, and luck on the part of the individual. Is the kind of effort relevant to the value we place on certain achievements?

The third kind of process assessment is directed at the suggestion that using biotechnology to enhance people is not the sort of thing physicians should do. The values or aims of the medical profession are held to be incompatible with performing enhancements. This can be a parochial concern in that a judgment that physicians should not perform enhancements leaves the question of the ethics of enhancement untouched. One can consistently be a supporter of capital punishment and yet hold that physicians should not be involved in either administering lethal injections or making the official pronouncement of death. In order to make this type of assessment have broader significance, one must argue that any profession that provides enhancements has suspect aims or values.

BIBLIOGRAPHY

1. C. Boorse, *Philos. Sci.* **44**, 542–573 (1977).
2. R. Wachbroit, *Philos. Sci.* **61**, 579–591 (1994).
3. E. Parens, *Enhancing Human Traits: Ethical and Social Implications*, Georgetown University Press, Washington, DC, 1998.
4. D. Allen and N. Fost, *J. Pediatrics* **117**, 16–21 (1990).

See other entries BEHAVIORAL GENETICS, HUMAN; GENE THERAPY, ETHICS, GERM CELL GENE TRANSFER; GENETIC DETERMINISM, GENETIC REDUCTIONISM, AND GENETIC ESSENTIALISM; see also HUMAN ENHANCEMENT USES OF BIOTECHNOLOGY entries.

HUMAN GENOME DIVERSITY PROJECT

HENRY T. GREELY
Stanford University Law School
Stanford, California

OUTLINE

INTRODUCTION

The Human Genome Diversity Project (HGDP), first proposed in 1991, has thus far been responsible for more controversy than research. It has raised many of the same concerns as the well-established Human Genome Project (HGP) but at the level of human groups rather than that of individuals. As of the date of this article, the future of HGDP remains uncertain, but it is certain that the study of human population genetics, with its implications for human groups, will continue. The history of HGDP and the discussions of ethical, legal, and political issues it has stimulated are a source of useful lessons in either case, lessons about the complexity of social consequences of genetic research on human groups.

HGDP

HGP plans to publish the nucleotide sequence of the human genome in the next few years. The underlying point to HGDP is that there is no one human genome; instead there are about six billion existing human genomes, one for every living member of *Homo sapiens*, each of whose genome is separate and at least slightly distinct from all others. (Even monozygotic, "identical," twins, who make up under 1 percent of our species will show some genetic differences caused by mutations during development.) HGDP seeks to create a resource for studying this diversity. Its goals, which are set out most clearly in the report of its founding meeting in September 1993 in Alghero, Sardinia, include collecting and preserving genetic samples from approximately 500 different human populations around the world, performing some genetic analysis of the samples, and making both the results of those analyses and portions of the samples themselves available to interested researchers (1). HGDP, and the scientists behind the project have pursued these goals with limited success since 1991.

Scientific Background

The human genome is made up of approximately three billion base pairs of DNA, spread over 46 chromosomes. Individual human genomes do not vary by much. Genomes from two people, from anywhere in the world, vary on average at about one base in a thousand along any given stretch of their DNA. Within the regions of the genome that code for protein, the variation is closer to one in ten thousand. Most human genetic variation falls in regions of the genome that have no known function. Variations in these regions may be completely without consequence. Even variations that fall within the coding region of genes may be unimportant, either because they do not change the protein product, as when a single nucleotide substitution does not change the amino acid coded for, or because they change in the protein product in ways that make no apparent difference — that do not change the individual's phenotype. Other variations will affect the person's phenotype: in ways that may be negative, such as a genetic disease; in positive ways, such as increased resistance to a disease; or in ways that, as far as can be seen, are neutral, such as eye color.

The consequences of these variations will often depend on the individual's environment; having one copy of the gene for sickle cell anemia may cause mild health problems but provides some protection against malaria. In areas where malaria is rare, it may be a disadvantage; in areas where malaria is common, an advantage. The study of variations in individual genomes that have phenotypic consequences is the traditional study of human genetics.

If humans did not vary genetically, classical genetics could not see anything. Hair color, eye color, blood type, strongly genetic diseases — all these genetic traits are linked to the existence in individuals of genetic variations.

Particular genetic variations, or genetic "markers," are not distributed randomly among the world's peoples. The percentage of any human group with a particular genetic variant, or marker, may differ from the percentage in another group. The genetic contribution to skin coloration, for example, clearly varies among humans in ways that correlate, albeit imperfectly, with culturally defined ethnic groups. The same is true of other genetic variation, some of which has observable consequences, such as blood types, and most of which does not. The study of the patterns of genetic variation among human groups is called human population genetics. After the rediscovery of Mendel's work at the beginning of the twentieth century it became possible to study human variation at the genetic level even though human genes, at that point, remained abstractions. The key was to find some observable physical characteristics that were inherited in the manner described by Mendel and hence, presumably, determined by genes. By observing the variations in those characteristics, one was observing variations in the underlying genes.

This research began around 1915 with studies of the prevalence of ABO blood groups in different populations (2,3). And different patterns were discovered (4). Type A blood was most common in Northern Europe (although not found in a majority of the population even there). The peoples of Southern and Eastern Europe had a higher percentage of type B blood than those of Northern Europe, although, again, only a minority of them had type B blood. Native Americans overwhelmingly carried type O blood. The classification of humans into different *biological* groups based on their observable physical variations had been undertaken in Europe at least since Linneaus began biological classification of animals in the eighteenth century. The usual result was "scientific proof" that Europeans were "biologically superior." Human population genetics, which appeared to offer a way of classifying human groups based on characteristics that were not environmental, could be enlisted into such an effort. Thus some Nazi propaganda talked of the importance of blood type A Nordic peoples holding back the flood of blood type B blood from the "inferior" peoples of south and east.

In fact, over time studies of human genetic variation using classical markers have revealed that humans are not genetically very distinct and that most of the variation that exists is found within human populations. An estimated 85 percent of all human genetic variation exists within populations; only 15 percent reflects statistical differences between populations (5,6). Where differences among groups exist, they are usually *not* in the presence or absence of particular variations, but in their frequency. These differences are the products of statistical analysis and meaningful only for the groups, not for the individuals within them. For example, all human populations seem to include some members with each of the ABO blood types: O, A, B, and AB. Around the world, the A blood type is found in about 30 percent of humans. In some populations, particularly in the Americas, it is found in only a few percent. Among Armenians, it is found in just under 50 percent of the population. Any one person with an A blood type is highly unlikely to be Armenian; any one Armenian is more likely than not to carry some other blood type. If, however, 50 percent of a town of 3000 people have the A blood type, the inference that most of the town is Armenian — or made up of people related to the Armenians — may be worth investigating.

For most of the twentieth century, a shortage of classical genetic markers — phenotypical variation that is inherited in a Mendelian manner — limited direct empirical research into human genetic diversity. The theory of population genetics, however, both human and nonhuman, fruitfully expanded through the middle of the twentieth century. Then, the finding by Avery that genes were made of DNA and the discovery by Watson and Crick of the structure of DNA led ultimately to an era where much more genetic variation was much more directly observable through analysis of variations in the DNA itself. This not only allowed examination of more genetic traits but permitted for the first time the examination of genetic variation in regions of the genome that had no effects on phenotype. As these regions contained the vast bulk of human genetic variation, the ability to study such variation was greatly enhanced.

Such studies could have many uses. Perhaps the most interesting would be to provide additional evidence concerning human history and evolution. It will rarely be the case that one bit of genetic variation will be very informative, but by analyzing the patterns of variations in many different locations in the genome, population geneticists can, in some cases, estimate the closeness of the relationships between human groups. Eventually they may be able to generate a "phylogenetic tree," showing the relationships between human populations as putative descendants of common ancestors. This kind of information could be used as evidence of human migrations. The evidence can be used for very general questions, such as testing the "Out of Africa" hypothesis of human history, or for very narrow ones, such as exploring the history of the Japanese. The migrations studied might be recent, as in the Native American movements to the Great Plains after the acquisition of the horse made bison a more easily exploited resource. They can be more distant, such as, perhaps, the spread of Polynesians through the Pacific, Bantu-speakers through sub-Saharan Africa, or Indo-European speakers through Europe. In some of these cases, the evidence may be negative, showing that these changes in culturally defined groupings involved changes in cultures but did not result from migrations of genetically related people, which would itself be an interesting finding. Or the migrations can be still more distant, such as the evolution and spread of humans across the globe from their African origins.

This general approach of using variations to trace history is neither novel nor foolproof. The same kind of approach has been used to trace the history and changes of different texts, the development of languages, and, with molecular and nonmolecular evidence, the evolution of many different living things. This phylogenetic approach

may prove particularly difficult with humans, who did not separate into reproductively isolated populations. Genetic variations have flowed between human groups prodigiously in the last few centuries; it is certain that some level of gene flow between many groups has long existed. Today's culturally defined "populations" may not have substantial genealogical, and hence genetic, connections (7,8). The absence of such connections, though interesting on its own account, would undercut the historical value of population genetics (and, as noted later, cause some ethical problems). Genetic evidence of history and evolution is one line of evidence, to be considered with historical, ethnographic, linguistic, archaeological, and other kinds of evidence. It is, in the abstract, no more or less powerful or accurate than any of the others. But it does offer a different, and independent, line of evidence, which is of great interest.

The study of human genetic diversity has other uses. It may be used to answer some questions in cultural anthropology, such as patterns of marriage in the caste system in India. It might be used to test more general propositions in population genetics — after all, we know far more about the history and mating habits of *Homo sapiens* than we know about *Drosophila* in the wild. Or it could be of some help in biomedical research. If, for example, a variation in a "candidate gene" is identified as linked to a particular disease, the ability to check the prevalence of that variation in a population where the incidence of the disease is known might provide useful hints for further research.

A final, and more symbolic, value concerns the definition of the "normal" human genome. Genetic variation within the human species is small — especially variation that clusters along population lines — but it does exist. If "the" human genome were defined by the genetic variations most common in the countries where HGP is taking place, it would greatly overrepresent people of European ancestry, who make up less than 20 percent of the world's population. Variations not found in those of European ancestry, even if common around the world, might be viewed as "abnormal." For example, the ability to digest lactose, the sugar found in milk, is very rare among adult mammals. Among humans, most North Americans, and most Northern Europeans, can digest milk well as adults, but that ability is not common in the rest of the world's humans (9). Adult lactose intolerance is considered an abnormal condition in the United States, even though it is found in most humans. A specieswide resource of human genetic variation could help counteract similar parochial misunderstandings of the human genome.

Human genetic variation can serve these goals only if it is known. Current efforts to study it run into the so-called empty matrix problem (5,10). Genetic variation has been studied in many populations around the world for most of the twentieth century, either as classical genetic markers (like blood types) or, more recently, as DNA variations. But genetic variations that have been examined in one population will not have been analyzed in others, so those populations cannot be compared. Additionally, for many populations the samples examined were quite small. Human population geneticists ideally would like to have a large number of samples from a large number of populations, all analyzed for the same large set of genetic variations. It was this desire that led to the birth of HGDP.

HGDP Forms, 1991 to 1993

HGDP's parents were two population geneticists: Luca Cavalli-Sforza at Stanford and Allan Wilson at University of California, Berkeley. Their discussions of the "empty matrix" problem took on new life at in the early 1990s when the Human Genome Project began operating. They conceived the idea of HGDP as an important, and relatively inexpensive, supplement to HGP. With Charles Cantor, Robert Cook-Deegan, and Mary-Claire King, they wrote the first call for the project, published in *Genomics* in 1991 (10).

Wilson's illness and subsequent untimely death from leukemia in July 1991 prevented him from playing a large role in the project, which had an important consequence for its shape. Wilson and Cavalli-Sforza had taken very different views over HGDP's sampling strategy. Wilson had wanted to sample on a grid basis, taking a certain number of randomly chosen people from each of a series of squares laid down over a world map. This sampling method was used by population geneticists studying wild populations of drosophila or other nonhuman species. Cavalli-Sforza was more interested in sampling based on existing populations, both for logistical reasons and for the information such sampling could provide about population history. Cavalli-Sforza might well have won this argument in any case, but Wilson's death ensured that result (1).

The nascent project began to attract other supporters, from anthropology as well as genetics, and from overseas as well as from the United States (11–13). Sir Walter Bodmer, then President of the Human Genome Organisation (HUGO), appointed a committee to study the idea of a Human Genome Diversity Project in 1992. Cavalli-Sforza chaired the committee (1).

In early 1992 the American members of this committee — Cavalli-Sforza; Marcus W. Feldman, a population biologist from Stanford; Kenneth K. Kidd, a geneticist from Yale; Mary-Claire King, a geneticist then at the University of California at Berkeley (and now at the University of Washington); and Kenneth M. Weiss, a physical anthropologist at Pennsylvania State University — received funding from the U.S. federal government for planning HGDP (1). The funds, amounting to about $55,000, came from the National Science Foundation (NSF), the Department of Energy (DOE), and from two parts of the National Institutes of Health (NIH): the National Institute for General Medical Science and what was then the National Center for Human Genome Research (now the National Human Genome Research Institute). The committee used the planning funds to sponsor three workshops: one at Stanford in July 1992, one at Penn State in October 1992, and one at NIH in February 1993.

The first planning workshop focused on collection methods. In considered what kinds of samples would be needed and how many samples should be obtained from each population. A crucial issue at this stage was whether the project should collect DNA samples or lymphoblast cell-lines. DNA, purified from blood or from samples

scraped from the interior of the cheek, was easy to process and inexpensive, but would provide only a small amount of DNA from each participant. The latter, white blood cells transformed by viral infection to have an indefinite life span, were more expensive and technically more difficult but held the promise of an inexhaustible supply of DNA from each sample. The workshop recommended a combined strategy of acquiring a small number of cell-lines from each population along with a much larger set of DNA samples (1,14).

The second planning meeting focused on anthropology and sampling strategy. Anthropologists met at Penn State to discuss what anthropological questions could be addressed by a resource of human genetic diversity and what *kinds* of populations would be most useful to sample. The participants talked about wanting to get samples from populations that represented major linguistic groups in a region, that had interesting cultural or linguistic aspects, or that could be used to answer specific anthropological questions. They also noted the value of collecting samples from isolated populations that were rapidly disappearing as distinctive cultures, largely through assimilation. These populations they referred to as "isolates of historical interest." Using these guidelines, the workshop divided into groups based on geographical specialization, with orders to produce a list of 500 populations as examples of the types of populations such a project might want to sample. The workshop felt that the list would be useful in demonstrating to funding agencies that the project thought through these sampling issues. [In the event, the anthropologists did not get their list below 700 populations (1,15).]

The third workshop, held for two and a half days in February 1993 at NIH, had three parts. The first day was returned to issues of sample types and the debate over cell-lines. The second day was devoted to a discussion of ethical and human rights issues raised by the project (16,17). On the third morning, the project organizers met with representatives from possible federal funding sources, including NSF, DOE, and several institutes of NIH. The morning started under a cloud, as the previous evening had featured President Clinton's first State of the Union address, in which he stressed the importance of reducing government expenditures. It did not get better for the project organizers, as the federal funding agencies listened to their plans for HGDP with interest but made no commitments.

In September 1993, HGDP held what became its founding meeting, in Alghero, Sardinia. This four-day meeting was supported by the remaining funding for planning workshops, with additional support from the Porto Conte Research and Training Laboratories Foundation, the European Commission, the Soros Foundation, and HUGO Europe. The researchers gathered at this meeting agreed on an organization and a substantive outline for HGDP (1).

The organization was to work at two levels. An international executive committee, affiliated with HUGO, would exercise oversight over the entire project. Thirteen members were appointed to this committee, from four continents. The international executive committee could have subcommittees and would, the meeting decided, establish at least two: one on informatics and one on ethics. Both the fund-raising and the actual operations of the project were to take place at a regional level, run by committees made up of scientists living and working in those regions. The regions were envisioned as having continental or near-continental scale.

The substantive work for these committees was to fall into three main categories: collection, preservation, and analysis (with subsequent data base entry) of DNA samples. The project adopted as an interim goal the collection of samples from 500 different human populations. Using the number of distinct languages as a rough proxy for the number of populations, this would be a sample of around 5 to 10 percent of human populations. The samples from each participating population were to include about 15 cell-lines and many more samples of purified DNA. Basic ethnographic information, derived from a standard questionnaire, would also be obtained from each individual providing samples.

These samples would then be preserved at repositories. The group at Alghero concluded that there should be more than one repository, in order to provide backup storage for samples, and that regional repositories should be considered. Samples of DNA from the repositories would be provided at cost to qualified researchers on request. What qualifications were necessary was not entirely settled, but the intent was to prevent the wide spread of these samples to cranks.

The samples were also to be the subjects of analysis, at the repositories and elsewhere. The repositories were expected to analyze samples for a standard set of markers. Researchers who accepted samples from the repositories were to do so subject to a condition that they return the results of their analyses of the samples to the repositories. The analysis of the samples was to be placed into a database by the project, which was also to be open to all qualified researchers.

The Alghero meeting endorsed an estimate that, worldwide, the project would take about five to seven years and cost about $5 million to $7 million per year. By the end of 1993, regional committees had formed in North America, South America, Europe, and Africa, while organizing efforts had begun in Oceania. The structure for HGDP proposed at the Alghero meeting was adopted by HUGO in January 1994. But attention had already been drawn to HGDP and opposition had started to build. For the rest of the decade, HGDP would make very little progress.

HGDP Stalls, 1994 to 1999

HGDP began to attract press attention in the 1992 planning workshops. The October 1992 workshop at Penn State in particularly was featured in *Science* magazine, where one of the sidebar stories had the inflammatory headline about "Endangered Populations" (18). Press attention brought less welcome attention, notably from a nongovernmental organization headquartered in Canada called the Rural Advancement Foundation International (RAFI) (19). RAFI had begun in the 1970s as a nonprofit group focusing on developing world agriculture, largely

in opposition to the so-called Green Revolution. Over the years RAFI had become concerned about what it termed "bio-piracy," collections of plant material from developing regions by Western seed and pharmaceutical companies. These collections, RAFI and others charged, were often used to produce new products with no financial or other sharing with the peoples from whom the plants had been taken. Starting at a Pan American Health Organization meeting in spring 1993, RAFI began to spread the word among nongovernmental associations concerned with indigenous peoples that HGDP was another, deeper attempt at bio-piracy, taking, this time, not indigenous people's plants but their own human genes (20). RAFI took the sample list of populations, compiled at the Penn State workshop in October 1992, and publicized it on the Internet as HGDP's "hit list" (21). This caused understandable concern to groups that suddenly read of themselves as the "targets" of an international genetics project from whom they had never heard.

RAFI also tried to tie HGDP to examples of what it claimed was bio-piracy by the U.S. government. In September 1993, RAFI protested a patent application filed by the United States on a white blood cell (lymphoblast) cell-line derived from a woman from the Panamanian Guaymí people. The cell-line was infected with a human retrovirus called HTLV-2, which made it interesting as a source for virus and viral antibodies. In October 1995, RAFI protested against a patent that had been granted to the United States for a cell-line derived from the white blood cells of a man from the Hagahai population in Papua New Guinea. This cell-line contained a related human retrovirus, HTLV-1. In both cases RAFI won. The United States abandoned the Guaymí cell-line patent application and abandoned the Hagahai cell-line patent application. In both cases RAFI tried to connect HGDP to these applications, claiming, in the Hagahai case, that "The thin veneer of the HGDP as an academic, non-commercial exercise has been shattered by the US government patenting an indigenous person from Papua New Guinea" (22,23).

RAFI's actions sparked a surge of condemnations of HGDP by various nongovernmental organizations, including some representing, or purporting to represent, indigenous peoples (One writer has counted 13 such resolutions (24); another, earlier article had a list of 11 (25).) Opposition to HGDP came from organizations concerned about indigenous peoples, about patent rights and the developing world, and about genetic technologies more generally. Most of the condemnations were taken after hearing only from opponents of the HGDP and they often demanded that the Project, which had not yet begun any collections, stop or reverse its work. Thus, for example, in 1994 the Foundation for Economic Trends, founded and controlled by Jeremy Rifkin, a well-known opponent of biotechnology, formally petitioned NIH to cease all of its (nonexistent) funding for the HGDP's collecting activities (26).

At least in part as a reaction to this activist opposition, the HGDP's North American Committee decided to spell out its ethical positions more fully. This was not the first time HGDP had discussed ethical issues. It had sponsored a full-day workshop on these issues at NIH in February 1993 (16). The organization's founding meeting in Alghero, Sardinia in September 1993 included a discussion of ethics and required that the International Executive Committee create an ethics subcommittee (1). The meeting's report devoted several pages to ethical concerns. It classified as major concerns collection issues, intellectual property; racism, xenophobia, and hypernationalism; and public understanding. The report included 10 proposed ethical guidelines that the meeting had accepted:

1. The HGD project and its participating researchers must always respect the human of the sampled individual and the cultural integrity of the sampled population. . . .
2. Informed consent is both an ethical imperative and a legal requirement. The HGD project must satisfy both conditions. . . .
3. Researchers should actively seek ways in which participation in the HGD project can bring benefits to the sampled individual and their communities. Examples of such benefit include health screening, medical treatment or educational resources.
4. One way to avoid bringing harm to the sampled individuals or their communities is by protecting the confidentiality of those sampled and, in some cases, of their entire community.
5. Although very unlikely, it is neverless [sic] possible that the results of the HGD project may lead to the production of commercially beneficial pharmaceuticals or other products. Should a patent be granted on any specific product, the project must work to ensure that the sampled population benefit from the financial return from sales.
6. Human history — and the human present — is full of racism, xenophobia, hypernationalism, and other tragedies stemming from beliefs about human populations. In the past, some of those tragedies have been perpetrated by, or aided by, the misuse of scientific information. All those involved in the HGD project must accept a responsible to strive, in every way possible, to avoid misuse of the project data.
7. Many people in the world have, at best, a limited understanding of human genetics. Some fear the consequences of human genetic research, in part because of the limits of their understanding. To scientists involved in the HGDP project, such fears may not seem justified or even, in some cases, fully rational by the concerns are very real of the people involved and they must be addressed. It is essential that a worldwide "public awareness" program is included within the project to educate people about its aims, methods and results.
8. Inevitably, the ethical issues faced by the project will evolve over time. The issues must therefore be kept under continual review. The widest possible consideration of the issues should be encouraged.
9. The transfer of technology to developing regions of the world, which is an integral part of the

proposed project, should contribute positively to the development of self-sufficiency in these regions. The help given should not be superficial and of only short-term usefulness.

10. There should be a feed-back of information to populations that participate in the HGD project, most especially about any aspect of the project in which a particular interest had been expressed.

The report from the Alghero meeting was not published until early 1995, and even then, it received little attention. It also became clear to the North American Committee that the project, at least in North America, needed a more concrete position on a number of ethical, legal, and social issues. As a result, in 1995, using funds provided by a grant from the John D. and Catherine T. MacArthur Foundation, the North American Committee developed its own draft "Model Ethical Protocol for the Collection of DNA Samples" (27). The Model Ethical Protocol was largely completed by the fall of 1995 (17). It was posted on the project's Web site in 1996 and published in a law review in 1997. As discussed in more detail below, the 15,000-word Model Ethical Protocol offers detailed guidance on many aspects of DNA collection for HGDP. It introduced two particularly significant innovations into this kind of research: the concept of "group consent" and the use of contracts to give participating populations some control over subsequent uses of their materials and derived data. The Model Ethical Protocol has been adopted by the project's North American Committee to guide collections in North America; it has not adopted by the project overall.

The ethical, legal, and social issues raised by HGDP began to be discussed outside HGDP and its opponents. UNESCO's International Bioethics Committee heard discussion of HGDP in the fall of 1994 and appointed a committee to study the issue. That committee, chaired by Dr. Darryl Macer, gave a mixed report in November 1995 (27–29). HUGO, which had appointed the original HGDP committee in 1992 and had adopted the Alghero meeting's recommendations in 1994, asked its Ethical, Legal, and Social Implications Committee to study the issues raised by the project. The Committee, chaired by Professor Bartha Knoppers, considered the project at an October 1995 meeting and produced a set of ethical principles, subsequently adopted by HUGO, on both HGP and HGDP (30). The ethical issues of HGDP were discussed at a workshop at Mt. Kisco, New York, in November 1993 (sponsored by the Wenner-Gren Foundation), at conferences at Stanford in November 1995, and in Montreal in September 1996 (31), at the meeting of the International Association for Bioethics in San Francisco in November 1996, at the annual meeting of the American Association for the Advancement of Science in February 1998 (32), and at a conference at the University of Wisconsin-Milwaukee in February 1999 (33), among other venues. Gradually a literature began to build about the ethical, legal, and social issues raised by HGDP specifically and by human population genetics more generally.

While the ethical (and political) conversation about the project was moving forward, the project itself was largely stalled for lack of funding. The project's organizers, having rejected the idea of any commercial funding, looked primarily to the U.S. government for funds. (HUGO itself never had sufficient funds to consider supporting HGDP and, in fact, has struggled to sustain itself.) The early interest shown by that government, including a largely favorable congressional subcommittee hearing in April 1993 (34), did not translate into substantive funding. DOE's part of HGP made clear quite early its lack of interest in HGDP. NIH was less immediately dismissive. Its portion of HGP, the National Human Genome Research Center (since 1997, the National Human Genome Research Institute) acknowledged the eventual importance of studying genetic variation but did not choose to invest any funds in HGDP. The National Institute for General Medical Science, which seemed more open to HGDP, also failed to fund the project. Only NSF, and in particular, its program in physical anthropology, was encouraging about funding HGDP. The entire annual budget for the physical anthropology program at NSF in the mid-1990s, however, was only about $2 million. It could only fund HGDP if it received a major infusion of funds.

As they had done a decade earlier with HGP, the federal agencies with the most interest in HGDP decided to ask the U.S. National Research Council to report on the idea of the project. The National Research Council, the report-writing arm of the U.S. National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, requires funding for its reports. The NSF, NIGMS, and the office of the Director of the NIH agreed in late 1994 to contribute about $400,000 to fund an NRC committee report on HGDP in the expectation, shared by HGDP advocates, that the Committee's report would either launch or bury the project. The Committee's report, which was finally released in November 1997, did neither.

In the event, the NRC committee was not appointed until early 1996 (35). Chaired by prominent geneticist Dr. Jack Schull, the 17-member committee included three ethicists, Professor George Annas, Dr. Eric Juengst, and Dr. Katherine Mosely. The committee held public meetings in April, July, and September 1996 at which both organizers and opponents of HGDP spoke. In November 1996 it began its deliberations, which took 12 months to yield a report. The report, when published, satisfied neither the proponents nor the opponents of the project — and may have confused many of its readers. The world's two leading scientific magazines, *Science* and *Nature*, reported on the committee's work with diametrically opposite headlines: "NRC OKs long-delayed survey of human genome diversity," claimed *Science* (36), while *Nature* wrote "Diversity project 'does not merit federal funding'" (37), (a headline *Nature* later retracted) (38).

The committee's chair, Dr. Schull, ended up writing correction letters to *both* journals (39,40), but, in fact, the committee's report seemed to provide, in different sections, support for both journals, particularly as diversely interpreted by some committee members. The committee stated that the plans for HGDP were too vague to be the subjects of a specific evaluation, and as a result it would

evaluate the overall concept. It concluded that although there were serious reasons for ethical concern, the federal government might fund such research but only through U.S. researchers who would be subject to U.S. rules on protection of human subjects. The overall message of the long-awaited NRC report was so equivocal, however, that it failed to achieve its sponsors' overall goal — to provide a clear verdict, up or down, on HGDP.

As of early 2000, HGDP — proposed in 1991 and organized in 1993 — remains largely unfunded. Its committee structure has in large part withered. The annual meetings of its International Executive Committee proposed at Alghero in 1993 became, as a result of lack of funds, one meeting in London in September 1994 and one meeting as part of a Cold Spring Harbor conference in October 1997. The North American Committee has remained active as a committee, thanks in part to the funds provided by the 1994 MacArthur grant. HGDP Regional Committees in Southwest Asia and in China, with local funding, have collected some DNA samples from their regions. The South American Committee disbanded and the others are inactive.

The NSF provided some funding for HGDP-related activities in 1997, when it awarded about $600,000 in grant funding for "pilot projects" on various issues related to the HGDP. These funds were explicitly not to be used for collecting DNA. Some continued interest in HGDP and its goals has been sustained by two conferences at Cold Spring Harbor Laboratories on Human Evolution, in October 1997 and April 1999. And, in 1999, HGDP has agreed with the Centre pour l'Étude du Polymorphisme Humain (CEPH) in Paris that CEPH will store 1000 cell-lines from existing collections and make DNA samples available to qualified researchers, thus advancing some of the goals of HGDP.

Outside HGDP the subject of human genetic diversity has become of greater interest. In 1998 the NHGRI, with the end of the sequencing phase of the HGP in sight, created an initiative to use samples from diverse populations to search for "single nucleotide polymorphisms," or SNPs. These SNPs are expected to be useful for locating genes of medical interest, and NIH sought to create a resource of SNPs in the public domain before private firms patented too many of them. (A consortium of pharmaceutical companies and foundations has embarked on a similar effort, again with the goal of putting SNPs in the public domain (41).) The NIH samples, selected from Americans of European, African, Asian, and Native American ancestry, are available from a public repository but with no identifying ethnic information, even at the continental level (42). In fact, investigators taking samples from the repository are required to promise that they will not attempt to identify the ethnic background of the person who gave the sample (43).

Meanwhile, human population genetics continues to be done, for anthropological and other purposes. Recently published research has used the analysis of DNA samples to provide evidence about the origins of Native American (44), Japanese (45), and Chinese populations, among others (46,47). But the empty matrix problem remains — there still exists no broadly derived set of human DNA samples that have been analyzed for a standard set of markers. And no special ethical oversight exists for the ongoing efforts to collect samples for the study of human population genetics.

ETHICAL, LEGAL, AND SOCIAL ISSUES RAISED BY HGDP

As a large-scale genetics research project, aimed at collecting DNA samples from thousands of people, HGDP has raised most of the issues that are raised in similar research. These concerns include, among other things, ensuring the confidentiality of individual information, avoiding undue inducements for research subject participation, the use of previously collected samples, and dealing with the possible return of medically significant information to research participants. The wide variety of cultural backgrounds of anticipated participants in HGDP makes resolving those issues unusually complex, but the basic problems are the same ones faced by any research related to human genetics. Answers to those problems are the same ones faced by any research related to human genetics. Answers to those problems in any context may be contentious, but the HGDP's answers, as proposed in the Alghero report and especially in the Model Ethical Protocol, are unusually detailed but not extraordinary (27). One difficult area not covered at length in the Model Ethical Protocol is the use of previously collected samples, an issue that remains as vexing and complicated for HGDP (48) as it does for medical research more generally (49,50).

The issues unique to the HGDP arise from the nature of the "research subjects" of HGDP: not primarily individual humans, but human groups. The group nature of this research that makes the ethical, legal, and social implications of HGDP so fascinating and so difficult. It also makes the proposed solutions to those issues, particularly the North American Committee's requirement of "group consent," complex and controversial, in some respects perhaps more controversial than the research itself.

Group Concerns

A host of concerns have been raised on behalf of the populations that might participate in the project. Although many different issues are involved, they all fall into two main categories: fears that genetic information will harm the groups that take part and concerns that the groups will be financially exploited with respect to their genetic resources. These concerns have, not surprisingly, been raised largely on behalf of indigenous groups that have suffered from oppression and continue to exist under European domination. Thus the concerns have been raised most actively by groups speaking for Native Americans throughout the Western hemisphere, Australian aborigines, and the Maori, the Polynesian minority in New Zealand. Although the concern about exploitation seems to exist in some other settings, the fears of harm are strongest among the populations with the least power.

It is worth noting that HGDP does not intend to sample solely populations that are indigenous, small, or powerless.

Its goal is to collect a roughly proportionate sample of the world's human populations. Some of them will be small; on the other hand, the Han (ethnic Chinese) are a population of great interest to population geneticists, and they make up about 20 percent of all humans. Thus much of the intended work of HGDP could be achieved without sampling the populations that might be most concerned about the project.

Direct Harms. Critics of HGDP have argued that the project could bring a wide range of harms to participating populations. These range from harms to the population's cultural, to harms in its members relationships to the broader society, to political costs. Each is discussed below.

Two kinds of possible cultural harms have been suggested. One is that participation in this kind of project may violate either a particular group's culture or, more broadly, general indigenous norms (51–54). It has been argued that the assertion of domination and control over nature implicit in modern science is, in itself, antithetical to indigenous cultures. More specifically it is also claimed that some aspects of the DNA collection process will violate cultural norms, such as those that may relate to the treatment of blood or hair. The great degree of variation in human cultures makes it prudent to be skeptical of claims of "universal" cultural norms for indigenous peoples, but certainly some groups will hold such views.

A second type of cultural harm looks instead to the results of the project. Some indigenous activists have argued that scientific evidence from HGDP, by contradicting local cultural histories and origin stories, may undermine the authority and power of the culture. Thus many Native American populations have oral histories that place their origin in American locations. Genetic evidence that the ancestors of Native Americans migrated to the Western Hemisphere from Siberia could shake community members' faith both in that origin story and in the entire culture (51). Supporters of the project counter that these claims need to be examined carefully. Anthropologists have long argued, based on many lines of evidence, that Native Americans migrated to the Western Hemisphere from Asia. HGDP might add some evidence to that conclusion, but its supplemental effect seems unlikely to be great. In addition, it is urged, many people have shown a great ability to disregard scientific evidence that conflicts with their origin myths. In the United States, which provides support for extensive research into the evolution of modern humans, over 40 percent of the population continues to believe that humanity was created in the Garden of Eden about 6000 years ago. Finally, one may question whether cultures necessarily want to preserve their myths unchanged. Nonetheless, in some circumstances this may be a realistic threat to a population's culture.

Those concerned about HGDP also point to broader social harms revolving around concerns about discrimination. As is the case with individuals, groups might be considered genetically susceptible to particular diseases or conditions. This could lead to discrimination against group members in employment, insurance, or other social activities. It might stigmatize the group or support a racist belief in the group's inferiority.

HGDP's supporters use several arguments to try to undercut this criticism. First, HGDP is not looking for disease-related genes; it will focus on random markers that will not reveal this kind of information. Second, the relevance of these concerns they depend crucially on a population's situation. Commentators have discussed in detail the possible and actual existence and significance of discrimination in insurance and in employment in the United States. (And, in the American context, the motivation for employment discrimination often stems from the employer's payment of health insurance costs (55).) In societies where every individual is guaranteed health coverage — which comprise all wealthy countries other than the United States as well as some middle- and lower-income countries — this discrimination becomes irrelevant. In some traditional societies where Western medicine itself may be unavailable, both health insurance and employment discrimination may be irrelevant.

More fundamentally, though, supporters of the project contend that discrimination fears exist in an individual context because genetic information might reveal something about risk that cannot otherwise be known. At the level of human populations, genetic analysis would not often provide that kind of information. Such an analysis might indicate that from genetic causes, the Irish have several times the average levels of the genetic variation that, when an individual inherits two copies of the variant, causes the genetic disease phenylketonuria. It would not indicate which Irish people were at risk and which were not. And, more important, the levels of these risks will already be known directly from public health statistics and epidemiological research. One does not need to know the distribution of genetic variations associated with disease to know disease rates in populations; those rates can be examined directly. It is conceivable that for some disorders, cases linked to genetic variations will have differences from nongenetic cases that would be important to insurers or employers. One could hypothesize that for example, breast cancer cases related to mutations in BRCA1 were less easily and cheaply treatable than other breast cancers. In that case, knowledge that a population had an unusually high rate of those mutations might be relevant. But, for the most part, the prevalence in a population of genetic variations linked to disease will add nothing to the risk information already available from observation of the disease incidence.

The issue of stigmatization is more complicated. Project supporters point out that it is hard to think of examples any populations that are stigmatized because of higher rates of disease. Europeans have higher than average rates of cystic fibrosis, many Africans and Southeast Asians have higher rates of hemoglobinopathies, Ashkenazic Jews have unusually high rates of Tay-Sachs disease. Alcoholism may be one of the few examples of a disease or condition that has stigmatized some populations, such as Irish, Russians, and Native Americans. Again, one might think that if a population is going to suffer stigmatization because it carries disease-associated genetic variations at an unusual level, then it would already been stigmatized for having a high rate of the disease itself. The fact that

the cause is genetic, though, might lead to a higher social impact. The concept of "genetic essentialism" — the idea that one's genes are one's essence (56,57) — could lead people to believe that a population with a high rate of genes associated with a disease is somehow inherently flawed, or more flawed, than a population with a low rate of that genetic condition. Of course, every population is likely to have somewhat higher and lower genetic risks for different diseases, but one might focus on a particularly stigmatizing disorder, such as schizophrenia. On the other hand, a genetic explanation might have opposite effects. If a higher rate of disease in a population is believed to be genetic, the individuals suffering from the disease may be viewed more favorably on the theory that the disease is not the result of their actions. (This relies on the general popular assumption, often false, that diseases with genetic "causes" do not also have, in the same individual, environmental "causes.") This issue has been interestingly discussed in the context of a possible genetic association with sexual preference (58); whether a conclusion that a population's higher rate of a condition has genetic roots would increase, decrease, or leave the same any stigma attached to the group because of the condition is unclear.

The asserted political harms vary. One set of concerns revolves around land claims. If rights to land are affected, legally or politically, by the history of a population's occupation of the land, genetic evidence that the population migrated from elsewhere, or direct evidence from ancient DNA that a different population occupied the land in the past, might have some political relevance (8). Project supporters argue, on the other hand, that few if any peoples are thought to be truly indigenous to any location — under current thinking about human evolution, only some African groups could even possibly make that claim. In many regions of the world, migrations and changes of homeland have happened within historical memories — in too many places, within the living memories. For legal and political significance, the relevant time frame is important; it is hard to imagine many situations where DNA evidence would have current relevance. And, if it were, it might be thought as likely, in a given situation, to favor a population's claims as to harm them. But if one takes the view — which might be quite reasonable for some subordinated populations — that the dominant culture will twist any new "scientific" evidence to its political benefit and your detriment, this concern becomes more understandable. Even if genetic evidence is not directly relevant, it might be used to weaken the political standing for some land claims. This fear of this kind of general political effect may be one of the sources of the controversy over the so-called Kennewick man, an ancient skeleton found in the northwestern United States that some allege has "European" features.

Also related to land claims, concrete or broad, is a possible concern about membership. Some worried that an outside government might impose a genetic test for membership on the population, adding and subtracting members without the group's consent.

Critics of HGDP cite the possibility that genetic variations might be used for biological warfare against sampled populations as yet another political concern. Project supporters claim that such ethnically targeted biological warfare seems scientifically implausible for two reasons. First, the enormous overlap in genetic variations between populations would mean that such a weapon would not affect many in the targeted group and would affect many outside the targeted group. The twentieth century has seen all too many more discriminating and "efficient" methods for genocide. Second, scientists do not know how to kill a cell based on variations in its genetic material. If they did, all infectious disease organisms and all tumors could be handily defeated. Nonetheless, some recent publications have fed this fear (59–61).

Exploitation. Issues of exploitation of participating populations are somewhat different from those of direct harms to those populations. These concerns focus on the possibility that participating populations might be robbed of something of value as a result of the project — "their genes" (20). This view is fed by the great interest, popular and financial, in biotechnology. It is also exacerbated by stories, which can reach the level of myths, about past "victims" of predatory commercial biotechnology, such as the tale of John Moore's spleen (62) or the patent applications for cell-lines derived from indigenous Panamanians and Papuans. There are also parallels, close and distant. The term "bio-piracy" was coined, apparently by RAFI, to describe the use of plants and plant genes from the developing world by commercial firms from the developed world for pharmaceutical and agricultural profit. HGDP may have appeared just an extension of that process to human genes. More generally, though, many indigenous cultures have suffered from exploitation by outsiders in recent history. The slogan that "You stole our land, you stole our resources, and now you want to steal our very genes" has power. Finally, this fear of exploitation flourishes as a result of limited understanding of human genetics. People are ready to believe that their population has its own "genes," let alone its own special variants of human genes. From the identification of a gene in a small number of people, to an understanding of its function, to its application as a profitable pharmaceutical may appear an easy series of events to the outsider. The idea that one's group has "special genes" of great value and power is also a flattering one. For all these reasons people concerned with indigenous groups or, more broadly, the developing world might conclude that HGDP was an effort to exploit the commercial value of human genes from outside the developed world — after all, if those genes did not have value, why would anyone look for them?

The reality is less financially promising and more complicated. Populations do not have unique "genes," although they may have an unusually high (or low) percentage of certain variants of human genes. Identifying a genetic variation as associated with a disease (or with protection from a disease) is a long and complicated process, requiring extensive medical as well as genetic work with individuals and their families, both affected and unaffected with the disease. The jump from the discovery of a disease-related variation to a commercial product is enormous and has rarely been made. None of these activities has any relationship to HGDP. The

project has disclaimed any commercial interests, backing, or connections. But, as critics of the project point out, scientists with commercial connections would, under HGDP's initial plan, have access to samples and data collected by HGDP.

In fact, neither a renegade HGDP nor scientific third parties would be likely to be able to derive information with strong medical, and hence commercial, value from the HGDP samples. HGDP will not obtain the kind of medical data about participants that would be necessary to make the samples it collects useful for this kind of medical research. Without knowing which participating individuals had diabetes and which did not, for example, research into connections between genetic variations and diabetes is not possible. HGDP's organizers have stated that some information of value to medical researchers might be created by the project. For example, if a researcher had identified a "candidate gene" for a particular disease, HGDP's resources might be able to tell her whether the high risk variant was more or less common in populations with known epidemiological risk. But this kind of information seems likely to have little direct commercial value — it will neither prove nor disprove a genetic connection but might narrow the possibilities. Thus HGDP has consistently evaluated the likely commercial value of the samples and data it collects as quite low. It has not, however, been able to rule out the faint possibility that some samples or data could end up having commercial value for some third party user of the information. It is this indirect and unlikely contingency that could give rise to concern about the financial exploitation of participating populations.

HGDP's Responses to Group Concerns

HGDP has been aware of some of the group concerns raised by its research since at least early 1993. It tried to address many of them in the Alghero meeting and report. Some of its solutions in that report include respect for the cultural integrity of participating populations; the resolution to protect individual and, in some cases, group confidentiality; and the renunciation of commercial connections; and, when appropriate, a sharing of financial benefits. The Alghero report expressly recognized that "the ethical issues faced by the Project will evolve over time" and called for their continual review. The most detailed such review from within the project has come from the North American Committee and its Model Ethical Protocol. This document accepted, for North American HGDP work, two broad responses to group concerns: group consent and contractual restrictions on subsequent uses of samples and data. Its idea of group consent has spawned nearly as much controversy in ethical circles as HGDP itself — and more academic literature.

Group Consent. The Model Ethical Protocol requires that where feasible, those collecting DNA samples in North America for HGDP obtain not only the informed consent of individual participants but also of their population (27,63). This collective consent is to be sought from the group's "culturally appropriate authorities," as defined by the group itself. Where permission is denied, the project would not accept any samples from members of the population. The protocol recognizes that such consent will not be feasible for groups that do not have an authority structure, like Irish-Americans or Ashkenazic Jews. In those cases it requires full discussion within the direct community in which the collection is to be done — the local town or religious or cultural association — as well as dissemination of the facts about the planned study broadly to those who identify with the group to enable them to try to influence their colleagues to participate or not in the project.

The North American Committee put forward several arguments in favor of group consent. At the most basic level, its claim is that human population genetics research, by its very nature, has the population as its subject as well as the participating individuals from that population. Every member of that population may be affected, positively or negatively, by the results of the research and thus the population, collectively, should have a chance to decide on its participation. Just as individual informed consent lets people decide whether a particular medical procedure or research project meets *their* balance between costs and benefits, under group consent the group as a whole could hear about the possible harms discussed above and decide whether to take those risks. If a group decides that HGDP involves too great a risk of cultural disruption, negative land claim evidence, or any other harms, it will simply not participate. Individual informed consent is also justified as respecting the autonomy and personhood of the patient or research subject. Group consent acknowledges the cultural reality of ethnic groups. This may have particular force with populations that have been dominated by others. By seeking and respecting their decision, one respects their cultural autonomy. This kind of respect may be welcome by almost any group, but is particularly appropriate with Native American tribes in the United States. Legally, federally recognized tribes are sovereign governments existing in a complex political relationship with the federal government. Culturally non-Indian peoples often do not recognize that legal reality. By expressly giving them a chance to say no to the research with their populations, whether or not conducted on the reservation and hence within their political jurisdiction, the Model Ethical Protocol would uphold and extend that autonomy. And, as HGDP seeks to sample only about 5 to 10 percent of the world's human populations, many groups can decide not to participate without jeopardizing the project's goals.

This idea of group consent is not new. Ethnographers and epidemiologists have long known that the kinds of intensive studies they conduct in the field often require, as a practical matter, a great deal of community consensus building and the approval of those with authority, formal or informal, within the group. In the United States many federally recognized Native American tribes, which are sovereign governments within their reservation boundaries, have in recent years established their own institutional review boards for assessing research proposed on the reservation (64). The idea of collective approval is new to some areas of research, but not to all. In its August 1999 report on the use of human biological materials, the United States National Bioethics

Advisory Commission recognized the significance of group interests and the possible value of group consent, but noted that current law regulating human subjects research does not require that these collective issues be addressed (50).

Group consent has precedents, but it also has problems, both practical and philosophical. Perhaps the largest problem, and one that fits both categories, revolves around the definition of the "group" whose consent is required. Eric Juengst has argued that this is a fatal flaw (8,65). Genetic information about one Navajo village, for example, may have implications all Navajos, on the reservation or off. It may also say something about other speakers of languages in the Na Dene language family, which is spread widely in western North America. It could demonstrate something distinctive about all Native Americans. Or it may have no implications for any groups larger than the individuals involved. Until the research is completed — along with complementary research on other populations — the scope of its effects cannot be known. If the justification of group consent is to have the approval of those who may be affected by the research, it cannot succeed. The relevant group, which would be population defined by a relative closeness of genetic relationship, cannot be defined in advance other than to say, based on the existing knowledge of population genetics, that it does not coincide perfectly with any culturally defined ethnic group.

Juengst's criticism is clearly correct, though its impact is not as clear. Defenders of the group consent can argue that although the definition of the relevant group will never be perfect, it will be a good start. The Model Ethical Protocol says that researchers should determine the relevant group in consultation with the community with whom they are working. If they view a larger grouping as necessarily involved in the research, then it should be so involved. Using the relevant community's own definition upholds, at least, the autonomy justification for group consent even if it does not fully encompass the consent of those affected. Juengst suggests replacing group consent with consideration of group interests during individual consent. Thus, rather than ask the group for consent, individuals asked to participate in the research would be told that the research may have implications for specific groups they belong to and that they should consider those implications when deciding whether to participate. This kind of directed consideration could play a useful role, although it too would be imperfect. The particular individuals approached for the research might not reflect well the concerns of the larger community. And the process of consent — the discussion and debate — may well be much broader (as well as longer and more arduous) in group meetings than in individual informed consent sessions.

Group consent has been subject to a number of other concerns, as well as arguments about the significance of the problem it attempts to solve (66–68). The determination of the "culturally relevant authorities," whose approval will be required may prove difficult (69). Federally recognized Indian tribes, in the United States and in Canada, have governmental structures that will provide a starting place for such authorities. Other communities will not usually have formal governments and finding any "culturally appropriate authority" may be difficult. Even recognized Indian tribes may have a variety of nongovernmental structures claiming authority. If there is a dispute within the population about which people or organizations have authority, how should researchers resolve it? This may be particularly difficult if participation in the research becomes, itself, a divisive issue within the population. One can even imagine one faction using the researcher's acceptance of them as authoritative as a weapon in a power struggle within the group. In some cases, where the structure of authority is sufficiently unclear, researchers may have to walk away from the research entirely. Other criticisms are less pragmatic. Juengst has suggested that group consent is suspect because it asks cultures to become complicit in research that may harm them, which seems a strained way to view giving choices to groups (67). Presumably the interests of the groups would not be better served by a process that imposed research on them whether they liked it or not. Some researchers, including some on the North American Committee, have been troubled that group consent would deprive individuals of their "right" to take part in research on the basis of the group's fears, even if the individuals had removed themselves from the group. Finally, it has been pointed out that the logic of group consent may extend uncomfortably far. Traditional research on human genetics uses "groups" made up both of people suffering from disease and of families. Should disease organizations or extended families also be accorded the power of group consent (63)? The issues raised by a group consent requirement are indeed broad, so broad that it has not been formally adopted yet by any HGDP regions other than the North American Committee. And a Canadian effort to revise their codes of research ethics, though initially proposing a strong "group consent" requirement, removed the discussion from their final version, leaving only a short section on special issues in dealing with the native peoples of Canada (70).

Some have argued that the process of group consultation and discussion should be required even if strict group "consent" is not sought. In a series of articles, Morris Foster and coauthors have argued for a form of "community review," which involves the community where the research is proposed without formally requiring community approval (71–73). Foster has combined this approach with a model agreement for regulating the rights of communities that agree to be the subjects of research (74).

One other response to the risks of group harms deserves mention. Foster has at least occasionally suggested that the identities of populations participating in genetic research should be hidden (33). If no one knows what group provided the samples, the information cannot be used against them by outsiders: in employment, insurance, or politics. (The internally derived harms, such as the undermining of the population's culture, could still occur.) HGDP proposed a limited form of this response, through "fuzzing" the precise identity of a participating group. For example, rather than identify a specific village, perhaps even by use of the global positioning system, HGDP could say that the samples came from a village within a broader region. Of course, withholding information about the group's identity may have its own bad consequences,

both for science, which may not proceed accurately without sufficient specificity, and for the group, which could lose some benefits of the research it not identified. Besides, if the level of group anonymity were too great, it would render impossible some of the historical research that is a key purpose of HGDP. Nonetheless, some amount of group anonymity is a protective tactic that should be considered in this kind of research — and perhaps should itself be a subject of discussion in the consent process, group or individual.

Contractual Defenses against Exploitation. The founding meeting of the HGDP endorsed the idea that participating populations should share in the financial gains, if any, resulting from the HGDP. Its report states: "Although very unlikely, it is neverless [sic] possible that the results of the HGD Project may lead to the production of commercially beneficial pharmaceuticals or other products. Should a patent be granted on any specific product, the Project must work to ensure that the sampled population benefit from the financial return from sales" (1). The Alghero report, however, provided no suggestions for how the Project might ensure such a result.

Again, the Model Ethical Protocol sought to implement that principle. It did so through the medium of contracts (27). HGDP plans to operate both DNA sample and cell-line repositories and a database containing the analyses of HGDP samples. Samples and information in both facilities would be available to qualified researchers. The Model Ethical Protocol states that those facilities should only allow access to researchers who agree, by contract, to restricted commercial uses of the information they obtain from the repository or database. Contracts governing this kind of access, called "materials transfer agreements" in the case of samples and "database access agreements" for the computerized information, have long existed, initially to protect the repositories and databases from liability. The Model Ethical Protocol would expand those agreements to provide that all of the HGDP's samples and information could only be used in ways consistent with terms and conditions limiting their use. These terms and conditions, the protocol suggests, would usually be set by the participating population as part of the process of group consent. Thus a population might choose to forbid any commercial use or any patenting of their samples, the information, or products derived from the samples or information. Alternatively, it might authorize such uses on the payment of a specified royalty to the group or on the completion of a subsequent express written agreement. Or it might allow any uses. Anyone who wanted to use samples or information provided by that group would have to agree to abide by its terms.

Like group consent, contractual limitations on commercial use also have problems, both in their inception and in their implementation, but, unlike group consent, both the idea and its problems have been little discussed. At least four topics need to be raised. First, the topic of possible commercial gains would need to be discussed very carefully in the group (or individual) informed consent. Otherwise, the (very faint) hope of financial rewards might unfairly influence a group to participate in research. Undue inducements to participate in human subjects research are not allowed; despite the researchers' disclaimers, a group might wrongly believe that it has a strong chance of becoming wealthy from its "gene royalties," which, as the researchers would recognize, would be very unlikely ever to exist. Second, some default standards would have to apply to populations for which group consent was not obtained because it was infeasible. These standards would not only have to determine what would be a "fair" return but would also have to consider the problem of who should receive or spend any such sums, as, by definition, no "culturally appropriate authority" would exist for a group from which no consent was sought. Third, attention would have to be paid to who could sue to enforce these contractual clauses. The population might have the best incentive to sue, but it would have poor access to information about the use of its materials and often poor access to the legal system. HGDP would have better access to both but less incentive to protect the population's rights. Finally, would these contractual clauses actually be enforced? Although the status of DNA as "property" remains unclear, there seems no reason to think that courts would not enforce this kind of contract. The main problem is whether anyone would recognize that a pharmaceutical company had started marketing a product based, at least in part, on research done many years earlier using some HGDP samples? This kind of enforcement problem could be enormous, but it might find a solution in the very structure of the pharmaceutical market. Bringing a drug to market is a very long and extremely expensive proposition, costing in the United States an average of several hundred million dollars. The existence of even a possible claim that such a product breached a contract with HGDP and participating populations might throw such a "cloud on the title" of the pharmaceutical company as to ensure that it would not proceed without negotiations.

Racism

In the nineteenth and early twentieth century, the biological sciences, including the newly born field of genetics, helped legitimize and reinforce racism by providing "scientific proof" of the inferiority of disfavored human groups (75–77). This history, shameful to contemporary geneticists, combines with the continued persistence of racist stereotypes to make genetic research into "racial" or ethnic characteristics still politically and ethically charged (78–80). HGDP risked being caught in these controversies around the genetics of race. Although its organizers denounce the idea that human races have *any* biological or genetic meaning, let alone that one "race" is genetically "superior" or "inferior" (6), the project's very name emphasizes that it is looking for genetic differences — and its research agenda makes it clear that those are differences between human populations. The jump from talking about "populations" to being seen as talking about "races" is short. This issue affected HGDP in at least three different ways.

First, some critics accused HGDP of being itself in the thrall, consciously or not, of nineteenth-century visions of genetically defined races. Its plan to compare the patterns of genetic variation frequencies of different culturally defined human groupings was seen as reflecting

a belief by the project organizers that such "pure" groups, or, in a term used by the project itself initially, population "isolates" really existed (81–84). Critics viewed this as both scientifically naive and socially harmful. Even structuring the project around an examination of these culturally defined populations would, in their view, necessarily bias the results. From this perspective, Allan Wilson's "grid" plan of sampling had great benefits. Second, other commentators pointed out that even if the HGDP organizers held a scientifically correct vision of the limited connection between cultural groups and genetic populations, the public did not (8). HGDP, by even looking at genetic differences between group, would reinforce for the public the ideas that such differences existed and that they were important. Finally, HGDP organizers themselves recognized that, however meritless scientists considered "scientific racism," the project's results could be misused for racist ends. In an event used at the time of the greatest discussion of this point, Bosnian Serbs might seize on some slight genetic variation between themselves, as a group, and Bosnian Muslims, as a group, and claim "scientific" support for ethnic cleansing.

The project's supporters disclaimed any belief in genetically "pure" populations, let alone in genetically defined human races. They countered that, although culturally defined groups were not the same as genetically defined populations, they were often genealogically more closely related to each other than to outsiders. Thus the statistical analysis of patterns of genetic variation between different groups had led to useful results in prior research. Further they called the Wilson grid plan impracticable — after all, which one person would represent the genetic variations found in New York or in San Francisco? And, they urged, because the project's data would all be public, the data could be analyzed by those skeptical about the sampling strategy to detect any bias that might exist. The project's organizers admitted the second and third risks: the unconscious reinforcement of a false concept of "genetic race" in the public mind and the possibility of racist or nationalist misuse of the project's data or findings. To counter those problems, the organizers called both for general education and for "ready response teams" of scientists prepared to refute false claims (1,27). The project's supporters went farther and argued that HGDP's results would make "scientific racism" even more untenable by demonstrating the great genetic similarities among humans. This response drew its own reply, accusing HGDP of exaggerating the effects its findings would have in countering racism (7).

One other claim, at least related to racism, has been advanced against HGDP. Its opponents argue that it did not take indigenous groups seriously and did not include them in its planning. This, it is urged, reflects anything from racism and colonialism (51,52), to a scientific arrogance toward research subjects (7). HGDP's organizers respond that indigenous peoples were not excluded from the project planning and point, among other things, to numerous meetings with indigenous groups and to two Native Americans on the project's North American Committee. It is true, though, that at least in North America, the project decided not to emphasize contacts with indigenous activist groups. Instead, it decided to focus its contacts on people living in native communities, after the project was funded to begin collection and so concrete measures could be discussed. It would, in that way, build connections with the very people whose participation in the project it would seek. And it would avoid encounters with political activist groups, often distant from the Native American communities, whose opposition to the project was viewed as highly likely. Whether this strategy was wise remains unclear, particularly when the funding that would have allowed concrete contacts with local communities — and, the project had hoped, would have provided examples of successful collaborations — did not arrive.

FUTURE OF HGDP

HGDP's future remains unclear. As of early 2000, no funding for substantive work had yet been received from the U.S. federal government. No DNA samples have been collected under the auspices of HGDP except in China and in Southwest Asia, with local funding to those regional committees. Discussions are under way for the creation of general repository for some genetic samples, but no database for the results of analysis of HGDP samples had been created. The project continues to inspire controversy, among some ethicists, some nongovernmental organizations, and some indigenous groups. It also continues to inspire its organizers' hopes. Only time will tell whether it thrives or shrivels.

Whatever happens to HGDP, the future of studies of human genetic variation seems quite clear. The kinds of isolated, individual studies that have been going on for 80 years are continuing and expanding. Perhaps more important, the importance of genetic variation for understanding genetic links to disease is increasingly recognized. Once HGP completes sequencing "the" human genome, two next steps seem obvious: determining the function of the identified genes and understanding the consequences of the variations found in the six billion human genomes. Already NIH has created a resource of 450 DNA samples from ethnically diverse residents of the United States to be used to find SNPs (single nucleotide polymorphisms). This kind of variation is mainly important to provide landmarks in the genome; other examples of human genetic variation will almost certainly be of medical and scientific interest. In the long run, as the studies of genetic variation for nonanthropological reasons get broader, the samples and data that HGDP seeks to gather may appear from other sources.

This kind of future, however, would be a continuation of the past. Unless DNA samples were available for analysis or were analyzed against a standard set of markers, the empty matrix problem that led to the organizers to propose HGDP would continue. And supporters of the project argue that the ethical, legal, and social concerns would be heightened, not diminished, by the end of HGDP. Rather than having a project that insisted on compliance with certain ethical standards, DNA collection and research would continue to be done in hundreds of different laboratories, using many different sets of ethical rules. Opponents, on the other hand, can hope

that communities can more effectively resist many small laboratories than one worldwide "project." And they may believe that an end to HGDP will lead to protections better than those HGDP claimed, unconvincingly to them, to plan to impose. One way or another, the issues will remain as relevant. Whether they would be as visible and as fully debated in the absence of a formal HGDP is, at best, unclear. But, whether it ever is completed or not, HGDP and the controversy around it has at least helped illuminate the complex issues — ethical, legal, social, and political (85,86) — raised when human genetics moves from the individual or family to the population.

ACKNOWLEDGMENT

Professor Greely has been associated with the HGDP since 1993 and chairs the Ethics Subcommittee of the Project's North American Committe. In that role he was the principal author of that Committee's Model Ethical Protocol.

BIBLIOGRAPHY

1. Human Genome Diversity Committee of HUGO, *Human Genome Diversity (HGD) Project, Summary Document*, HUGO Europe, London, 1995, Available at *http://www.stanford.edu/group/morrinst/HGDP.hmt*
2. L. Hirschfeld and H. Hirschfeld, *Lancet* **2**, 675–695 (1919).
3. W.H. Schneider, *History and Philosophy of the Life Sciences* **18**, 7–33 (1996).
4. A.E. Mourant, *The ABO Blood Groups: Comprehensive Tables and Maps*, Blackwell Scientific, Oxford, 1958.
5. L.L. Cavalli-Sforza, P. Menozzi, and A. Piazza, *History and Geography of Human Genes*, Princeton University Press, Princeton, NJ, 1994.
6. L.L. Cavalli-Sforza and F. Cavalli-Sforza, *The Great Human Diasporas: The History of Variation and Evolution*, Addison-Wesley, Reading, MA, 1995.
7. M. Lock, in B.M. Knoppers, ed., *Human DNA Sampling: Law and Policy — International and Comparative Perspectives*, Kluwer Law, The Hague, 1997.
8. E.T. Juengst, *Kennedy Inst. Ethics J.* **8**, 183–200 (1998).
9. W.H. Durham, *Coevolution: Genes, Culture, and Human Diversity*, Stanford University Press, Stanford, CA, 1991, pp. 226–285.
10. L.L. Cavalli-Sforza et al., *Genomics* **11**, 490–491 (1991).
11. K.M. Weiss, K.K. Kidd, and J.R. Kidd, *Evolu. Anthropol.* **1**, 78–80 (1992).
12. J.R. Kidd, K.K. Kidd, and K.M. Weiss, *Hum. Biol.* **65**, 1–6 (1993).
13. L. Roberts, *Science* **252**, 1614–1617 (1991).
14. L. Roberts, *Science* **257**, 1204–1205 (1992).
15. L. Roberts, *Science* **258**, 1300–1301 (1992).
16. H.T. Greely, *Human Genome Diversity Project — Summary of Planning Workshop 3(B): Ethical and Human Rights Implications*, 1993.
17. H.T. Greely, in B.M. Knoppers, ed., *Human DNA Sampling: Law and Policy — International and Comparative Perspectives*, Kluwer Law, The Hague, 1997.
18. L. Roberts, *Science* **258**, 1300–1301 (1992).
19. Rural Advancement Foundation International, *Home Page*, Available at: *http://rafi.ca*
20. Rural Advancement Foundation International, *RAFI Communiqué*, May 1993.
21. Rural Advancement Foundation International, *HGDP Hit List*, Available at: *http://rafi.ca*
22. Rural Advancement Foundation International, *Indigenous Person from Papua New Guinea Claimed in U.S. Government Patent* (1995), Available at: *http://www.rafi.ca*
23. H.T. Greely, *Cultural Survival Q.* **20**(Summer), 54–58 (1996).
24. L.A. Whitt, *Oklahoma City Univ. L. Rev.* **23**, 211–259 (1998).
25. D.R.J. Macer, in B.M. Knoppers, ed., *Human DNA Sampling: Law and Policy — International and Comparative Perspectives*, Kluwer Law, The Hague, 1997.
26. Foundation on Economic Trends, the Rural Advancement Foundation International, the World Council of Indigenous Peoples, the Third World Network, and the Cordillera Peoples' Alliance, *Petition to Dr. Harold Varmus*, June 30, 1994.
27. North American Regional Committee of the Human Genome Diversity Project, *Houston Law Rev.* **33**, 1431–1473 (1997).
28. UNESCO, International Bioethics Committee, Subcommittee on Bioethics and Population Genetics, *Bioethics and Human Population Genetics Research* (November 15, 1995), Available at: *http://www.biol.tsukuba.ac.jp/~macer/PG.html*
29. D.R.J. Macer et al., *Nature* **379**, 11 (1996).
30. B.M. Knoppers, M. Hirtle, and S. Lormeau, *Genomics* **50**(3), 385 (1998).
31. B.M. Knoppers, ed., *Human DNA Sampling: Law and Policy — International and Comparative Perspectives*, Kluwer Law, The Hague, 1997.
32. M.S. Frankel and A. Herzog, *Human Genome Diversity Project: The Model Ethical Protocol as a Guide to Researchers*, American Association for the Advancement of Science, Washington, DC, 1998.
33. Workshop on Anthropology, Genetic Diversity, and Ethics Workshop, *Meeting Summaries* (1999), Available at: *http://www.uwm.edu/Dept/20th/projects/GeneticDiversity/index.html*
34. Senate Committee on Governmental Affairs, *Human Genome Diversity Project: Hearing before the Committee on Governmental Affairs*, United States Senate, 103rd Congr. 1st Sess., April 26, 1993.
35. National Research Council, *Evaluating Human Genetic Diversity*, National Academy Press, Washington, DC, 1997.
36. E. Pennisi, *Science* **278**, 568 (1997).
37. C. MacIlwain, *Nature* **389**, 774 (1997).
38. *Nature* **390**, 221 (1997).
39. W.J. Schull, *Science* **279**, 10 (1998).
40. W.J. Schull, *Nature* **390**, 221 (1997).
41. E. Marshall, *Science* **284**, 406 (1999).
42. F.S. Collins, L.D. Brooks, and A. Chakravarti, *Genome Res.* **8**, 1229 (1998).
43. Available at: *http://locus.umdnj.edu/nigms/comm/order/assurance.html*
44. D. Hurtado and R. Braginski, *Science* **283**, 1439 (1999).
45. D. Normile, *Science* **283**, 1426 (1999).
46. L.L. Cavalli-Sforza, *Proc. Nat. Acad. Sci. USA* **95**, 11501 (1998).
47. A. Spaeth, *Time*, January 17, 45 (2000).
48. H.T. Greely, in R. Weir, ed., *Stored Tissue Samples: Ethical, Legal, and Policy Implications*, University of Iowa Press, Iowa City, IA, 1998.

49. E.W. Clayton et al., *J. Amer. Med. Assoc.* **274**, 1786–1792 (1995).
50. U.S. National Bioethics Advisory Commission, *The Use of Human Biological Materials in Research*, National Bioethics Advisory Commission, Bethesda, MD, 1999.
51. D. Harry, *Abya Yala News* **8**, 8 (1994).
52. J. Armstrong, *Nat. Cath. Rep.*, January 11–12, 27 (1995).
53. A.T.P. Mead, *Cultural Survival Q.* **20**(Summer), 46–51 (1996).
54. R. Liloqula, *Cultural Survival Q.* **20**(Summer), 42–45 (1996).
55. H.T. Greely, in D.J. Kevles and L. Hood, eds., *The Code of Codes*, Harvard University Press, Cambridge, MA, 1992.
56. D. Nelkin and M.S. Lindee, *The DNA Mystique: The Gene as Cultural Icon*, W.H. Freeman, New York, 1995.
57. S. Wolf, *J. Law Med. Ethics* **23**, 345 (1995).
58. J. Halley, *Stanford Law Rev.* **46**, 503 (1994).
59. M. Dando, *Biotechnology, Weapons, and Humanity*, British Medical Association, London, 1999, pp. 53–67.
60. U. Mahnaimi and M. Colvin, *Sunday Times (London)*, November 1 (1998).
61. W.J. Broad and J. Miller, *N.Y. Times*, Sunday, December 27, Sec. 4, p. 1 (1998).
62. *Moore v. Regents of the University of California*, 51 Cal. 3d 120; 793 P.2d 479; 271 Cal.Rptr. 146 (1990).
63. H.T. Greely, *Houston Law Rev.* **33**, 1397–1430 (1997).
64. W.L. Freeman, in R.F. Weir, ed., *Stored Tissue Samples: Ethical, Legal, and Policy Implications*, University of Iowa Press, Iowa City, IA, 1998.
65. E. Juengst, *Am. J. Hum. Genet.* **63**, 673–677 (1998).
66. C. Weijer, G. Goldsand, and E.J. Emanuel, *Nature Genet.* **23**, 275 (1999).
67. P.R. Reilly, *Am. J. Hum. Genet* **63**, 682–685 (1998).
68. P.R. Reilly and D.C. Page, *Nature Genet.* **20**, 15–17 (1998).
69. R.A. Grounds, *Cultural Survival Q.* **20**(Summer), 64–68 (1996).
70. Canada Tri-Council Working Group on Ethics, *Code of Conduct for Research Involving Humans*, Minister of Supply and Services, Ottawa, 1997.
71. M.W. Foster et al., *Am. J. Hum. Genet.* **64**, 1719–1727 (1999).
72. M.W. Foster, A.J. Eisenbraum, and T.H. Carter, *Genet. Test.* **1**(4), 269–274 (1997).
73. M.W. Foster and R.R. Sharp, *J. Law Med. Ethics* **28**, 41–51 (2000).
74. M.W. Foster, D. Bernsten, and T.H. Carter, *Am. J. Hum. Genet.* **63**, 696–702 (1998).
75. S.J. Gould, *The Mismeasure of Man*, Norton, New York, 1981.
76. D.J. Kevles, *In the Name of Eugenics*, Knopf, New York, 1985.
77. D.B. Paul, *Controlling Human Heredity: 1865 to the Present*, Humanities Press International, Atlantic Highlands, NJ, 1995.
78. A.L. Caplan, in T.F. Murphy and M.A. Lappe, *Justice and the Human Genome Project*, University of California Press, Berkeley, CA, 1994.
79. P.A. King, in J. Kahn, A. Mastroianni, and J. Sugarman, *Beyond Consent: Seeking Justice in Research*, Oxford University Press, New York, 1998.
80. P.A. Baird, *Perspect. Biol. Med.* **38**(2), 159–166 (1995).
81. J. Marks, *Anthrop. News* **72**, April (1995).
82. J. Marks, *Human Biodiversity: Genes, Race, and History*, Aldine De Gruyter, New York, 1995.
83. M. Lock, *Soc. Sci. Med.* **39**, 603–606 (1994).
84. M. Lock, *Chicago-Kent Law Rev.* **75**, 83 (1999).
85. D.B. Resnik, *Politics Life Sci.* **18**, 15 (1999).
86. H.T. Greely, *Politics Life Sci.* **18**, (2000).

See other entries OWNERSHIP OF HUMAN BIOLOGICAL MATERIAL; see also PATENTS AND LICENSING entries.

HUMAN SUBJECTS RESEARCH, ETHICS, AND INTERNATIONAL CODES ON GENETIC RESEARCH

BARTHA MARIA KNOPPERS
University of Montreal
Montreal, Canada

DOMINIQUE SPRUMONT
University of Neuchâtel
Neuchâtel, Switzerland

OUTLINE

INTRODUCTION

At least theoretically, medicine has benefited and continues to benefit from the participation of sick or healthy volunteers, in human research. Medical practice and research are inseparable elements of modern medicine. Yet, despite the unprecedented success of

medicine and biotechnology in the second half of the twentieth century, human subjects research remains associated with the atrocities committed by Nazi doctors during World War II (1), the dreadful biological warfare experiments conducted by the Unit 731 of the Imperial Japanese Army (2), not to mention the infamous radiation experiments conducted and/or financed by the U.S. government and by governments of other countries (i.e., UK, Switzerland) at the height of the cold war (3). As Jean Bernard, the first chairman of the French National Ethics Committee, has said "Human experimentation is morally necessary and necessarily immoral." The high hopes and expectations in the constant progress of medicine, as expressed by the growing efficacy and quality of health care, have not overridden the fears of society that human beings could be or are being grossly abused for the sake of science. As noted by Jay Katz in the introduction to his comprehensive case book on human experimentation: "When science takes man as its subjects, tensions arise between two values basic to Western society: freedom of scientific inquiry and protection of individual inviolability" (4, p. 1). The primary goal of the regulation of human experiments is indeed to ensure protection of the rights and welfare of human subjects.

As in many other fields of medical practice today, there is a clear shift toward legal regulation (5). By way of national and international legislation, a growing number of detailed guidelines must be followed by all participants in research activities; this means investigators, sponsors, monitors, ethics review boards, research institutions (hospitals, universities, etc.), local and national authorities. The complex regulatory framework is aimed both at protecting the interests of the persons participating in research and at ensuring the quality of the research's results will meet with general public approval, indeed, that the research is "necessarily moral." Before considering in more depth the basic rules that apply to human subjects research as they are presented in international codes, we will review the nature and scope of current research. It is also important to understand how these rules evolved historically as well as the underlying ethical, professional, and legal issues.

It is against this background that the basic principles of research in human genetics should be examined. Of particular interest is the emergence of international ethical norms governing human genetics in the area of human subjects research in biotechnology (6). Positions and proposals have emanated from international and regional bodies such as the United Nations Education Science and Culture Organization (UNESCO), the World Health Organization (WHO), the Human Genome Organization (HUGO), the Council of Europe (CE), the European Commission (EEC), and the Latin American Human Genome Program (PLAGH). We will limit our discussion to the last decade and demonstrate that in the gradual evolution of ethical norms governing human subjects, the area of human genetics research is witnessing a movement from general principles to more refined and complex approaches, sometimes in contradiction with one another. Moreover, because of the personal, familial, and social nature of genetic information, some controversial areas of application are "gen-ethics," which test the founding ethical principles governing human subjects research. The immediate conclusion to be drawn is that the increasing multiplicity, complexity, and specificity of ethics in human subjects research may well lead to losing sight of the fundamental ethical principles, if not undermine them.

NOTIONS AND TERMINOLOGY

Ties Between Medical Practice and Research

In the now classical essay "Training for Uncertainty," Renée Fox analyzed the process by which medical students learn to cope with their uncertainty which, by and large, remains the main (if not only) certainty in medical knowledge. Indeed, the proportion of medical care which relies on solid scientific or empirical evidence varies from only 10 to 50 percent depending on the authors (7). Three basic types of uncertainty can be identified:

> The first results from incomplete or imperfect mastery of available knowledge. . . . The second depends upon limitations in current medical knowledge. . . . A third source of uncertainty derives from the first two. This consists of difficulty in distinguishing between personal ignorance or ineptitude and the limitations of present medical knowledge (8, p. 20).

The first and third types of uncertainty can be limited by the proper selection, training, and continuing education of physicians, while the second type calls for research and the collection of empirical data. Research appears in the latter sense as a means to limit uncertainty and to acquire new, generalizable knowledge. Paradoxically, newly acquired knowledge is per se a source of uncertainty as it often raises more questions than it answers. The constant quest for new knowledge and the expected progress that should result from it is indeed an intrinsic element of scientific thinking (9).

Without elaborating on this point further, it is important to note that research is the basis of medical knowledge, and it is an activity that should be better promoted to give everyone the chance to benefit from it. Even if the recent history of medical research has been full of scandals and abuses, the positive dimension of research is beginning to be accepted. There has been a dramatic change in the general perception of research especially since the 1980s. The major change came from the AIDS epidemic which caused the gay community to put pressure on the scientific community and the government authorities to funnel research into that area. Feminist groups have also successfully promoted the idea that women should no longer be systematically excluded from research protocols, which formerly biased results. Finally, developing countries have asked to receive more support in research activities that are specific to their needs.

Distinctions Between Medical Practice and Research

In 1974 the U.S. Congress created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (the Commission). Its primary

task was to establish "the boundaries between biomedical and behavioral research involving human subjects and the accepted and routine practice of medicine" (10). After extensive consideration of the topic, the Commission concluded:

> For the most part, the term "practice" refers to interventions that are designed solely to enhance the well-being of an individual patient or client and that have a reasonable expectation of success. The purpose of medical and behavioral practice is to provide diagnosis, preventive treatment or therapy to particular individuals. By contrast, the term "research" designates an activity designed to test a hypothesis, permit conclusions to be drawn and thereby to develop or contribute to generalizable knowledge (expressed, for example, in theories, principles and statements of relationships). Research is usually described in a formal protocol that sets forth an objective and a set of procedures designed to reach that objective.
>
> When a clinician departs in a significant way from standard or accepted practice, the innovation does not, in and of itself, constitute research. The fact that a procedure is "experimental," in the sense of new, untested or different, does not automatically place it in the category of research. Radically new procedures of this description should, however, be made the object of formal research at an early stage in order to determine whether they are safe and effective. Thus, it is the responsibility of medical practice committees, for example, to insist that a major innovation be incorporated into a formal research project.
>
> Research and practice may be carried on together when research is designed to evaluate the safety and efficacy of therapy. This need not cause any confusion regarding whether or not the activity requires review; the general rule is that if there is any element of research in an activity, that activity should undergo review for the protection of human subjects (11, pp. 2–3).

This characterization of medical practice and research makes three basic distinctions. First, the primary goal of practice is to enhance the health and/or the well-being of an individual patient. By contrast, the investigator's goals include those of the research itself. He or she is somehow a double agent whose two masters are research and practice. Even where one does not behave contrary to the interests of one's subjects, the investigator does not act exclusively in their interests. Levine speaks in this regard, of "practice for the benefits of others" (12) which includes not only research with human subjects but also organ or tissue donation, vaccine programs, and participation in the training of health professionals. Second, the doctor–patient relationship is highly personal in the sense that all the activities of the practitioner should be based exclusively on the specific needs and interests of the patient. By contrast, an investigator must strictly follow the procedures fixed in the research protocol. Third, research should be based, in principle, on a written protocol defining its purpose, goals, and means. This research protocol is essential in assessing scientific validity at every step of the research process. It is necessary not only to guarantee the quality and reliability of the research results but also to protect the human subjects against unnecessary and unpredicted risks and burdens.

Questionable Terminology: Therapeutic Research

In 1984 Taylor and his colleagues published an enlightening study analyzing the reasons physicians tend to avoid including all eligible patients in relevant research (13). Of particular import was their discovery that only 27 percent of participating physicians maximized inclusion. The authors cited the following as the most common reasons for this finding:

> (*1*) concern that the doctor–patient relationship would be affected by a randomized clinical trial (73 percent), (*2*) difficulty with informed consent (38 percent), (*3*) dislike of open discussions involving uncertainty (22 percent), (*4*) perceived conflict between the role of scientist and clinician (18 percent), (*5*) practical difficulties in following procedures (9 percent), and (*6*) feelings of personal responsibility if the treatment were found to be unequal (8 percent) (13, p. 21).

This research points out the uneasiness of doctors acting as investigators to the detriment of their role as healers. If the term "therapeutic research" helps physicians to better accept their ambivalent role, it is a source of confusion for the patients–subjects. Research and therapy are fundamentally different. It is contradictory to speak of "therapeutic research," and this term should be avoided (14). The term is ambiguous as it implies some therapeutic benefits for the research subjects, regardless of the fact that benefits are, by definition, hypothetical. It creates confusion about the exact role of physicians involved in research activities who do not clearly disclose the fact that in the research setting they are not acting as healers but as investigators. By the term "therapeutic research," the physician is inviting the patient to participate in research, and the emphasis is on his or her latter role rather than the former. The patient may not be aware of this fact, that he or she is involved in research and that a valid consent is needed.

Categories of Research with Human Beings

The activities covered by the expression "human subjects research" are in fact very broad. A first distinction can be drawn from the term "human subjects." Does it refer solely to research with persons, meaning a person between the moment of birth and the instant of death or should it be extended to research done on fetuses and embryos as well as on human cadavers? Research with fetuses and embryos is, in principle, treated separately from research with human subjects. In general, vulnerable groups are provided special attention and protection in the regulation of research. This is the case for children, psychiatric patients, and prisoners. Minority members are also subject to specific regulation, not only to protect them from research risks and abuses but also to ensure that they have access to research protocols specific to their needs. In summary, the regulation of human subjects research varies according to the vulnerability of the group of subjects involved as well as the nature and degree of expected risks and discomfort.

Concerning the nature of research risks, a distinction should be made between (*1*) clinical research, namely research done "at the bed side," implying direct contact

between the research subjects and the investigator, (*2*) epidemiological research which is based on medical data usually collected for other purposes than the research itself, and (*3*) research on biological material of human origin. Human subject research usually refers to clinical research and the regulation of research is generally designed to address the specific problems raised by this type of research. Yet there is increasing attention being directed toward the two other types of research. Epidemiological research raises particular problems for the respect of privacy and confidentially. Its regulation is linked to the problem of data protection. It also provides directives on the way to collect and assess the informed consent of the subjects. Research on biological material lies somewhere in the middle, depending on whether the material being used has been collected for research purposes. We will see that the development of gene therapy and the mapping of the human genome feed a growing concern about how to regulate this type of research.

ETHICS, PROFESSIONAL RULES, AND LAW

After World War II, 23 doctors, physicians, and high ranking officials of the Ministry of Health in the Nazi government were brought before an international court founded under the principles of public international law. During this trial, which was held in Nuremberg, ten basic principles to be followed in the conduct of research were enumerated. This became known as the Nuremberg Code. Yet there has remained uncertainty as to the Code's specific nature whether its principles are ethical, legal, or both. The Nuremberg Code is usually presented as the first international code of ethics in the field of research. It has had tremendous influence on the codes and rules that followed especially in the drafting of the Declaration of Helsinki adopted by the World Medical Association in 1964. One question remains unanswered, however, at least for the medical profession. Is it strictly a codification of ethical principles or is it legally binding? Interestingly most lawyers would agree on the legal nature of the Nuremberg Code (15), while physicians would rather consider it an ethical code.

Historically two sets of rules dealing with human research had been promulgated earlier in Germany, and they were, to some extent, more detailed than the Nuremberg Code. They are the 1900 Directives to the Directors of Clinics, Out-Patient Clinics and Other Medical Facilities formulated by the Prussian Ministry of Religious, Educational, and Medical Affairs and the 1931 Guidelines on Innovative Therapy and Scientific Experimentation of the Reich Minister of the Interior. These rules were either simply not applied by doctors or were applied in such a confidential way that they could hardly be considered as the expression of a common practice by the medical profession. The Nuremberg Code was also largely inspired by guidelines published only a few months earlier by the American Medical Association. But while the judges who elaborated the Nuremberg Code felt that they were only codifying or summarizing the common rules of the medical profession, physicians have long refused to admit that this intrusion in the regulation of their research activities had any legal consequences. By coining the Nuremberg Code as an ethical code, they were emphasizing that it had no binding force, and therefore that its application was open to interpretation. Its only binding force was derived from the fact that it was to be considered as the expression of the state of the art by the professionals. From the beginning of modern regulation of human research subjects, rules were confined by the medical profession to the restrictive scope of professional ethics. Such interpretation was necessary, if not indispensable, to maintaining the independence of the researchers from any unwanted intervention by lawyers and the legislature.

We agree with Dr. Leo Alexander when he states: "The Nuremberg Code covers all these contingencies in a much more specific manner than subsequent formulations such as the Helsinki Resolution (sic)" (16, p. 396). In fact the first goal of the Helsinki Declaration was not to protect the freedom and rights of the human subjects but rather to allow the continuation of human experimentation: "Because it is essential that the results of laboratory experiments be applied to human beings to further scientific knowledge and to help suffering humanity, the WMA has prepared the following recommendations. . .." This difference in apprehending the regulation of human subjects research whether it is viewed in a legal or in a medical perspective should be better acknowledged. To a great extent there is a misunderstanding about the fact that when there is no specific legal norm, one should apply the general principles of the law. In short, the absence of statutory laws does not mean the absence of law. This has created the illusion that "ethical codes" were filling a lacunae in the law, while they were only making the regulatory framework for research more complete and precise.

INTERNATIONAL CODES IN THE REGULATION OF HUMAN SUBJECTS RESEARCH

Introduction

The regulation of human subjects research is characterized by myriad national and international rules that can be qualified as ethical, professional, or legal. Two of the most preeminent of these norms have already been mentioned, namely the Nuremberg Code and the Declaration of Helsinki. Concerning the latter, it should be noted that it was revised in 1975 (Tokyo), 1983 (Venice), 1989 (Hong Kong), and 1996 (Somerset West). Other documents of reference at the international level should also be cited: Article 7 of the United Nations Covenant on Civil and Political Rights (1966), the International Guidelines for Biomedical Research Involving Human Subjects first adopted in 1982 and revised in 1993 by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with WHO. The so-called Belmont Report issued in 1978 by the U.S. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research should also be mentioned. It defends a conception of bioethics that has greatly influenced modern ethical reasoning and is based on the balancing of four basic ethical principles: the principles of justice,

respect for persons, beneficence, and nonmaleficience. This is only a short list of the existing guidelines (18). There are guidelines that concern specific types of clinical trials, for instance, the Declaration of Madrid adopted in 1996 by the World Psychiatric Association or the various guidelines for good clinical research practice (19) primarily designed for drug trials such as the WHO Guidelines for Good Clinical Practice (GCP) for trials on pharmaceutical products (1995) (17) and the "International Conference on Harmonisation" (ICH) Guideline for Good Clinical Practice (ICH–GCP). Other guidelines have only a regional scope, for instance, the Nordic Guidelines for Good Clinical Trial Practice (1989), or the present draft European Union Directive on Drug Trials (20) as well as the CE Convention on Human Rights and Biomedicine (21), not to mention the numerous national laws and codes of ethics. To complete this broad picture of the regulation of biomedical research, two observations still need to be made: First, these codes and guidelines, such as the Declaration of Helsinki and the CIOMS guidelines, are being regularly revised; second, if some of these codes and guidelines should be considered as mainly ethical and/or professional norms, others, such as most of the GCP guidelines, are more or less statutory rules.

Even if guidelines and laws are very heterogeneous, a consensus emerges from them all on some basic requirements regarding the protection of human research subjects during clinical trials. Among theses, the most commonly identified are the following:

- Review of the research project by a competent ethics review board (ERB)
- Fair selection of the human research subjects (special protection of vulnerable population)
- Free, informed consent of the human research subjects
- Respect for the privacy of the human research subjects and for the confidentiality of clinical data
- Favorable balance of harms and benefits
- Compensation for research-induced injury
- Sound, scientific design of the research based on sufficient data from previous nonclinical and clinical studies
- Qualification and experience of the investigator
- Adequacy of the resources available (time, staff, facilities, finances)

These norms are the core elements of all regulation concerning the protection of human research subjects. To better understand the nature and scope of the aforementioned codes, we propose to analyze them in light of the rules of informed consent and ethical review mechanisms. This will also serve as an occasion to clarify the specific links as well as discrepancies between these codes.

Informed Consent

There are few texts like the 'Nuremberg Code' that express with such strength and precision the principle of informed consent freely given by human subjects. It is worth quoting *in extenso*:

> The voluntary consent of the human subject is absolutely essential. This means that the person involved should have legal capacity to give consent; should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, overreaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him to make an understanding and enlightened decision. This latter element requires that before the acceptance of an affirmative decision by the experimental subject there should be made known to him the nature, duration, and purpose of the experiment; the method and means by which it is to be conducted; all inconveniences and hazards reasonably to be expected; and the effects upon his health or person which may possibly come from his participation in the experiment.
>
> The duty and responsibility for ascertaining the quality of the consent rests upon each individual who initiates, directs, or engages in the experiment. It is a personal duty and responsibility which may not be delegated to another with impunity.

It is interesting to note that the first version of the Declaration of Helsinki, adopted in 1964 by the World Medical Association and claiming a relationship to the 'Nuremberg Code', does not mention the rule of informed consent among its basic principles of research ethics. Furthermore, though it provides that the informed consent of human subjects should be obtained before beginning a study with no direct therapeutic benefit, the Helsinki Declaration only requires that "if at all possible, consistent with patient psychology, the doctor should obtain the patient's freely given consent after the patient has been given a full explanation" (23, p. 473).

This wording seems to admit the possibility of applying the therapeutic privilege in the field of clinical trials. By therapeutic privilege, we mean the physician's ability to withhold some information from a patient based on the rationale that it could be detrimental for the patient. Such an interpretation contradicts the first principle of the 'Nuremberg Code'. It has been partly abandoned in the later versions of the Helsinki Declaration, the 1996 version stating clearly that "in any research on human beings, each potential subject must be adequately informed.... The physician should then obtain the subjects' freely-given informed consent, preferably in writing." Yet a 1999 proposed revision raises some serious concerns as it would allow a waiver of written informed consent if the ethics review board determined that the research risks are slight or if the procedures to be used in the research are customarily used in medical practice without informed consent (24). If adopted, this would represent a step backward for the respect of the principle of free and informed consent of the human research subjects (25).

The respect of the human subject's autonomy, and thus of the principle of informed consent, is also a cornerstone of the CIOMS guidelines. Of particular interest is the fact that the CIOMS guidelines recognize that the Declaration of Helsinki is "the fundamental document in the field of ethics in biomedical research," and a copy

of the 1989 version of this text can be found in their annex. Nevertheless, they differ on several points, in particular, on the requirement of the subject's informed consent. Guideline 8 of the CIOMS guidelines which deals with the question of research involving subjects in developing communities provides that: "every effort will be made to secure the ethical imperative that the consent of individual subjects be informed. . . ." Yet, in the commentary following guideline 8, it states: "For example, when because of communication difficulties investigator cannot make prospective subjects sufficiently aware of the implications of participation to give adequately informed consent, the decision of each prospective subject should be elicited through a reliable intermediary such as trusted community leader. In some cases other mechanisms, approved by an ethical review committee, may be more suitable." Thus, after stressing the need to respect individual informed in accordance with the basic principle I, 9 of the Declaration of Helsinki as amended in 1996, the CIOMS guidelines open the door for a broad exception in case of research conducted in developing countries (26). Which norm should apply in a given situation? Should it be the first principle of the Nuremberg Code, the basic principle I, 9 of the Declaration of Helsinki, and also guideline 1 of the CIOMS guidelines, or on the contrary, guideline 8 of the CIOMS guidelines interpreted in light of its commentary? The situation becomes more complicated by the fact that the CIOMS guidelines do not make explicitly reference to the Declaration of Helsinki. To what extent then should the new amendments of the Declaration of Helsinki be taken into consideration by someone referring to the CIOMS guidelines? Such questions arise for many other codes and guidelines that refer to the Declaration of Helsinki or to the CIOMS guidelines, if not to both (27).

In short, the great variety of rules of conduct concerning the protection of human research subjects at the international level does not always contribute to a better understanding of what ethical rule should finally prevail in a given situation. The need for a more effective coordination of these numerous codes is a growing concern, especially at a time several of them are being revised. The very existence of so many documents of reference raises some questions on their role and objectives. As stated by Jay Katz concerning ethical codes of conduct more than 30 years ago:

> The proliferation of such codes testifies to the difficulty of promulgating a set of rules which do not immediately raise more questions than they answer. By necessity these codes have to be succinctly worded and, being devoid of commentary, their meaning is subject to a variety of interpretations. Moreover, since they generally aspire to ideal practices, they invite judicious and injudicious neglect. Consequently, as long as they remain unelaborated tablets of exhortation, codes will at best have limited usefulness in guiding the daily behavior of investigators (28).

If these codes are still relevant for the diminishing number of countries and regions of the world where there is presently no statutory regulation of clinical trials, there is concern about their effectiveness, especially in view of their built-in contradictions. In the near future, it is foreseeable that the issue will become more a matter of the proper training of all participants in clinical trials, the investigators, members of ethical review boards, and the human subjects. The existing codes could then be regarded more as a source of inspiration than of regulation. The need to frequently revise the codes to meet the demand of the research would be less important than the need to assess that everyone involved in research with human subjects is working according to high ethical standards. The question then would be not so much revising the codes but ensuring that the fundamental ethical principles that underlie them are effectively implemented.

GENETHICS

The continuing drive to greater levels of specificity has led to the proclamation of "gen-ethics," which is quickly moving toward more detailed "gen-policies." Some examples of the latter are the legal status of DNA, DNA banking, patentability, genetic research involving children, and, confidentiality. Before turning to specific "gen-policies," we will review the common interpretation of the ethical principles discussed in the first part of this article within the context of human genetics. Broadly speaking they are autonomy, privacy, justice, equity, and quality, all relating to respect for human dignity (29).

Autonomy

The principle of autonomy has found its most recent expression in UNESCO's 1997 Universal Declaration on the Human Genome and Human Rights. The simple yes–no of participation has moved to include not only the notion of free and informed consent to participation [Article 5(b)] but also the choice to be informed or not of the results [Article 5(c)]. The probabilistic nature of most predictive genetic information (to say nothing of its personal, familial, and social nature) requires that the expression of individual autonomy include this choice. The problem with this new approach, however, is that little is known of how much information must be communicated for this right to be exercised. Would too much information unduly restrict or undermine the "right" not to know, and too little that of a truly informed consent not to know? Another addition to the second requirement of consent is that of requiring a specific consent for the banking of a DNA sample, often collected and conserved in the past without such consent (30). Only newborn screening for treatable disorders constitutes a valid exception to individual informed consent according to WHO (31). Finally, another "new" consent appears in the preamble to the European Directive on the Legal Protection of Biotechnological Inventions (32). The preamble states that a patent application on an invention using human biological material of human origin must be from a person who has had the "opportunity of expressing [a] free and informed consent thereto, in accordance with national law" (para. 26). While laudatory in its transparency, the weakness of this formulation lies in the latter part that refers to national law. This

is because only few national laws have addressed the issue of payment to DNA "donors," although this may be covered by human tissue gift legislation that does not specifically exclude DNA from its mandate. No law specifically requires an explicit consent to eventual commercialization (33). Furthermore it remains to be seen if such consent will be necessary for each possible commercial application (impossible?) or whether a simple notification will suffice and, if so, constitute a preliminary condition of participation.

Privacy

Two genetic-specific interpretations of the principle of privacy merit mention here. The first is the increased emphasis on the confidentiality of genetic information especially as concerns third parties (owners and employers). Article 9 of the UNESCO Declaration maintains that exceptions to the confidentiality of genetic information can only be prescribed by law, thus excluding individual consent. Indeed, this could be because consent to access by insurers and employers is not totally free. Thus WHO has proposed: "Genetic information should not be used as the basis for refusing employment or insurance. Exceptions would have to be legally defined (34). The formulation found in the European Convention on Biomedicine (21) specifying that genetic tests should only be performed for "health purposes" (Article 12), that is, medical indications, is particularly enlightening. Yet, if this information is in the medical record due to participation in research, an individual applying for insurance or employment will certainly have "consented" to access to such confidential information. Hence the importance of keeping participation in genetic research out of the medical record.

Finally, WHO's 1997 Proposed International Guidelines (31) and HUGO's Statement on DNA Sampling: Control and Access (30) make a specific exception for professional disclosure to at-risk family members for serious, treatable, or preventable conditions where the patient or research participant refuses to do so. This issue merits further discussion.

Justice

The principle of justice includes the notion of distributive justice, and in the context of genetics this involves not only living persons (e.g., vulnerable populations) but also future generations. Both the inclusion and protection of vulnerable populations in genetic research is reinforced through the requirement of a legal authorisation by a third party and a direct benefit to the incompetent person or, in the absence of the latter, only research of minimal risk Article 17 (2)(ii) (22). Ironically, this protection, while necessary, may inadvertently create indirect discrimination on economic grounds. This is because only those with the legal foresight and financial wherewithal to provide in advance for such a legal mandate specific to research can be included. Thus a competent person would have to undertake the legal procedures to name a future legal representative with a mandate specifically for participation in research. It goes without saying that in the absence of this advance mandate, it is very difficult (emotionally and financially) for family members to deliberately have an incompetent family member legally declared incompetent so that person can benefit from participation in research!

As for the interests of future generations, Article 24 of the UNESCO Declaration mentions that germ-line therapy, which could irrevocably affect future generations "could be contrary to human dignity" (35). This addition was only added in July 1997 when the government representatives met to approve the Declaration prepared by the IBC. The Declaration of Manzanillo would not grant legal representatives the power to authorise interventions in the human genome but would entrust such decisions to a neutral body (Principle 6(e)).

In contrast, neither the 1997 Proposed International Guidelines of the WHO nor its 1999 Draft Guidelines on Bioethics refers to germ-line therapy. As mentioned, the Declaration of Manzanillo (1996) would establish a "neutral" multidisciplinary authorizing body to consider such requests (Principle 6 (d)) (37). This is a subject that should be at the forefront of public debate, or as was the case with cloning, legislation will be inadequate (due to hasty preparation), too comprehensive, or unduly restrictive.

Equity of Access

Both the Declaration of Manzanillo and the WHO Proposed International Guidelines stress the need to respect equity of access to services, that is, access according to need and not economic means. Closely related to the principle of justice, this principle could also serve as the ethical underpinning for the concept of benefit-sharing first found in HUGO's statement on the Principled Conduct of Genetic Research (33). Likewise the 1997 Proposed International Guidelines by WHO maintain that if genetic knowledge leads to the development of "a diagnostic test or new therapies, equity requires that the donors, or the community generally, should receive some benefit" (31). Finally, while not specifically mentioning the principle of equity or access to genetic research or testing, UNESCO's Universal Declaration does provide in Article 12 that the "benefits" of advances in biology, genetics and medicine should be available to all (35).

Quality Control

Building on the principle of "basic science" of the Helsinki Declaration (25), the emergence of the notion of competence and quality control as a prerequisite for permitting research in a given area is a welcome development. Both HUGO in its 1996 Statement on the Principled Conduct of Genetic Research and WHO's 1997 Proposed International Guidelines see the provision of quality control through accreditation and surveillance of laboratory services. Requiring a medical indication for access would also serve to limit the number of tests being offered or advertised by those not trained in the all-important communications or counseling areas. With the advent of predictive testing, counseling takes on extraordinary importance. Still missing, however (and this is not unique to genetics research) are the quality control

aspects (e.g., monitoring/surveillance), once a protocol has been approved.

In short, the translation of general ethical principles governing human research to human genetics has led not only to their further refinement but, at least as a leitmotiv, to their amplification. Indeed, in the area of consent, the emergence of the right not to know, of the need to obtain consent for DNA banking, and perhaps to patenting, if not at least to eventual commercialization, are welcome developments. The principle of the privacy of one's genetic makeup while reinforced as concerns employers and insurers is nevertheless severely tested by the legitimate needs of at-risk family members. In contrast, overprotection of incompetent research subjects in the requirement of a legally recognized representative for research purposes may discriminate against certain families and those with late-onset genetic conditions. Finally, the last two principles, of equity and quality, merit further definition in the context of genetic research. It would be salutary, to say the least, if it was the fear of abuse of testing and misuse of genetic information that has led to the further development of guidelines in this area, since quality control and ongoing surveillance and monitoring are problems endemic to all research.

GEN-POLICIES

The specific translation of ethical principles into gen-ethics further contextualizes into policies concerning issues, such as, the DNA banking, patenting, research involving children, and confidentality. At this more specific domestic level, the interpretation of gen-ethics is less harmonious. Diversity is a cultural and political fact, and it raises problems for multicentered, international trials.

Status

The issue of the status of human genetic material has both ethical and legal significance. Ethics are involved because status distinguishes the human "source" from other elements of the body or products used in research. This is especially evident in the "common" or "collective" characterization at the level of the whole human genome and that of "familial" even at the level of the individual. The law is involved because of the classical division between persons and things (37).

The characterization as "common heritage" at the collective level can influence the treatment of human genetic material in international trade and patenting disputes. The characterization as "familial" can serve to sensitize researchers and research participants to the familial implications of genetic testing. Such distinctions form the underpinning for the construction of an ethical framework that is not solely individualistic in nature, although the choice to participate in research or to be tested must necessarily remain so.

DNA Banking

Nowhere has respect for individual choice based on personal values blossomed more than in the options offered in the context of DNA banking. From the blanket yes or no of a decade ago, participants are now not only offered the choice to be part of disease-specific or more extensive research collaboration, to have their DNA anonymized or destroyed, or to set limits on length of storage, but are also warned of socioeconomic risks (e.g., employment/insurance), possible stigmatization, and commercialization. It is interesting that regardless of the status of the human genetic material (property or person), the same choices are offered.

Theoretically any person with a property right, as proposed in the model Genetic Privacy Act of the United States, must either renounce future financial interests or claim a benefit. Absent an extremely rare genotype, the latter approach would not be practical. The time it takes for a discovery, an invention, and possible profits usually spans a decade and involves thousands of participants in different countries. Only recently has there been any move to simplify and standardize DNA banking choices (38). More interesting is the suggestion (albeit with some dissent) by the National Bioethics Advisory Commission of the United States to allow participants to agree to other future research without further specification (39). This recommendation does not require anonymization as a condition of such an open-ended choice. Last, the international Ethics Committee of HUGO has proposed notification for the anonymous use of genetic material removed as part of routine care (30). This evolution is significant in that the "reification" and "sacralization" extremes in DNA banking policies seem now to have reached a compromise that respects individual choice but does not accord a higher status to DNA than to the person.

Patenting

UNESCO's Universal Declaration on the Human Genome and Human Rights epitomizes both the ethical and legal positions concerning the trend toward the commercialization of genetic research around the world. Article 4 maintains that "the human genome in its natural state shall not give rise to financial gains" (35). While genetic material in its natural state cannot be bought or sold, this should not be confused with the issue of patentability. As the 1998 European Directive on the Legal Protection of Biotechnological Inventions stated (Article 5):

> 1. The human body, at the various stages of its formation and development, and the simple discovery of one of its elements, including the sequence or partial sequence of a gene, cannot constitute patentable inventions (32).

Nevertheless,

> 2. An element isolated from the human body or otherwise produced by means of a technical process, including the sequence or partial sequence of a gene, may constitute a patentable invention, even if the structure of that element is identical to that of a natural element (33).

This ruling is buttressed by two international instruments that bear mentioning. Both the General Agreement on Tariffs and Trade (GATT) agreement and the North American Free Trade Agreement (NAFTA) include the

ethical filter of excluding inventions whose commercial exploitation would be contrary to public policy, a criterion also found in the European Patent Convention of 1973 (40).

Nevertheless, despite both the potential of this ethical filter and the clarification provided by the European Directive, it was not until 1999 that the U.S. Patent and Trademark Office stated that applicants should "explicitly identify a specific and substantial utility." Therefore, for genes and gene products, there is a need "to specify an immediate and identifiable benefit to the public" (41). Such clarification is necessary because patent applications on sequences and on genes have been seriously hampering research and possible therapeutic applications (33).

Genetic Research Involving Children

The earlier discussion of the principles governing the inclusion of vulnerable populations, such as, incompetent adults, raised the issue of potential discrimination by automatic exclusion in the absence of a "legally" appointed representative. When guidelines do not differentiate between incompetent adults and children, the same issue of possible exclusion occurs with children. In the majority of legislation or policy statements, the parents are legal guardians by law. Thus, while the potential inclusion of children is possible (and may even lead to overparticipation), provided their best interests are served by parental authorization, the door is not completely open. Indeed, not only the best interest criterion but also the notions of benefit and risk come into play.

As concerns benefit, the requirement is that in principle, some therapeutic benefit must be expected as well as a favorable risks–benefit ratio. Even in the absence of direct benefit, when the results are "capable of conferring benefit to other persons in the same age category or afflicted with the same disease or disorder or having the same condition ... [and] entails only minimal risk and minimal burden" [Article 17(2)] (21), the research may proceed. This is important in that many genetic conditions are specific to childhood (43) and the automatic exclusion from research would seriously harm both affected children and those potentially at risk. Indeed, overprotection (43), while an understandable reaction to the abuses of the past or even as a precautionary measure when faced with "genetic" unknowns, is becoming a major ethical issue. That being said, there are specific limitations on the genetic testing of children in the absence of available treatment or prevention. These limitations include no testing for the benefit of other family members or for carrier status, nor for late-onset, susceptibility, or predictive purposes.

These restrictions, however, imply the availability of a test, whereas most research aims to find a gene that is either a cause or a determinant among other factors. Thus, in theory, the restrictions above would apply in the situation where DNA is taken with a view toward developing tests that situate the issue within the wider realm of the legitimacy of genetic testing for carrier status or prediction, in general, and where there is no therapeutic potential (prevention, reproductive information, or treatment). It is here that anonymity (removal of all identifiers with some clinical and demographic information remaining) of DNA samples taken in research involving children might best serve their interests and still respect the above-mentioned principles. Samples that cannot be traced can still serve the pediatric community writ large. When the child reaches maturity/majority and treatment becomes available or even reproductive genetic information, the choice to be tested will be an autonomous and informed one.

Confidentiality

Equally sacrosanct to the principle of consent is that of confidentiality. The principle of confidentiality has acquired greater importance in human genetics where the risk of socioeconomic discrimination and stigmatization could take hold. Yet genetic information is at once personal, familial, and social. This raises two further thorny issues: the protection of genetic information in the research setting and disclosure to at-risk family members. The first remains to be settled, and the second is slowly emerging into a consensus.

Currently the protection of genetic information in the research setting is no different from other research findings. When relevant to treatment or prevention, genetic information is placed in the medical record. But in the absence of such medical significance, patient consent is sought for inclusion in the medical record. Indeed, because of the sensitivity of genetic information, most research records are kept under secure conditions. Increased legislative protection of medical records is needed before this practice will change. Any eventual "normalization" of genetic information as medical information depends on such legislative protection. As genetics is integrated more into mainstream medicine, it may well force this issue of medical confidentiality (45).

On disclosure to family members in the situation where an at-risk family member refuses to warn another member, there has been some international agreement about the ethical obligations, although the law is not clear. Drawing on the notions of genetic material or information as familial, on the obligation to rescue, on professional ethics, or on a legal basis for disclosure, the consensus can be summarized as follows: (*1*) if the patient has repeatedly refused to warn an at-risk identifiable family member, (*2*) if that member is at high probability of a serious disorder, and (*3*) if the disorder in question can be prevented or treated, the physician could ethically breach confidentiality but is not legally obliged. At a minimum, however, such disclosure could be considered a legal privilege or a defense to an accusation of breach of confidentiality (45). It goes without saying that the normalization of genetic conditions and information may well solve this serious dilemma of breaching medical confidentiality and allow families to be less stigmatized and more open.

CONCLUSION

The 1997 Universal Declaration on the Human Genome and Human Rights maintains in Article 12(b) that "Freedom of research, which is necessary for the progress

of knowledge, is part of freedom of thought. . . ." (35). There is no doubt that freedom is fundamental to progress and it should be protected and promoted even in a controversial area like human genetics. Still the therapeutic and intellectual benefits of research are challenged by the ethical principles put in place to ensure that voluntary participation is kept in balance with risk–benefit ratios. The major issue of today might well be not that of adding more "rules" but rather constructing the processes and structures that allow us to remember and to realize the basic principles in the first place. The ethical principles should hold true (whatever their translation into more specific policies) in all scientific domains. Like basic legal principles, the ethical foundations of research have survived the test of time. The procedures in place for the implementation of ethical principles must be developed and observed in this time of rapid scientific upheaval.

BIBLIOGRAPHY

1. R.J. Lifton, *The Nazi Doctors, Medical Killing and Psychology of Genocide*, Basic Books, New York, 1986.
2. S.H. Harris, *Factories of Death*, Routledge, London, 1994.
3. *Final Report of the Advisory Committee on Human Radiation Experiments*, U.S. Government Printing Office, Washington, DC, 1995, Available at: *http://www.seas.gwu.edu/narchive/radiation*
4. J. Katz, *Experimentation with Human Beings*, Russel Sage Foundation, New York, 1972.
5. *Bioethics: From Ethics to Law, From Law to Ethics*, Publications of the Swiss Institute of Comparative Law No. 30, Schulthess, Zurich, 1997.
6. S. LeBris, B.M. Knoppers, and L. Luther, *Houston Law Rev.* **33**, 1363–1395 (1997).
7. J. Ellis et al., *Lancet* **346**, 407–410 (1995).
8. R.C. Fox, in *Essays in Medical Sociology, Journeys into the Field*, 2nd ed., Transaction Books, New Brunschwig and Oxford, UK, 1988, p. 20 (first published in *The Student-Physician*, Robert K. Merton et al., eds., Harvard University Press, Cambridge, MA, 1957, pp. 207–241).
9. J.F. Lyotard, *La condition postmoderne*. Editions de Minuit, Paris, 1979.
10. Public Law 93–348.
11. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, Publication No. (OS) 78-0012, 2–3, DHEW, Washington, DC, 1978.
12. R.J. Levine, *Ethics and Regulation of Clinical Research*, 2nd ed., Urban & Schwarzenberg, Baltimore, MD, 1986, p. 7.
13. K.M. Taylor, R.G. Margolese, and C.L. Soskolne, *Engl. J. Med.* **310**(21), 1363–1367 (1984).
14. K. Lebacqz, *Villanova Law Rev.* **22**, 357–366 (1976–1977).

    F. Rolleston and J.R. Miller, *IRB: Rev. Hum. Subjects Res.* **3**(7), 1–3 (1981).

    A.M. Capron and R.J. Levine, *IRB: Rev. Hum. Subjects Res.* **4**(1), 10–11 (1982).

    R.J. Levine, *Ethics and Regulation of Clinical Research*, 2nd ed., Urban & Schwarzenberg, Baltimore, MD, 1986, pp. 8–10.

    D. Sprumont, *La protection des sujets de recherche*, Staempfli, Berne, Switzerland, 1993, pp. 33–37.
15. C. Bassiouni, T.G. Baffes, and J.T. Evrard, *Revue Int. Droit Pénal* (also exists in english) **51**, 285 (1980).

    M. Belanger, *Droit international de la santé*, Paris, 1983, p. 43.

    J. Katz, *J. Am. Med. Assoc.* **276**, 1662 (1996).

    G.A. Grodin, *Health Hum. Rights* **2**(1), 7 ff (1996).

    P. Arnold and D. Sprumont, in U. Trõhler and S. Reiter-Theil, eds., *Ethics Codes in Medicine: Foundations and Achievements of Codification since 1947*, Ashgate, Aldershot, UK, 1998, pp. 84–96.
16. L. Alexander, in *Ann. N.Y. Acad. Sci.* **169**(344), 346 (1970).
17. *WHO Technical Report Series*, No. 850, 1995, pp. 97–131.
18. *Informal listing of selected international codes, declarations, guidelines, etc., on medical ethics/bioethics/health care ethics/human rights aspects of health*, Periodically updated by Sev S. Fluss, c/o CIOMS, WHO, 1211 Geneva 27.
19. E.F. Hvidberg, *Drugs* **45**(2), 172 (1993).
20. United Nations Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *Good Clinical Practice*, (O.J. 97/C 306/10).
21. Council of Europe, *Int. Dig. Hlth. Leg.* **48**(1), 99 (1997).
22. H.D.C. Roscam Abbing, *European J. Health Law* **1**(2), 148 (1994).
23. World Medical Association, *N. Eng. J. Med.* **271**(9), 473 (1964).
24. World Medical Association, *Proposed Revision of the World Medical Association Declaration of Helsinki*, Santiago, Chile, 1999 (WMA document: 17.C/Rev 1/98).
25. See T.A. Brennan, *N. Engl. J. Med.* **341**(7), 527 (1999).
26. J. Legemaate, *Eur. J. Health Law* **1**(2), 161–165 (1994).
27. WHO GCP, provision 1.2, *Ethical principles*.
28. J. Katz, *Daedalus* **98**, 482–483 (1969), also published in: *Experimentation with Human Subject*, P.A. Freund, ed., George Braziller, New York, 1970, pp. 293–314.
29. B.M. Knoppers and R. Chadwick, *Science* **265**, 2035–2036 (1994).
30. HUGO Ethics Committee, *Genome Digest* **6**(1), 8 (1999).
31. World Health Organization, *Proposed International Guidelines on Ethical Issues in Medical Genetics and Genetic Services*, Geneva, Switzerland, 1997.
32. European Union, *Official J. Eur. Communities* **13**, L 213 (July 6, 1998).
33. B.M. Knoppers, *Nat. Gen.* **22**, 23 (1999).
34. World Health Organization, *Draft Guidelines on Bioethics*, Geneva, Switzerland, 1999.
35. United Nations Educational, Scientific and Cultural Organization, International Bioethics Committee, *Universal Declaration on the Human Genome and Human Rights*, Paris, November 11, 1997.
36. Latin American Human Genome Program, *Int. Dig. Hlth. Leg.* **48**(3–4), 424 (1997).
37. B.M. Knoppers et al., *Genomics* **50**(3), 385–401 (1998).
38. B.M. Knoppers, M. Hirtle, and K. Cranley Glass, *Science* **286**(5448), 2277–2278 (1999).
39. National Bioethics Advisory Commission, *Research Involving Human Biological Materials: Ethical Issues and Policy Guidance*, Rockville, MD, August 1999.
40. European Patent Organization (Administrative Council), *European Patent Convention*, 1973.
41. Paul Smaglik, *Nature* **403**, 6765 (2000).
42. World Health Organization, *Human Genetics and Noncommunicable Diseases*, Fact Sheet No. 209, 1999, *http://www.who.int/inf-fs/en/fact209.html*

43. M. Hirtle, *Health Law J.* **6**, 43–81 (1998).
44. Mark A. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997.
45. American Society of Human Genetics, Social Issues Subcommittee, *Am. J. Hum. Genet.* **62**(2), 474–483 (1998).

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, ETHICS, AND RESEARCH ON CHILDREN

LORETTA M. KOPELMAN
Brody School of Medicine at East Carolina University
Greenville, North Carolina

OUTLINE

INTRODUCTION

Research with children as subjects can help them individually and collectively. Those who participate can benefit through access to programs or treatments otherwise unavailable to them. In addition research can be socially useful to gain better information about children's diagnoses, therapies, and prognoses. Without research with children as subjects, science cannot address children's unique needs, conditions, or reactions. Doctors may be reluctant to prescribe therapies tested only on adults for children, and insurance companies may refuse to pay for their use because safety and efficacy has not been shown for children. Insufficient information can result in undertreatment of children's conditions or unexpected adverse drug reactions.

What should be done, however, when it is socially useful to conduct a study that will provide extremely worthwhile information, yet some of the children will not be helped and might even be harmed? Should we say that a child's well-being always takes precedent, even if the harm to the child represents only a slight inconvenience, and the information to be gained is urgently needed? On the other hand, what risks are too great to be tolerated, and how should we assess and balance potential harms and benefits? For example, some genetic research involves screening people for certain conditions by taking a small blood sample. While this carries almost no physical risk, a variety of psychosocial harms could befall the participants. Testing may reveal many things people would prefer to remain unknown. They may find that the person's biological father is not the person everyone supposes or detect diseases that are likely to occur later in life. People may find that they are likely to get a late-onset disease like Huntington's disease, schizophrenia, breast cancer, Alzheimer's disease, and so on, and their lives will be changed forever by the test results. Even though they are entirely healthy and can do nothing to prevent the disease, a positive test for one of these late-onset diseases may expose them to prejudice and discrimination. They may be stigmatized, finding people unwilling to marry them, give them jobs, or insure them. Who should decide when children should participate in these or other studies and what standards should be used to base these decisions?

CHANGING ATTITUDES ABOUT RESEARCH PARTICIPATION

Until recently, research was regarded as a dangerous activity and certain vulnerable groups, especially children, were often excluded from studies for their protection. Because children were kept out of studies, however, they often could not obtain untested or experimental therapies, or investigational new drugs (INDs). Such restrictions were criticized as unfair for unjustly excluding some groups from receiving needed care and treatment. For example, regulatory obstacles including federal and state rules restricted the use of untested drugs such as AZT on children with AIDS until some of the rules were changed. However well meaning, these protective measures kept many children with AIDS from getting the only drug that could help them (1). In some cases people's best chance for good medical care is through participation in trials and access to these INDs.

Another issue of justice arose from the practice of automatically excluding certain groups, such as women and children, in order to make the subject population homogeneous. Investigators sometimes restricted those whom they allowed in studies to make it easier to analyze results. People of different ages, races, ethnic groups, or genders sometimes react to interventions differently because of their unique cultural, biological, or behavioral variations. The more homogeneous the subjects, the easier it is for investigators to analyze findings. It also helps hold down costs because studies with similar people require fewer subjects. This practice, however, made it difficult to generalize the findings of studies to other groups. For example, test results of drugs used to treat depression on white, middle-class men might not apply to other groups, especially children, because the causes of depression vary.

While children and adults often have the same diseases, if studies are not done with children, it is uncertain if the results apply to them. Yet they are often treated with the same drugs. The American Academy of Pediatrics concluded that only a small number of drugs and interventions used on both adults and children have had clinical trials performed on pediatric populations. Moreover a majority of drugs on the market have not even been labeled for pediatric populations (2). The National Institutes of Health (NIH) found that 10 to 20 percent of research inappropriately excluded pediatric populations (3). Insufficient information about products can result in undertreatment of children's conditions and adverse drug reactions, because clinicians must choose between prescribing drugs that do not have good information about safety and efficacy, or using therapies that are potentially less effective. Careful studies show that drugs commonly used among adults sometimes harm children, and this could not be discovered until testing was done on this population. For example, Chloramphenicol, a common antibiotic, caused many deaths in neonates because their immature livers were unable to tolerate it (4). Including children in research, then, may provide opportunities for the individual child, as well as benefits for children as a group.

In recent years, federal agencies tried to address problems of inappropriate "protection" and inadequate testing of pediatric drugs and interventions. The NIH adopted a policy that "children (i.e., individuals under age 21) must be included in all human-subjects research, conducted or supported by the NIH, unless there are scientific and ethical reasons not to include them. This policy applies to all NIH conducted or supported research involving human subjects. . ." (3, p. 2). Justification for excluding children from the research would be that the topic is irrelevant to children's health and well-being, that legal or regulatory restrictions exist prohibiting the inclusion of children, that information about them is already available, or that a separate study for children would be inappropriate due to the rarity of the condition in this population, the few number of the children who would be affected, or because there is insufficient data to include them.

In addition the Food and Drug Administration (FDA) (5) now requires drug manufacturers to develop information to help provide better care and treatment of children, especially data regarding the most frequently used pediatric drugs. Like the NIH, the FDA requires manufacturers to test interventions on pediatric populations unless they show that the research would be unlikely to be useful to the treatment of children in any great numbers, or that including children in studies would be impractical, unsafe, unlikely to be unsuccessful, or otherwise unreasonable.

MORAL ASSUMPTIONS

In discussing what sort of research should be permitted using children as subjects, several assumptions are generally made (5). First, competent adults have responsibilities to ensure that children are not abused, neglected, exploited, or denied access to basic care. Second, different policy solutions regarding children should be evaluated in terms of their adherence to primary moral values. Among the most important of these are beneficence (what will benefit children), social utility (what will fulfill some social good), and justice (what is fair). Policies are judged superior when they are likely to benefit children and ensure that they are not exploited. Good standards fairly promote children's interests, well-being, and opportunities to flourish, as well as help children to develop their potential to become empowered and self-fulfilled. Promoting children's well-being and helping them develop their potential as self-determined individuals demonstrates a concern for their welfare and a willingness to address inequalities of the "natural lottery" (the inequalities caused by nature such as the inherited intelligence or diseases) and "social lottery" (the inequalities caused by social factors such as environment or illness that also affect intelligence and abilities).

To introduce the complexity of the moral issues and disputes about children's research, consider two examples. In each, important values must be ranked in order to assess what ought to be done. These values include the importance of pursuing knowledge, respecting people's rights, protecting minors' well-being, and conducting socially needed research.

Case 1

Dr. D, an endocrinologist, first met Martin two years ago as a happy-go-lucky 6-year-old when his concerned parents brought him to Dr. D's office. Martin was extremely short for his age, in the lower 2 percent for height, with a projected adult height of 5 feet 3 inches. The doctor concluded this is a normal genetic consequence of the boy's parents' short stature. They wanted their son to escape the discrimination they had suffered and sought to get growth hormone therapy for him. They were disappointed to learn they could not afford the $20,000 a year or more for up to 10 years that it would cost for this intervention. They had no insurance to cover such matters. Dr. D told them about a double-blind placebo controlled study at the NIH. In this study all the children receive injections three times a week until they reach their full height, perhaps 10 years later. Half the children get the growth hormone, but half get only a placebo (salt-water) injection, resulting in approximately 1000 injections over the years. The children are brought to NIH each year with expenses paid for, and have physical and psychological examinations, nude photographs, and X rays. Martin's parents enrolled him in the study because he has a 50 percent chance of getting the growth hormone. Two years later they return sad and frustrated because Martin, who hates the injections, is still in the lower 2 percent for height and has undergone no growth spurt. The study is underway to test the theory that the children getting the growth hormone will be taller as adults. While clinicians did not know if final growth is affected, they do know if someone has not had a growth spurt, it is unlikely he is getting the growth hormone. Dr. D suspects if he tells the parents this, they will remove Martin from the study. What should Dr. D say when the parents ask if Martin's lack of a growth spurt means he is probably not getting growth hormone but is in the control group getting injections of salt solution?

This case raises issues about the sometimes divided loyalties of clinicians toward the patients in studies, and about how they balance gaining information by supporting the study and acting in the best interest of the subject. The child, although short, was perfectly happy initially but is now sad perhaps because he thinks he needs treatments for his shortness. Having the opportunity to be in the study has arguably harmed him by making him feel there is something wrong with him that needs to be "fixed" by medicine. The general problem is whether doctors' obligations to their patients change when they are enrolled as subjects in research studies. If the physician's first duty is to act in the patient's best interest, and to be truthful, then perhaps Dr. D. should tell the parents that he does not believe that their son is getting the growth hormone because he has had no growth spurt. He knows that if many subjects withdraw from the study this could cause problems for the investigators and even undermine the study.

Case 2

Dr. M is a pediatrician who's taking care of Barbara, a healthy 12-year-old. Her paternal grandfather died of Huntington's disease after a long and debilitating illness where he became increasingly demented and unable to control himself. Barbara, who was horrified by her grandfather's decline, tells her parents she wants to enroll in a research project in which they test people who may get this inherited condition. She longs to know if she is free of this genetic condition and "like everyone else" or has to plan her life to do things she wants before she reaches middle age. Her father does not want to be tested, although he understands that if his daughter has a genetic predisposition for Huntington's disease, then he does as well. He agrees to her being tested because it is important to her but prefers not to know the test results. While she and her parents agree that she should be tested, her pediatrician, Dr. M, and the genetic counselors strongly object to testing her or recommending her as a research subject for this study. They cite policies barring minors from testing for late-onset genetic conditions (6) and have effectively opposed the investigators who would like to include minors in their study. The clinicians oppose late-onset testing and research with minors because they argue that if children test positively for diseases such as Huntington's, they may feel differently about themselves, suppose they are already sick when they are not, find it harder to get jobs or health insurance, and face other psychosocial risks of harm. Should restrictions be placed on parental authority to assess what is in their child's best interest and to give consent for their child to be tested or participate in studies?

The case raises concerns about parental consent, minor's assent, and genetic testing and screening. Especially difficult to assess are the harms and benefits of testing for diseases that will appear many years later. Who should decide whether testing should be done for these "late-onset" conditions? Are some studies too risky to permit minors to participate, even if the parents consent? Many adults handle information of late-onset disease badly, and this could cause children even more harm. Who should assess the potential benefits and psychosocial harms such as loss of self-esteem, bias, prejudice, and discrimination in employment or insurance? While many policies bar minors from testing for late-onset disease, some parents object to such interference with their authority to decide what is best for the children. Should late-onset disease testing be prohibited on minors, or left to the judgment of parents about whether it is in her best interest to get this testing? In addition, what role should older children have in making decisions? In case 2, the parents and the 12-year-old wish to have the child participate in research, but the clinicians strongly object. If the minor is 17 or on the threshold of full civil rights, her views might carry more weight. As children get older their views gain increasing moral importance.

The transition from dependence to independence is gradual in childhood. Infants should be fully protected, while well-adjusted minors on the verge of majority have a justified claim to understand and express themselves in actions affecting their lives. As children become older and can comprehend more, they should participate in important decisions about their lives, including whether they will serve as research subjects. The problem of whether to include children in decision making becomes acute when children are neither clearly incompetent nor clearly as competent as most adults. For this reason the children's assent is sought as they mature.

As with competent people, the ethical basis for research policy with persons lacking capacity to give informed consent concerns promoting their self-determination, fair treatment, and well-being. The difference is that children cannot be expected to promote their own self-interests. Consequently we have to consider when others, usually parents, will be permitted to consent to have their children participate in research.

CONSENSUS ON INFORMED CONSENT

The right to consent for treatment or research is a civil right, and people must achieve a certain degree of maturity before they gain it. As a matter of administrative convenience an arbitrary age such as 18 is selected as the time when a person is assumed to be sufficiently mature to be granted full civil rights. Until minors reach the age of majority, adults generally have legal authority to make decisions for them. Guardians have this authority because they are usually best suited to protect, identify, and act in children's best interest, and foster development to the point at which children can take responsibility for themselves. Yet guardians do not have an unqualified right to control their children's destiny, and their approval of research projects is insufficient to include minors in research.

A moral, legal, and medical consensus now exists that the competent, free and informed choices of adults must generally be followed in devising treatment plans or enrolling them or their children as research subjects (7,8). Ideally, people who are patients getting standard or experimental care share decision making with physicians as part of an ongoing process, where doctors clarify options and make recommendations. When the subject or patient is a minor, their parental consent is sought. Yet parental views do not carry the same weight when the consent is

given for dependents and the professionals believe that the choice for a minor are abusive or neglectful.

There is a moral, legal, clinical, and regulatory consensus about how to understand the meaning of informed consent. To give informed consent, persons must (*1*) have had disclosed to them all information material to the decision, (*2*) understand or comprehend the information disclosed, (*3*) act or agree voluntarily, (*4*) be competent to act or make the decision, and (*5*) authorize or consent to the procedure, act, or intervention (9). Overcoming this presumptive duty to gain informed consent may be accomplished by demonstrating that the study or treatment constitutes a well-established exception to the duty to gain consent. Incompetent patients cannot give informed consent, so this is a well-established exception. Other morally and legally valid exceptions include medical emergencies, public health emergencies, and patient's waiver of consent (8).

Unlike competent adults who can generally decide what constitutes their own best interest, children need others to protect them. For this reason children are not authorized to decide if they will participate in research. Children are vulnerable subjects. Potential research subjects are called "vulnerable" if they lack capacity to give informed consent or are likely to be coerced or manipulated to participate. Children and those severely impaired by mental illness or retardation are typically regarded as vulnerable because they lack capacity to give informed consent. (Others such as institutionalized subjects, prisoners, the poor, those desperately ill, members of the military, students, hospital staff, and laboratory assistants may be called vulnerable because, while able to give informed consent, they are vulnerable to coercion or manipulation.) The informed participation of vulnerable subjects is problematic, and enrolling them in research protocols often requires special justification or safeguards.

FOUR DIFFERENT POLICY INITIATIVES

There are several important research policy options offering different approaches to balancing what is fair, most protective of incompetent people's well-being, and most respectful of whatever self-determination they have or may develop (5,10). These four policies represent different regulative ideals because they balance these primary values differently and because they offer different authority principles (stating who decides) and guidance principles (substantive directions about how decisions should be made). The remaining discussion will be focused on these options.

Let Surrogates Decide Using Their Own Values

One policy allows guardians to give consent for children under their care as they do for themselves, without any additional restrictions. If they can consent for themselves to take part in a research program testing for late-onset genetic diseases (see case 2 in the discussion above), then, according to this policy, they should be able to authorize this for their children. Guardians are typically knowledgeable and concerned, defenders argue, and society should intrude as little as possible into the privacy of family life. Since guardians have the authority to make such life-shaping choices as to their child's religion and schooling, then, according to this view, guardians should also determine whether or not their child should participate in research.

Early in the twentieth century, it was assumed that guardians had almost unqualified rights to control their children's futures until adulthood unless they were emancipated through military service, marriage, or independence (11). One reason for this policy was that minors were judged incompetent to make rational decisions, so complete adult direction was regarded as necessary to help them develop their potential. Parents, usually the father, had this right. No consideration in the courts was given to minors' emergent rationality, hopes, or plans. Their views were not material to a decision, unless they were emancipated. It was considered unnecessary and irrelevant to find out, for example, how the child wanted to be treated (12).

In addition, minors were thought to virtually belong to their guardians until they reached the age of 21 (11). When there were disagreements about the treatment or placement of a child and the dispute went to the courts, the courts' role would be to decide who "owned" the child, since that person had authority to make decisions. It was a great tragedy if the guardian was abusive or neglectful, and as a result a child died or was maimed for life, but the courts took the view that the decision maker, usually the parents, had a right to decide (12). The attitude was that competent persons had a right to make even an unfortunate decision for themselves and their wards, and there was little to be done beyond trying to persuade them to act otherwise.

Critics began to challenge this position in the late nineteenth century, when a major legal and social shift occurred and as children's interests were increasingly considered when choices were made about them. They were granted entitlements to be protected from abusive or neglect as well as liberties independent of their parents. More recently this included setting research policy restricting when children could participate in studies, even if parents wanted them to be in the research. In addition the child's assent was sought as a precondition of research participation.

Guardians now have authority insofar as they promote the well-being of those under their care, and prevent, remove or minimize harms to them. Volunteering to put oneself in harm's way to gain knowledge may be morally admirable. Volunteering to put another in harm's way is not admirable, and may violate the guardian's protective role. Consequently this first policy option — to let parents decide as they would for themselves — is rarely defended today. Its echoes, however, may be heard in case 2, discussed above, where parents resent state restrictions on research with minors.

Nuremberg: Require Consent for All Research

Another policy requires consent from competent persons to enroll them as research subjects. This view is found in the first written international research code, the Nuremberg code (13). Composed at the end of World War II, the

Nuremberg code stands as a international response to the horrible, involuntary medical studies done by the Nazis in which many people, including children, were killed or permanently maimed. The Nuremberg code states: "The voluntary consent of the human subject is absolutely essential." It goes on to define consent in a way that has become fairly standard, as requiring legal capacity, free choice, and understanding of "the nature, duration, and purpose of the experiment; the methods and means by which it is conducted; all inconveniences and hazards reasonably to be expected; any effects upon his health or person which may possibly come from participation in the research" (13, pp. 181–182).

The Nuremberg code was a response to invasive, harmful and sometimes-deadly studies done on unwilling subjects. If taken as a general code for research (and it may not have been intended as such), this policy excludes subjects who lack capacity to give informed consent. The reason sometimes given for adopting this policy is that it violates people's rights of self-determination to use them in research if they cannot give consent for themselves to participate.

One difficulty with this policy is that research for children and other people who cannot give consent will stop and, consequently, so will advances in their care. Children have unique medical problems, and the results from studies with adult subjects may be inapplicable to children. Adults cannot be used to test the safety and efficacy of drugs for childhood schizophrenia, or for premature infants with respiratory distress and infections. To test the safety and efficacy of many standard, innovative, or investigational treatments for distinctive groups, some members of the groups have to be subjects.

The initial justification for excluding persons who lack capacity to give informed consent for research is to honor their rights and protect their welfare. This restriction, however, may prohibit studies that benefit children. Some are denied access to projects or investigational therapies that could help them as well. All children are denied good information about their conditions because there are no results applicable to pediatric populations. Thus it may not promote their welfare, individually or collectively, to forbid their participation. In addition participation in controlled research does not always violate the rights of persons who lack capacity to give informed consent. By excluding persons who are unable to give consent from relatively safe studies that advance knowledge or provide therapeutic benefit, children's needs and opportunities are not given full consideration. This policy would exclude all children from being enrolled in any research, however low the risk of harm. Naturally it would disallow studies outlined in each of the two cases mentioned earlier. More troubling is that it would not permit even safe therapeutic studies. A dying child could not, for example, obtain an experimental drug that might save her life.

Helsinki: Only Therapeutic Research Without Subject's Consent

A third policy option holds that persons who lack the capacity to give informed consent may be enrolled only in therapeutic studies. This view is represented in the next major international code for research to follow the Nuremberg code, the World Medical Association's Declaration of Helsinki. It states: "In the case of legal incompetence, informed consent should be obtained from the legal guardian in accordance with national legislation" (14, p. I, 11). The incompetent person must agree as well, when able to do so. It allows research with incompetent people, but "...only to the extent that medical research is justified by its potential diagnostic or therapeutic value for the patient" (14, p. II, 6). If, however, the medical research is nontherapeutic, then: "The subjects should be volunteers — either healthy persons or patients for whom the experimental design is not related to the patient's illness" (14, p. III, 2). People who cannot volunteer therefore cannot be subjects in nontherapeutic studies.

This policy option distinguishes clinical or therapeutic research (studies seeking generalizable knowledge but intending to provide medically acceptable therapy for the individual) from nontherapeutic biomedical research (studies seeking generalizable knowledge but not intended as therapy to benefit the individual directly). Therapy is designed to benefit the person, so drawing the line at therapeutic research for people who lack the capacity to give informed consent might seem a good solution.

This approach has difficulties. First, classifying research as either therapeutic or nontherapeutic may be misleading and arbitrary. Studies often have features that are not routine therapy such as extra procedures, tests, or visits to the doctor. Second, important medical research may hold out benefits other than therapy to subjects. Small children who are not patients may enjoy participating in studies where they are asked to do such things as stack similar blocks or identify animals from sounds that they made. Older children might, for example, enjoy an outing to a research facility or find the study interesting in itself.

The Helsinki policy, however, forbids nontherapeutic studies, even though some are important and safe. For example, it is very important to learn if children are developmentally delayed. But to determine this, investigators need to establish a normal range against which to compare results. This standard is obtained by collecting data from many children about their height, weight, vision, hearing, and so on. These children are not sick, so the studies cannot be regarded as therapeutic, and therefore on this policy cannot be permitted even if they are extremely important and safe.

Third, the distinction between therapeutic and nontherapeutic studies focuses on direct benefits to the individual. Yet nontherapeutic studies may indirectly benefit persons unable to give consent. When these studies obtain valuable information about them as a group, and involve little or no risk, it seems problematic to prohibit them. As we indicated, standards of typical growth and development help all children yet require testing on large numbers of normal children to generate data that help distinguish developmental delays or impairments in children from normal growth and development. If children ride tricycles, it is not research; if investigators observe when and how they do it to make generalizations about children, it is research that may not be burdensome to the child.

The Helsinki policy is controversial then because it excludes many important, low-risk studies. It also excludes, of course, the research studies suggested in two cases discussed above. In case 1 the child is normal, not sick, so this is not a therapeutic study; in case 2 no therapy is planned.

The initial justification for excluding persons who lack the capacity to give informed consent from nontherapeutic research was to honor their rights and protect their welfare. Safe, nontherapeutic research, however, seems neither unfair, nor a violation of the rights or welfare of people who lack the capacity to give consent. Failing to do safe, but important, studies might be unfair and violate their rights and welfare, since it fails to consider all their needs. Thus, when nontherapeutic studies are not potentially harmful or inconvenient and when the subjects want to participate, it is not clear that their rights and welfare are always violated or that they are treated unfairly.

U.S. Rules, CIOMS and Other Guidelines: Assess the Ratio of Likely Harms to Benefits

A fourth approach allows research with incompetent persons if it holds out benefit or does not place them at unwarranted risk of harm, discomfort, or inconvenience. To try to balance the social utility of research with respect and protection of incompetent people, this option stipulates that the greater the risk, the more rigorous and elaborate are the procedural protection and consent requirements. The U.S. federal government (15, §46.404–7) reflects this policy option in its codes for research, involving children as well as adults. The Council of International Organizations of Medical Science (CIOMS) (16) and health policy groups in other countries have also adopted this general approach (10).

There are advantages to focusing directly on the benefits and harms of procedures or interventions in distinguishing the permissibility of research. It avoids the three problems discussed earlier concerning other options. It offers special protections for vulnerable subjects but permits some research with children as subjects. It also allows that there may be benefits other than therapy to consider in assessing potential harms and benefits to children participating in studies. It does not make the sometimes troubling distinction between therapeutic and nontherapeutic research in order to determine which studies are acceptable for children as recommended in the Helsinki approach. As we saw above, classifying studies as therapeutic or beneficial can be misleading or arbitrary. It can be arbitrary because studies have potential harms unrelated to this classification such as extra tests or visits. It can also be misleading if people assume therapeutic studies are always safe or beneficial. Calling something "therapeutic" may hide potential harms, disadvantages, or other nonbeneficial features, creating an inappropriate bias for participation that would be revealed by careful risk assessment.

In contrast, the fourth policy option focuses on risk assessment to balance the social utility of encouraging studies with the protection of people's rights and well-being. It stipulates that, whenever possible, the incompetent persons should give their assent or affirmative agreement to participate.

In using the likely harms to benefit calculation, the U.S. regulations specify four categories of research with children (15). As risks increase, the regulations require increasingly more rigorous documentation of appropriate parental consent, children's assent, direct benefits to each child, and benefits to other children with similar conditions. Local Institutional Review Boards (IRBs) can approve studies only in the first three categories.

The *first* category permits research with no greater than a minimal risk provided that the study makes adequate provisions for consent from at least one parent and the child's assent. The *second* category of research permits approval of studies with greater than a minimal risk if the risk is justified by the anticipated benefit to each subject, the risks in relation to these benefits are at least as favorable to each subject as available alternatives, and provisions are made for consent from one parent and the child's assent. The *third* category permits research with a minor increase over minimal risk that holds out no prospect of direct benefit to the individual subject where the study is like the child's actual or expected medical, dental, psychological, or educational situation, is likely to result in very important information about the child's disorder or condition, and provisions are made for parental consent, typically consent from both parents, and the child's assent. Investigators using this category might be permitted to conduct, for example, additional lumbar punctures on children with leukemia to help study their disease. Consent from both parents is required if practicable.

Research that cannot be approved under the first three categories might be approved under a *fourth* category if it presents a reasonable opportunity to understand, prevent, or alleviate a serious problem affecting the health or welfare of children, and the study is approved by the secretary of the Department of Health and Human Services (HHS) after consulting with a panel of experts about the study's value and ethics and determining that adequate provisions have been made for the parental consent, typically consent from both parents, and the child's assent. Using this category, investigators might, for example, gain approval to conduct studies on normal, healthy children intended to prevent expression of late-onset genetic diseases. In the United States, IRBs cannot approve studies that have more than a minor increase over a minimal risk which do not hold out benefit for the children. As Table 1 shows, IRBs must seek approval from the federal government to conduct them.

Unfortunately, this fourth policy leaves key terms poorly defined and consequently allows broad interpretations about what risks of harm are warranted. The pivotal concepts of "minimal risk" and "a minor increase over a minimal risk" are problematic. The regulations state: " 'Minimal risks' means that the risks anticipated in the proposed research are not greater, considering probability and magnitude, than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests." (US 45 CFR46 102i) (15).

Table 1. U.S. Regulations Requirements for Research with Children

Research Category	Risk-of-Harm Category	Requirement for Children's Participation
I (46.404)	Minimal risk	1. IRB approval 2. Child's assent, and 3. Informed consent from at least one parent or guardian
II (46.405)	More than minimal risk, with prospects of direct benefit for each subject	1. IRB approval 2. Child's assent 3. Informed consent from at least one parent or guardian 4. Risk justified by anticipated benefit to each subject, and 5. Anticipated benefits to each subject at least as favorable as that presented by available alternate approaches
III (46.406)	More than minimal risk, without prospect of direct benefits to each subject	1. IRB approval 2. Child's assent 3. Informed consent from both parents or guardians 4. Risk represents a minor increase over minimal risk 5. Likely to yield generalizable knowledge about child's disorder or condition that is important for the understanding or ameliorating of disorder or condition, and 6. Intervention or procedure presents experiences to child that are reasonably commensurate with those in child's actual or expected medical, dental, psychological, social, or educational situations
IV (46.407)	Research not otherwise approvable	1. IRB approval 2. Child's assent 3. Informed consent from both parents or guardians 4. IRB finds that research presents a reasonable opportunity to further the understanding, prevention, or alleviation of a serious problem affecting the health or welfare of children, and 5. Study gains the approval of the Secretary of HHS after consultation with a panel of experts in pertinent fields and following opportunity for public review and comment

Source: CRF **45**, 46, 404–7 (1991).

The first part of the definition is vague because daily risks include dangers from riding in cars, flying in airplanes, and living in a world filled with nuclear and conventional weapons. How well do we know the nature, probability, and magnitude of these "everyday" risks and why should they serve as a baseline to estimate minimal research risk? Depending on where one lives, daily risks can be life-threatening. It seems considerably easier to determine that having a 4-year-old stack blocks is a minimal risk study than to determine the nature and magnitude of whatever risks people normally encounter (10).

The second part of the definition seems to set a standard for physical interventions that have a minimal risk. The test is whether the activity is like that of a routine examination. Accordingly, IRB or Research Ethics Committee (REC) members may not approve as minimal risk research with such procedures as X radiography, bronchoscopy, spinal taps, or cardiac catheterization because they are not part of routine examinations. IRBs and RECs, however, can approve studies which have a minor increase over minimal risk, and some of these procedures have been approved as having only a minor increase over a minimal risk. Without standards for risk assessment, how effective are these guidelines? Not surprisingly there are considerable differences of opinion about whether procedures such as lumbar punctures, multiple placebo injections, arterial punctures, and gastric and intestinal intubations are regarded as risky (10).

Moreover this definition of "minimal risk" and subsequent judgments about what constitutes a minor increase over minimal risk offers no guidance about how to assess psychosocial risks. These include invasion of privacy, breach of confidentiality, labeling, and stigmatization. In "routine" visits doctors and nurses "ordinarily encounter" discussions of family abuse, sexual preference, and diagnoses that could affecting people's abilities to get jobs or insurance. As we saw in our two examples, these can be grave risks.

Freedman, Fuks, and Weijer (17) respond that identifying everyday risks we all encounter is not difficult, although it may be hard to quantify them. Yet they acknowledge the potential problem of using different standards for subject inclusion because of intercultural variation regarding "everyday risks." Their solution is to promote "intercultural ethics" where the norms of *all* the cultures participating would have to be honored to conduct any crosscultural studies. Clearly, however, this proposed solution has not been adopted. This is obvious from the multinational AIDS research conducted in developing countries but funded by nations that would not be able to conduct them in their own countries because they violate research policy. Nations from North America and Europe,

especially the United States, supported this research and tried to justify it in terms of local standards and conditions. There is an ongoing debate over permitting greater consideration of local standards or conditions (18,19).

Second, their assumption that a consensus exists regarding what paradigms to use in assessing the crucial upper levels of justifiable risks of harm also seems unwarranted. There are considerable differences among pediatric experts, in both treatment and research settings, about how to assess the risk of such procedures as lumbar punctures, multiple placebo injections, venipunctures, arterial punctures, and gastric and intestinal intubations (20). Investigators and others have concluded that better standards of risk assessment in children's research need to be formulated (21,22,10). More recently this lack of consensus and regulatory guidance within the research community was extensively discussed in testimony before National Bioethics Advisory Commission (NBAC) recorded in "Regulatory Understanding of Minimal Risk" where members of NBAC agreed this was a problem (22).

The fourth approach represented by the U.S. research rules, CIOMS regulations, and those of other countries seems to many to balance the need to protect the rights and welfare of people who lack the capacity to give informed consent with the need to encourage research. It may, however, be popular because it is so vague that it permits different understandings of what constitutes acceptable risks of harm to them. Each of the other three policy options, although flawed, offer clear guidance about what is permissible. This is not the case with the fourth standard, and debates arise about how to use this policy in certain cases.

DEBATES ABOUT BALANCING HARMS AND BENEFITS

The general acceptance of the fourth policy alternative, with its attendant ambiguities about what constitutes a warranted risk of harm, has engendered disputes about how to understand and balance harms and benefits in permitting research involving children. Recall that in case 1, the parents focus on what is best for their child, and they balance their views about what helps him against the harms that will fall to him if he is in the study but does not get the growth hormone. It is not unreasonable for parents to want to improve their child's life, using the values they think most important. Martin's parents are willing to take risks to have him avoid the discrimination they faced because they are short, but unwilling to have him suffer without a chance of some advantages. If they decide there are insufficient benefits to him from being in the study, then the harms from the hundreds of placebo injections, inconvenience, medicalizing of his shortness, and so on, tip the balance and, using their values, justify taking him out of the study.

The investigators, however, might consider a different set of values, harms, and benefits. They might want to consider not just the harms and benefits to a few people in the study but to all people in the future who might take this drug. Investigators might want to emphasize the virtues of keeping one's bargains of staying in the study until the end. Investigators sometimes stress the long-term benefits of their research in possibly developing new and better therapies in the future. On the other hand, if we speculate on possible long-term benefits from a not yet developed therapy, then we probably should also consider and balance them against the possible long-term harms from it as well. This calculation becomes further complicated because it is hard to factor in special interests, biases, prejudices, and wishful thinking. In addition, if open information makes those in one arm of a study want to withdraw, one must question if there is a genuine balance of harms and benefits in the other arms of the study.

The research rules give little guidance about what goals to use to determine harms and benefits and how to balance them. For example, in case 1 it is crucial to determine the benefits of a good study against the harms to the child, including getting so many placebo injections, and the long- and short-term harms to the doctor–patient relationship. How do we establish the different goals and assessments about harms and benefits? One might conduct an opinion poll about this. If so, what groups should be surveyed? Should they be among parents, doctors, short parents, investigators, or the general public? Clearly we can influence the results of the opinion poll by the group we pick to survey.

Different goals and assessments of harms and benefits are also apparent in case 2, where the parents mean well and want to help their child find out if she has a late-onset disease, hoping, of course, that she will find out that she does not. Unlike case 1, where parents can withdraw their child from a research study, in case 2 they want to get their child into a study if that is the only way that they can get her tested for a late-onset disease. There are different assessments of harms and benefits, with the investigators, parents, and the potential subject minimizing the importance of the predicted harms from testing for Huntington's disease and the pediatrician and genetic counselor refusing to accept their assessment.

These debates arise because reasonable and informed people of good will have different goals and views about what constitutes a harm and a benefit, and about how to rank or balance them. In addition they have different expectations of the future. In what follows, two areas of debate about goals, benefits and harms are discussed: (*1*) distinguishing treatment, prevention, and enhancement, and (*2*) research to develop germ-line gene therapies.

Distinguishing Treatment, Prevention, and Enhancement

New genetic technologies create problems about how to use them. One proposal is to distinguish treatment, prevention, and enhancement and permit use of these new techniques only to prevent or treat medical conditions like sickle cell disease; but not to enhance normal people's height (as in case 1) or their beauty, strength, and intelligence, and the like. On this view we should simply prohibit development of new technologies for the purposes of enhancement. Yet this viewpoint is questionable. In some sense to prevent or treat a disease already enhances one's life, making it difficult to say how these distinctions should be determined. For example, putting fluoride in

the water enhances people's lives by reducing dental pathology. Treating infectious diseases also enhances lives by making people better and preventing others from getting sick. Since such enhancements are not controversial, it appears there are no principled objection to enhancement understood as the prevention or treatment of disease. On the other hand, making normal children taller, as in case 1, is more controversial because shortness is not a disease like dental pathologies or infections. Martin's parents understand the projections are that he will be only 5 feet 3 inches, and while possibly agreeing that shortness is not a disease, they see this is a disadvantage for a man in our society.

Some enhancements are even more controversial than those aiming to bring someone into normal range of height so he will escape discrimination, as the parents in case 1 desired. For example, some enhancements seek to give people special advantages. For example, parents might wish to improve their son's opportunities to become a professional basketball player and are disappointed to learn their child's projected height is only 6 feet 3 inches. They want him to get growth hormone so he can have a better chance of becoming the next Michael Jordan. Many people might object to this use of technologies on several grounds. First, not everyone has access to this technology, so it is unfair for some to get this advantage. Second, if everyone had access to the technology, no one would have the sought-for advantage. It is not possible to enhance the height of all short, normal children because no matter what resources are used, there would still be children in the lowest 2 percent for height. Third, it exposes the child to some risks for controversial or questionable motives. Finally, it is a poor use of an expensive technology when there is so much need in the world. The debate continues over what limits should be placed on parental authority, investigators, and the use of state funds to develop, test, and use new technologies for enhancement, especially where they involve possible harms to the child.

In trying to resolve the debate by prohibiting these technologies for enhancement but permitting them for prevention and treatment, consider the following example. Two boys have a projected height of 5 feet 3 inches. One has short parents, while the other has a tumor. Should one be refused growth hormone therapy because it would constitute enhancement, while the other gets it because it is treatment? Some would argue they both should get the therapy so they both can reach normal height, gain a normal opportunity range, and escape the social disadvantages from prejudice (7). Some might object, however, to either of them getting it, not on the ground that this may be enhancement rather than treatment but because both are in the normal opportunity range for height; they might agree if either child's projected height were far less.

Trying to settle the debate in terms of someone's normal opportunity range raises difficult problems about what constitutes a normal range, as well as the limits of social obligations to find the money to get children to some norm (23). Some might object to giving children a very costly drug using state money when they have a projected heights of 5 feet 3 inches, since being 5 feet 3 inches is not, in their view, abnormal. That is, most options are still open to people with a projected height of 5 feet 3 inches. Many women get along perfectly well at this height, finding no great hardship in being 5 feet 3 inches. Moreover, if the problem is social prejudice, we should fix that, not the child, especially given the great needs to use state resources for basic health care for children. Little girls with Turner syndrome have extremely short stature if they do not get growth hormone, so public funds are routinely used in treating their short stature with growth hormone. Arguably, if any child were projected to have such short statue as these girls, they too should have the growth hormone. But some uses of these new technologies seem like cheating. The use of anabolic steroids, for example, to enhance athletic performance is prohibited as seeking unfair advantage over competitors.

Ultimately, resolution of these debates may involve factors other than finding precise delineations of enhancement, prevention, and treatment. Not only is it very difficult to draw clear lines between them but, in the end, some believe it may not be very important (7). What may be more important is whether it permits someone to gain a normal opportunity range, without undue risk of harm, at a cost society can afford (7,23). The issue may turn on how we rank values and on the balancing of potential harms, benefits, and costs.

Research to Develop Germ-Line Gene Therapies

Human germ-line therapies are controversial because they might modify traits that offspring would inherit from parents. They can be beneficial for all, for example, by curing a condition such as sickle cell disease for patients and their descendants. Or, they might harm the person and subsequent generations. How do we assess the goals and potential harms and benefits of research that could harm not only the individual but his or her offspring? The debate also raises issues about the proper moral restraints on science. Eric Juengst (24) summarizes arguments regarding the goals and potential benefits and harms in developing germ-line gene therapies. The first is its potential to be medically useful. Research into germ-line gene therapy will develop ever more therapies and techniques to cure conditions, rather than merely treat the symptoms. Second, developing these techniques may be the only effective way to treat some of these conditions when palliative or symptomatic therapeutic interventions are unavailable. Third, such measures may be efficient in that by developing germ-line gene therapies, one would prevent the transmission of diseases between generations and therefore avoid treatments for many others in the future. Finally, many resist limiting these therapeutic techniques on the ground that it is an unwarranted restriction of scientific freedom. The federal regulations delineate the limits of research involving human subjects, and if research falls within these regulations, it is unreasonable to limit free inquiry.

On the other hand, there are arguments opposing development of germ-line gene therapy techniques, especially with children as subjects. The first is the unknown, unexplored, and unpredictable risks to the persons and their offspring. Given the uncertainty of these risks of harm, there is no way to determine exactly

whether the harm–benefit ratio would be justified or not. A second issue concerns the problem of whether we are developing techniques for therapy or for enhancement. As we discussed earlier, it is difficult to distinguish between medical and nonmedical uses of these techniques. Third, some object that since these techniques could harm future generations as they try to perfect the development of these techniques, they would be putting future individuals, who gave no consent, at risk. Fourth, there is the issue of the allocation of resources. It would be costly to develop these new germ-line gene therapy techniques, and some argue that we have higher social priorities. Finally, some are concerned that we would change the germ line for all time by these techniques, altering inheritance intentionally and perhaps capriciously.

To conclude, research involving children can foster different and important values including benefit for the children who are subjects, medical advances, prudent use of public resources, health care planning, and public health. With more information, better diagnoses, treatments, and prognoses, as well as informed consent, are possible. In some cases, however, these values can conflict, posing difficult choices on how to acknowledge people's rights, protect their interests, gain good information, and do what is useful socially. Different research policies offer different solutions on how to balance these values, and these policies have their own different strengths and weaknesses. Solutions were judged superior when they fairly promote children's well-being and opportunities to develop and flourish.

BIBLIOGRAPHY

1. L.M. Kopelman, in J.F. Monagle and D.C. Thomasma, eds., *Health Care Ethics: Critical Issues*, Aspen Publishing, Gaithersburg, MD, 1994, pp. 199–209.
2. American Academy of Pediatrics, Committee on Drugs, *Pediatrics* **95**(2:2), 286–294 (1995).
3. U.S. National Institutes of Health, *Policy and Guidelines on the Inclusion of Children as Participants in Research Involving Human Subjects*, March 6, 1998. Available at: *http://grants.nih.gov,grants/guide.notice-files-98-024.html*.
4. U.S. Department of Health and Human Services Food and Drug Administration, *Regulations Requiring Manufacturers to Assess the Safety and Effectiveness of New Drugs and Biological Products in Pediatric Patients*, [FR Doc. 98-31902] OC 98142, Docket No. 97N-0165, pp. 66631–66672. Available at: *http://www.verity.fda.gov/search97cgi/s97_cgi.exe?action=View&VdkVgwKey*.
5. A.E. Buchanan and D.W. Brock, *Deciding for Others: The Ethics of Surrogate Decision Making*, Cambridge University Press, Cambridge, UK, 1989.
6. E.W. Clayton et al., *J. Am. Med. Assoc.* **274**(22), 1786–1792 (1995).
7. D.W. Brock, *Bioethics* **9**(3, 4), 269–275 (1995).
8. R.R. Faden and T.L. Beauchamp, *History and Theory of Informed Consent*, Oxford University Press, New York, 1986.
9. T.L. Beauchamp et al., *J. Clin. Epidemiol.* **44S**(1), 151S–169S (1991).
10. L.M. Kopelman, *J. Med. Philos.* (2000), forthcoming.
11. L.S. McGough, in *Encyclopedia of Bioethics*, rev. ed., vol. 1, Simon & Schuster, New York, 1995, pp. 371–378.
12. A. Holder, *Legal Issues in Pediatrics and Adolescent Medicine*, 2nd ed., Yale University Press, New Haven, CT, 1985.
13. The Nuremberg Code, Nuremberg Military Tribunals, 1949, pp. 181–182.
14. World Medical Association, Declaration of Helsinki [1964] 1989, *Recommendations Guiding Medical Doctors and Biomedical Research Involving Human Subjects*, Adopted by the 18th World Medical Assembly, Helsinki, Finland, and amended in 1975, 1983, and 1989.
15. U.S. Federal Regulations, 45 CFR §46 (1991).
16. Council for International Organizations of Medical Science (CIOMS), *International Ethical Guidelines for Biomedical Research Involving Human Subjects*, Geneva, Switzerland, 1993.
17. B. Freedman, A. Fuks, and C. Weijer, *Hastings Center Rep.* **23**(2), 13–19 (1993).
18. R.J. Levine, *N. Engl. J. Med.* **341**(7), 531–534 (1999).
19. T.A. Brennan, *N. Engl. J. Med.* **341**(7), 527–531 (1999).
20. J. Janofsky and B. Starfield, *J. Pediatrics* **98**, 842–846 (1981).
21. A.D. Lascari, *J. Pediatrics* **98**, 759, 760 (1981).
22. National Bioethics Advisory Board, Testimony before the Human Subject Subcommittee of the National Bioethics Advisory Commission Meetings, January 8, 1998, in Arlington, VA. Availability at: *http://bioethics.gov/transcripts/index#jan98*.
23. N. Daniels, *Just Health Care*, Cambridge University Press, Cambridge, UK, 1985.
24. E.T. Juengst, *J. Med. Philos.* **16**(6), 587–592 (1991).

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, ETHICS, COMPENSATION OF SUBJECTS FOR INJURY

TERRENCE F. ACKERMAN
University of Tennessee
Memphis, Tennessee

OUTLINE

INTRODUCTION

Participation in clinical research sometimes results in pain, physical disability, or even death for human subjects. For several decades scholarly commentators, medical researchers, and government advisory committees have debated the advisability of establishing a national program to provide compensation for injured research subjects. The

substance of this debate has focused on several critical questions. What is the basis and extent of society's moral obligation to compensate subjects for research injuries? Does the incidence and magnitude of research injuries justify the establishment of a national compensation program? What are the essential components of a morally acceptable and practicably workable compensation plan? Disagreements about our moral obligations, the paucity of empirical data about research injuries, difficulties in devising a workable system, as well as the press of other public priorities, have thwarted any clear social consensus on the merits of a compensation program. Nevertheless, the unavoidable fact that some persons are injured in medical research for the benefit of society has sustained the policy controversy.

SOME INITIAL DISTINCTIONS

In exploring the issues regarding compensation, several distinctions are crucial to the analysis. First, it is important to distinguish research injuries for which subjects possess legal means for securing redress and those for which legal recourse is not available (1,2). Injuries may occur as a result of researcher negligence when subjects have not been adequately informed of the risks involved or when due care has not been exercised in protecting them from risks. In these instances injured persons may bring civil suits for damages against investigators. In other limited cases, such as workers' compensation for federal employees, persons injured in research may use existing administrative mechanisms for securing compensation. By contrast, most research injuries occur without any fault on the part of researchers. There are several reasons. Most interventions employed in clinical research, such as chemotherapy or tissue biopsy, possess a definite risk of harm even when competently performed. Oftentimes the use of new drugs or medical devices results in harm that could not be anticipated based on prior knowledge. Moreover, many research subjects have diseases or disabilities that result in a heightened susceptibility to the harmful effects of medical interventions. Yet very few research institutions have formal compensation programs. As a result the debate about compensation has focused on research injuries that occur without researcher negligence and for which no redress is legally available.

Important to the contours of the debate is a second distinction between therapeutic and nontherapeutic research procedures (3). Therapeutic research procedures are performed in order to benefit the individual subject, as well as to produce generalizable knowledge. For example, administration of a new drug in a randomized clinical trial comparing it to standard treatment is a therapeutic research procedure. By contrast, when a medical intervention is performed solely to produce generalizable knowledge, without being intended to benefit the subject, it is a nontherapeutic research procedure. For example, serial venipunctures performed to profile the pharmacokinetics of a new drug in healthy volunteers constitute nontherapeutic research procedures. This distinction is crucial to disputes about the limits of the moral obligation to compensate injuries, the extent of the need for a program, and the practicality of any scheme for identifying compensable injuries.

A third distinction concerns the reasons for which subjects participate in clinical research. Some may participate primarily to promote their own interests. Other subjects may participate in research studies primarily to contribute to the welfare of society. This distinction cuts across the prior one between therapeutic and nontherapeutic research procedures. Most subjects enter trials of therapeutic procedures to secure personal benefits, especially the opportunity to receive a new treatment that may be more efficacious or safe than existing therapies. But participation in a clinical trial may offer no special advantage, as when investigators are comparing the efficacy and safety of two standard treatments. On the other hand, subjects who participate in studies involving only nontherapeutic research procedures anticipate no medical benefit. Yet in many cases they may receive an attractive payment for their participation. An important point of contention is whether personal sacrifice for the common good is a necessary condition for a valid claim against society for compensation of injuries.

A final distinction concerns the difference between medical research that is undertaken at the behest of society and research not similarly sanctioned. In the former category there is research conducted by employees of the federal government, such as physician-investigators employed by the National Institutes of Health (NIH). In addition much medical research in the United States is supported through grants from the federal government to investigators working in public and private academic institutions. There is also considerable medical research conducted by private medical companies required by regulations of the Food and Drug Administration (FDA) for the development of new drugs and medical devices. Whether medical research is federally conducted, supported, or regulated, it is commissioned through the official agencies of society. By contrast, considerable medical research is funded through private sources, particularly charitable foundations. While such research contributes to the common good, it is not officially sanctioned. An important issue is whether an obligation to compensate injuries should be limited to subjects in research sanctioned by the official agencies of society or should extend to all research that contributes to the common welfare.

HISTORY OF THE COMPENSATION DEBATE

Between 1945 and 1965 annual expenditures of the NIH increased exponentially from $701,800 to $436,000,000. This increase reflected an exploding confidence of the American public that medical research could yield enormous advances in the prevention and treatment of disease. By 1960 governmental support for clinical research sustained a vast infrastructure of research programs at medical colleges in the United States.

Despite the proliferation of medical research employing human subjects, there was only sporadic discussion of issues related to the rights and welfare of human subjects. Although the ethical guidelines for the conduct of human

research formulated at Nuremberg were considered relevant to the American enterprise, it was generally believed that the moral conscientiousness of individual investigators was sufficient to assure that subjects were adequately protected. The absence of public reports of research injuries reinforced the informal approach.

This public and professional confidence began to unravel in the 1960s with a series of revelations about studies in which persons were uninformed that they were subjects and/or exposed to unjustifiable risks of harm (4). Public concern was awakened after severely deformed infants were born to mothers who took the sedative thalidomide during pregnancy without being aware of its investigational status. This tragedy helped solidify support for the 1962 Harris-Kefauver Amendments to the Food, Drug and Cosmetic Act, which included the requirement that informed consent be secured from prospective subjects for the testing of investigational drugs. In 1964 it was revealed that two New York physicians, while studying cancer immunology, had injected live cancer cells under the skin of elderly, debilitated patients without their knowledge. This was followed in 1966 by the publication, in the *New England Journal of Medicine*, of Henry Beecher's paper, "Ethics in Clinical Research," which described 22 studies published in leading medical journals that involved the exposure of subjects to unjustifiable risks of harm (5). Finally, in 1972, front-page newspaper articles detailed the infamous Tuskegee Syphilis Study in which 400 poor, black males in Alabama suffering from syphilis had been left untreated for 40 years to study the natural history of the disease (6). Medical research was now acknowledged to involve palpable human costs.

These developments fostered a growing recognition that adequate protection for the rights and welfare of human subjects required the creation of formal social controls over clinical research. Beginning in 1966, the Public Health Service (PHS) established protective mechanisms involving two main components (7). First, human research activities conducted or supported by PHS were required to undergo prior review by committees established at each institution engaging in clinical research. Second, these committees were to approve research only if adequate provision was made for securing the informed consent of subjects or their legally authorized representatives and only if the risks were justified by the anticipated benefits to subjects and others. These essential elements of social control were refined over the next decade and a half, assuming much of their present form by 1981 with the implementation of the recommendations of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (hereafter, the National Commission) (8).

At the same time, a third mechanism of social control was being explored in the legal and medical literature: no-fault compensation for injured research subjects. Some early commentators viewed compensation as a means of fulfilling a societal obligation to those injured in contributing to the common good (9). Initially, however, compensation was often conceptualized as a mechanism for controlling risks (10,11). Compensation programs funded by research institutions would encourage more careful scrutiny of the risk/benefit profile of proposed research studies and would necessitate closer surveillance of the safety of subjects during their participation. Moreover, if injured subjects were fully compensated for medical expenses, lost wages, and related expenses, society would be forced to acknowledge the full costs of medical research.

Beginning in the 1970s a series of government advisory committees formally examined the compensation issue. With the strengthening of other mechanisms for controlling research risks, these groups began to focus on the moral obligation of society to compensate research injuries. For example, the Tuskegee Syphilis Study Ad Hoc Advisory Panel in 1973 endorsed the concept of compensation as follows:

> No policy for the compensation of research subjects ... has been formulated, despite the fact that no matter how careful investigators may be, unavoidable injury to a few is the price society must pay for the privilege of engaging in research which ultimately benefits the many. Remitting injured subjects to the uncertainties of the law court is not a solution (12).

Based on this rationale, the committee recommended establishment of a no-fault compensation program for injured research subjects. Similarly NIH accepted this rationale in three proposals submitted to the Secretary of Health, Education, and Welfare (HEW) in the early 1970s for implementation of a no-fault compensation program administered by the government for injuries sustained in research supported by federal funds. However, acceptance of these proposals was considered premature. Their fiscal implications had not been assessed, alternatives to a federally administered program had not been explored, and the need had not been clearly established (13).

These shortcomings in prior proposals prompted the Secretary of HEW in 1974 to create a task force to examine in greater detail "whether and how to compensate subjects injured in the course of research." The HEW Secretary's Task Force on the Compensation of Injured Research Subjects (hereafter, the HEW Secretary's Task Force) conducted an 18-month study in which it elicited papers on the ethical and legal aspects of compensation, commissioned an empirical study on the extent of research injuries, and consulted with officials from the insurance industry regarding the feasibility of a compensation program underwritten by private companies. In its final report, the Task Force strongly endorsed the proposition that society has a moral obligation to compensate injured research subjects (13). Moreover it maintained that this obligation extended to subjects of both therapeutic and nontherapeutic research procedures. While the Task Force report recommended that subjects in research conducted by PHS be included under the provisions of the Federal Employees Compensation Act, it suggested that compensation for subjects in PHS-supported research be provided by the research institutions receiving financial support. A similar recommendation was made regarding compensation for subjects of research on new drugs and medical devices regulated by FDA.

Although the recommendations of the Task Force were endorsed in June 1977 by the National Commission, they were not readily embraced by the research community (14). This hesitancy did not focus on the ethical argument. Rather, widespread concern was voiced regarding the ability of research institutions to secure private insurance which met their needs for protection, while also not imposing costs or limitations on coverage that would thwart valuable research. These concerns were reflected in the report of the Task Force itself, which found private insurers extremely reluctant to consider compensation coverage, given the diversity of research institutions, studies and subjects, as well as the absence of actuarial data on which to formulate premiums.

While the controversy continued regarding the practical contours of a compensation program, the National Commission addressed the issue from another standpoint. In its report on institutional review boards, the Commission outlined those items of information that a reasonable person in the prospective subject's position would need to know in deciding about participation in research. The Commission proposed that a reasonable person would need to know whether medical treatment and compensation is available in the event of injury (15). In January 1979, this recommendation was adopted by the Department of Health, Education, and Welfare (DHEW), which now required that, for research involving more than minimal risk, institutions receiving federal funds for research must inform prospective subjects regarding the provision of medical treatment for injuries, the availability of monetary compensation, and the identity of the appropriate individual to contact in the event of injury (16). This requirement of informed consent disclosure remains in effect today.

With the substantive issue of a compulsory compensation program still unresolved, the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (hereafter, the President's Commission) decided in 1980 to undertake a fresh investigation of the issue. In developing its recommendations, the Commission reviewed papers on the ethical and legal issues, reports on the few existing compensation programs, an analysis of the risks of common research procedures, and assessments of alternative program designs (14). The conclusions it reached were more narrowly drawn and tentative than those of the HEW Secretary's Task Force. The Commission affirmed that society has an obligation of justice to compensate injured research subjects. However, it considered this a prima facie obligation that might be outweighed by other duties of social justice if the need for a program was negligible or the costs of administering it were disproportionate to the need. Moreover the Commission restricted the scope of the obligation to injuries incurred as a result of nontherapeutic research procedures. With respect to therapeutic research procedures, the Commission found little evidence that subjects who seek therapy in the research setting expose themselves to risk exceeding that presented in the nonresearch setting. In addition the Commission asserted that distinguishing excess injury in subjects who are seriously ill patients posed administrative burdens that could not be justified in light of the relatively weaker moral claim to compensation for injuries resulting from therapeutic research procedures. The Commission also identified numerous practical aspects of administering a program whose appropriate form could not be settled without some limited experience in implementing such a program. Lastly, the Commission believed that insufficient data on research injuries existed to establish a compelling need for a program. Based on these factors, the Commission recommended that the (now) Department of Health and Human Services (DHHS) conduct "a small, controlled experiment to determine whether a formal program is needed and, if so, the most fair and efficient means of providing compensation." The report of the President's Commission was conveyed to President Reagan in June 1982. In an era marked by shrinking federal programs, no action was taken on its recommendations.

In the intervening years the issue of compensation for research injuries has been revisited only once at the federal level. In 1995 the Advisory Committee on Human Radiation Experiments issued its report on radiation research on human subjects sponsored by the federal government and conducted during the period 1944 through 1974 (17). The Advisory Committee recommended that the federal government financially compensate subjects or next of kin in cases where the government deliberately withheld information from individuals or families regarding the nature of the research. It also recommended compensation for subjects who were physically injured in studies involving either no prospect of direct benefit to the subjects, or the use of controversial interventions that were misrepresented as standard practice. However, the Advisory Committee did not examine general issues regarding implementation of a no-fault, federal compensation program for injured research subjects.

MORAL BASIS OF COMPENSATION

Moral arguments favoring compensation of injured research subjects follow two strategies. Some proponents have argued that establishing a compensation program will have consequences that improve the overall cost–benefit ratio of clinical research. Other analysts have contended that compensation is owed to injured research subjects as a moral right, regardless of its beneficial consequences. This moral right reflects the obligation of society to justly distribute the benefits and burdens of cooperative social activities like clinical research.

The view that a compensation program is justified by its cost/benefit consequences was articulated by early theorists who viewed it as a mechanism for the social control of research (10,11). Several potential benefits of no-fault compensation were envisioned. One is that compensation would encourage investigators to engage in less dangerous research, because the costs of compensating injured research subjects would be drawn from the total funds available for research. Also reduction of risk in planning research would be facilitated by accumulating data regarding the compensation cost profile of common research interventions used in specific subject groups. In

addition compulsory compensation would encourage the allocation of more resources for the early identification and treatment of injuries incurred by subjects. Another goal achieved would be increased public support for the research enterprise, deriving from awareness that appropriate resources are devoted to the needs of injured subjects. Moreover, availability of compensation would encourage participation in research by persons who might demur in the absence of protection against personal loss. Finally, a compensation program might facilitate the conduct of some clinical research that is unusually risky but promises significant advances in medical knowledge. Thus, on this view, society ought to provide compensation for research injuries because it will lead to safer research, encourage societal support and participation, and facilitate the conduct of some risky but highly promising studies.

There are serious shortcomings in this approach to grounding a societal obligation to provide compensation. One problem is that the argument depends on unproven factual assumptions about the incentives for researchers and subjects created by a compensation program. For example, the argument assumes that the availability of compensation would increase recruitment of subjects, even though there is no evidence that persons now decline participation due to the absence of this guarantee. Similarly the claim that research will be practiced more safely if compensation costs are included in the total funding for studies presupposes that current review mechanisms and the conscientiousness of investigators are inefficient in minimizing risks. A second problem is that the incentives created by a compensation program may depend on its exact administrative configuration, and particular schemes may inconsistently both promote and undermine specific goals cited. For example, if a federal compensation program is financed through funds separate from research budgets, it may have little deterrent effect on investigators contemplating the use of risky research procedures. At the same time, greater public awareness about research injuries might discourage research participation, irrespective of the availability of compensation. A third problem is the most decisive. While the incentives of a compensation program may decrease the total harm caused to human subjects and encourage public support for and participation in clinical research, the moral claim of injured subjects against society is valid even if these beneficial consequences are not realized. That is, it is the injury itself that seems to trigger the obligation to provide compensation. Therefore the moral basis of the obligation to compensate injured research subjects must reside in considerations other than the valuable consequences of a compensation program.

The second major approach to delineating a moral basis for compensation proceeds from this important insight. This view asserts that compensation is owed to injured research subjects as a matter of social justice. According to one common interpretation, justice requires that the benefits and burdens of certain cooperative social endeavors be distributed in ways that give all persons an equal opportunity for a good life. When persons are injured in these activities, compensation may help to restore the equality of their opportunity vis-à-vis other members of society.

A crucial question is what cooperative social endeavors create this obligation of justice. In a position that strongly influenced both the HEW Secretary's Task Force and the President's Commission, Childress argues that three features of an activity create a societal obligation of compensatory justice (18). First, the injured party has accepted or has been compelled to accept a position of risk. Objective risks that the party would not otherwise have encountered are created by this position. Second, the activity is for the benefit of society, although any particular person's motives may not be to benefit society. Third, society, through its government or agencies conducts, sponsors or mandates the practices in question. For example, these conditions obtain with respect to persons who serve in the military forces and create an obligation to compensate service-connected injuries. They also apply to the injuries incurred by some research subjects.

Several crucial points should be noted about the scope of the obligation of compensatory justice as specified by Childress. The obligation to compensate research subjects is triggered by the injury without regard to negligence by investigators. Compensation should be provided on a no-fault basis. In addition the obligation to compensate pertains to injuries derived from assuming a position of objective risk that benefits society. The subject's motives to secure a promising new treatment or to receive an attractive cash payment are irrelevant to the right of compensation for injury. Moreover, the obligation to compensate applies to injuries associated with both therapeutic and nontherapeutic research procedures, because both involve a position of risk for the benefit of society. Lastly, the societal obligation to injured subjects applies only to cooperative endeavors officially conducted, supported or mandated by society. Thus research conducted or supported federal agencies, or regulated by FDA are covered by the obligation of compensatory justice. Research privately supported or conducted is not officially endorsed by society and does not create a societal obligation to redress the injuries of subjects.

Despite its intuitive appeal, there are two serious difficulties with this argument. The first relates to the role of consent by prospective subjects to participation in research. When persons knowingly and freely participate in cooperative ventures possessing a risk of harm, other parties are not normally held responsible for injuries that materialize without negligent behavior. This notion is captured by the maxim, *volenti non fit injuria*—there is no injury to one who consents. Moreover, as required by current federal regulations, prospective subjects must be advised as to the availability of compensation. If subjects accept participation in research knowing that it is not available, then their consent is secured with disclosure of the information that a reasonable person would need to know. The choice to participate without this guarantee would seem to relieve any prima facie obligation to redress resulting injuries.

Some commentators respond that the actual process of informed consent is too imperfect to absolve society's obligation to compensate research injuries. The report of the President's Commission delineates some of these

imperfections (14). One is that in some research the risk profile of interventions is not yet established. Prospective subjects must decide about participation without being aware of the nature of unknown or unanticipated risks. In addition experimental evidence suggests that persons tend to discount the importance of substantial harms with a very low probability of occurrence. Finally, it is difficult to convey to prospective subjects a full appreciation for the ways in which specific injuries might be harmful to their interests. In light of these imperfections in knowledgeable decision making, it might be argued that subjects do not really intend to waive their security against the personal losses imposed by injuries.

The argument from imperfections in the consent process is badly weakened by its paradoxical implications. As Engelhardt has pointed out, we are asked to imagine a consent process in which the decision making of prospective subjects is adequate enough to validate their decision to participate in research but not adequate enough to validate their willingness to forgo compensation for injuries (19). Yet the decision to expose oneself to the injury in the first place does not seem to require less ability to comprehend facts and to deliberate about their consequences than the decision to accept research participation without the promise of compensation. Thus, if prospective subjects have sufficient capacity to competently decide about participation in research, they are also competent to accept participation without compensation for injuries.

Another response to the argument that consent relieves the obligation to compensate focuses on the morality of the request rather than the validity of consent (20). We may grant that knowing and willing agreement by subjects relieves society of its obligation. Nevertheless, it may be morally wrong of society to even make such a request. If justice requires that subjects be compensated for injuries incurred in officially sanctioned medical research, then society should not be free to ask persons to forgo compensation. The request for a waiver of this right should be limited to situations in which weightier obligations constrain the use of societal funds for compensation. While persons should be free to decline compensation after they have been injured, society must make the offer as a matter of compensatory justice.

A more serious problem with Childress' argument for the obligation of compensatory justice relates to its scope. For Childress the obligation to compensate research injuries is triggered when persons assume a position of risk in activities that benefit society and are sanctioned through its official agencies. Injuries resulting from both therapeutic and nontherapeutic research procedures are compensable. Likewise injuries are compensable whether the primary motivation of subjects is to contribute to the interests of others or to promote their own interests. The essential difficulty is explaining why the assumption of risk triggers an obligation to compensate injuries when subjects' participation is premised on the judgment that the risks are outweighed by anticipated personal benefits. If subjects are not making a sacrifice for society, then compensation for their losses is not owed them as a matter of justice. The scope of the obligation must be restricted to circumstances in which the positional risk is assumed without the anticipation of offsetting personal benefits.

An example illustrates the problem (21). The federal government sponsors a national lottery to fund educational programs. These programs contribute to the general welfare. Participants in the lottery assume a position of risk. As a cooperative social enterprise, the lottery satisfies the three conditions for compensation of injury cited by Childress. If a losing participant has purchased so many tickets that he faces financial ruin, it follows that he ought to be compensated for his losses. If this implication is unacceptable, then it appears that the scope of the obligation to compensate injuries as sketched by Childress is too broad.

The same difficulty arises with regard to clinical research involving the evaluation of therapeutic interventions (22). Childress assumes that subjects are placed at special risk by participation. However, it is generally agreed that it is not morally permissible to undertake a trial of a new therapy unless prior evidence suggests that it is likely to be at least as efficacious and safe as alternative treatments acceptable to the subject. Moreover, because shortcomings in the efficacy and safety of standard treatments provide the stimulus for evaluating new therapies, it is usually the case that preliminary evidence suggests that the new treatment may be superior. In addition treatment in the research context often includes more intensive monitoring and nursing care for subjects. In many research centers there are also more numerous and specialized ancillary personnel available, including allied health professionals, social workers, and psychologists. Lastly, new treatments being investigated are often provided without charge to the subjects. Thus subjects assume a position of risk to secure overriding personal advantages associated with participation. If there is no sacrifice involved in the subject's participation, then an obligation of compensatory justice does not apply in the event of injury.

Most commentators have assumed that the same factors do not apply to injuries resulting from nontherapeutic research procedures. Subjects injured by these interventions cannot anticipate commensurate medical benefits. If justice requires that the impact of these burdens be ameliorated, then it is appropriate to redress these injuries. But even here, the motivation of subjects complicates the moral equation. In some cases subjects participate in research involving nontherapeutic procedures to secure an attractive payment. Rather than being an act of sacrifice, participation reflects a calculation of overriding personal benefit. Nevertheless, in many cases, research involving nontherapeutic procedures does not involve payment for participation, and subsequent injuries to subjects create the societal obligation to provide compensation.

If many subjects do not incur special risks, then the obligation to compensate injuries is more narrowly circumscribed than envisioned by Childress and the HEW Secretary's Task Force. However, there are ways to blunt the force of this argument. First, the risk–benefit profile of a new therapy undergoing evaluation is, *ex hypothesi*, not yet clearly established. Although preliminary evidence may suggest that it will be at least as favorable as

alternative treatments, the new therapy may prove inferior in a controlled clinical trial. Second, there is always the danger that unknown or unanticipated risks of new therapies will emerge during or after the conduct of clinical trials. Participating subjects expose themselves to these uncertainties that are not present in the nonresearch setting. Third, for subjects who incur injuries in excess of what might be anticipated in the nonresearch setting, it can be maintained they have clearly made a sacrifice that contributes to the general welfare of society. If these considerations are decisive, then the obligation to compensate research injuries should cover therapeutic as well as nontherapeutic interventions.

Even if the argument for an obligation of compensatory justice is compelling, it establishes only a prima facie societal obligation to compensate research injuries. There may be other societal needs, or other obligations of social justice, which have moral priority in the use of limited societal resources. Thus the case for compensation may depend on the extent of the problem of research injuries, as well as the importance of this problem relative to other pressing social needs.

EXTENT OF RESEARCH INJURIES

In the period right after World War II, medical research was frequently depicted as a perilous enterprise exposing subjects to grave risks of harm that would not otherwise be encountered. However, this assumption was not based on empirical data about the frequency and magnitude of injuries. Even today there is only one large descriptive study of injuries in clinical research and the accumulated experience of a few small compensation programs. Available information raises serious questions about whether the frequency and magnitude of research injuries establishes a compelling need for a national compensation program.

A large empirical study was performed by Cardon and colleagues for the HEW Secretary's Task Force (23). Data were compiled by telephone from 331 investigators conducting research on 133,000 human subjects. Overall, injuries were reported in 3.7 percent of all subjects. Eighty percent of these injuries were classified as trivial by the investigator, while nearly all the remaining were temporarily disabling. Permanently disabling and fatal injuries together accounted for about 1 percent of the total injuries.

Separate analyses were performed for therapeutic and nontherapeutic research procedures. Of 39,216 subjects undergoing therapeutic research procedures, 10.8 percent were reported injured. Trivial injuries were incurred by 8.3 percent of subjects and temporarily disabling injuries by 2.4 percent, with less than 0.1 percent suffering either permanently disabling injuries or death. In actual numbers, 13 subjects were permanently disabled and 43 died. Of 93,399 subjects undergoing nontherapeutic research procedures, 0.8 percent were injured. Trivial injuries accounted for 0.7 percent of all subjects, temporarily disabling injuries 0.1 percent, and permanently disabling injuries less than 0.1 percent. Only one subject was permanently disabled, and there were no fatalities.

The investigators also determined that injuries resulting from therapeutic procedures were clustered in clinical trials of cancer therapies. These trials accounted for 37 of 43 fatalities, 9 of 13 permanently disabling injuries, and 648 of 937 temporarily disabling injuries. Except for these cancer trials, there were few serious and/or irreversible injuries reported. Moreover, investigators were asked to enumerate actual injuries, without regard to whether the same results might have been expected from treatment in the nonresearch setting. This point is significant with respect to research on cancer treatments. Standard treatments for cancer carry substantial risks of disabling injuries or death. The study design did not permit investigators to determine whether the frequency and magnitude of injuries in cancer trials was greater or less than expected in the nonresearch setting.

Finally, investigators broached the question of how the risks of harm to subjects undergoing nontherapeutic research procedures compare to the risks of everyday life. Annual rates of accidental injuries that are temporarily disabling, permanently disabling and fatal, per 100,000 Americans, were found to be about 50, 2, and 0.6, respectively. Rates for the same types of injuries in approximately 93,00 research subjects were 37, 1, and 0. If the average duration of participation in research employing nontherapeutic procedures is 3 days, or 1/100 of a year, then these rates do not seem significantly greater than the risks of everyday life.

The President's Commission also reviewed data on research injuries from existing compensation programs in the United States. McCann and Pettit analyzed data from the University of Washington for 1972 to 1981 during which an estimated 356,000 subjects were covered by the university's compensation program (24). In the year prior to initiation of the program, 10 out of 14,942 subjects (0.07 percent) experienced adverse effects. None were partially or permanently disabled, and no deaths were reported. For the first eight years of the compensation program, 144 out of 356,000 subjects (0.04 percent) experienced temporary disability, none were permanently disabled, and two died. In addition a report was received from the Quincy Research Center in Kansas City, Missouri, whose subjects were covered by a workers' compensation program (25). The data covered 2596 normal volunteers and 2478 patients. Clinically adverse events resulted in the hospitalization of 1.43 percent of patients and 0.2 percent of normal volunteers, while fatalities occurred in six patients and no normal volunteers. Analysis suggested that no deaths in the patient group, and only 7 of 36 hospitalizations could be related to the research interventions. Finally, the Commission reviewed the results of the drug testing program using normal volunteers in Michigan prisons for the years 1964 to 1976 (26). The authors reviewed 805 protocols involving 29,162 research participants. They reported that 64 subjects experienced adverse reactions (0.2 percent), but there was only one permanently disabling injury and one fatality (in a patient receiving a placebo). Thus data from these research programs corroborate the findings of Cardon and associates that serious research injuries are very infrequent.

An important policy issue is whether the extent of research injuries is sufficient to transform a prima facie moral obligation to offer compensation into a legitimate social priority. As outlined earlier, the moral argument for compensating research injuries is most compelling with regard to subjects harmed by nontherapeutic procedures. However, the evidence generated by Cardon et al. suggests that the frequency and magnitude of injuries resulting from nontherapeutic procedures is no greater than the risks of accidental injury in daily life. The moral argument for compensating injuries caused by therapeutic procedures is less compelling, because subjects often accept the risks of harm to secure the offsetting benefits of therapy in the research setting. Here the Cardon study found a greater frequency and magnitude of injuries to subjects. Nevertheless, these injuries were clustered in cancer treatment trials, and no attempt was made to determine whether the extent of injuries exceeded that anticipated in the nonresearch setting. Thus available empirical data does not provide convincing evidence that the prima facie obligation to compensate injured subjects should be a social priority.

The President's Commission concluded that existing evidence was insufficient to establish a compelling need for a national compensation program (14). In recommending experimental, pilot programs for compensating injured subjects, the Commission indicated that gathering more comprehensive data on the extent of research injuries should be a crucial component of this endeavor. In this way the data necessary to an informed social policy decision on the relative need for a compensation program could be generated.

ELEMENTS OF A COMPENSATION PROGRAM

Even if the moral argument for compensation is persuasive and the prevalence of research injuries significant, numerous questions remain regarding the design of a morally acceptable program. A leading question involves how compensable injuries are to be defined. The challenge is to isolate those harms to subjects that exceed what would have occurred if they had not participated in research. With regard to nontherapeutic research procedures, the resolution is relatively straightforward. Any injuries that are reasonably related to the interventions should be compensable. Because the procedures are performed for reasons unrelated to the welfare of subjects, any injuries suffered would not have occurred but for their participation in the research.

The problem is considerably more complex in dealing with therapeutic research procedures. Subjects undergoing therapeutic procedures are patients. As patients, they may incur harms from the progression of their disease. In addition, even if they do not participate in a research study, they receive treatments that may cause harm. Therefore it becomes difficult to specify which injuries would not have occurred but for participation in research.

In its recommendations, the HEW Secretary's Task Force proposed the following criterion for compensable injuries:

> Human subjects who suffer physical, psychological, or social injury ... should be compensated if (*1*) the injury is proximately caused by such research, and (*2*) the injury on balance exceeds that reasonably associated with such illness from which the subject may be suffering, as well as with treatment usually associated with such illness at the time the subject began participation in the research (13).

Although intuitively plausible, application of the "on balance" test in the clinical setting presents at least four serious problems. These problems are well-illustrated by clinical trials of cancer treatments, where physical injuries often occur (22).

First, it is frequently not possible to determine whether the disease or the treatment has caused an identified harm. For example, in treating acute leukemia, both the leukemic process and chemotherapy may suppress bone marrow function in a way that disposes to serious infection or bleeding. Second, the "on balance" test relies on a distinction between investigational and standard treatment. However, in some forms of disseminated adult cancer, no standard treatment of choice has emerged. It may not be clear with what treatment an investigational therapy should be compared to determine excess injury. Third, the "on balance" test requires comparison of the overall frequency of various harms and benefits associated with research and conventional therapies. But it is uncertain how to compare incommensurable harms and benefits. For example, does compensable injury occur when an investigational chemotherapy for bone cancer increases the median length of survival by four months but involves a higher incidence of acute hearing loss and severe nausea and vomiting than standard therapy? A final problem is that all chemotherapies for disseminated cancer involve serious risks and fail to achieve remission in a definite percentage of cases. Thus, although the harm–benefit ratio of an investigational therapy in a series of subjects may be less favorable than with standard treatment, it is not possible to determine whether the outcome for a particular subject has, "on balance," resulted in more harm than standard treatment. Thus the President's Commission found that this test for identifying compensable injuries posed "enormous burdens of administration" that could not be justified in light of the less compelling moral claim to compensation of subjects undergoing therapeutic research procedures (14). Of course, if compensable injuries are restricted to nontherapeutic research procedures, these problems are avoided.

A second important issue involves the types of injuries for which compensation will be provided. The HEW Secretary's Task Force recommended coverage for physical, psychological, and social injuries. However, DHEW's interim rule restricted informed consent disclosure regarding the availability of compensation to physical injuries. The President's Commission argued that compensation should focus solely on physical and psychological injuries. The Commission defined "social" injuries as harms to subjects' reputation, personal relationships, or legal status resulting from unauthorized disclosure of personal information gathered in research. It suggested that problems

related to confidentiality could be handled effectively in the prospective review of protocols (14).

The moral analysis of the obligation of compensatory justice provides some guidance in selecting relevant categories of injuries. The principle of justice requires that all persons have equal opportunity to pursue their life plans. Categories of injuries are pertinent if failure to compensate them might seriously compromise the opportunity of injured subjects vis-à-vis other persons. Serious and irreversible physical injuries obviously satisfy this criterion. Likewise serious psychological injuries, such as increased frequency of clinical depression resulting from an experimental drug for bipolar affective disorder, may substantially impair the opportunity of subjects to pursue their life plans. Moreover, even if we restrict "social" injuries to unauthorized disclosure of confidential information, there is no reason in principle to make these injuries ineligible for compensation. Injuries to reputation, personal relationships, or legal status may gravely impair the ability of persons to pursue their goals.

Nevertheless, compensatory justice constitutes only a prima facie societal obligation. Other requirements of social justice, or other compelling social needs may present weightier moral demands on limited resources. These limitations may necessitate that the relative burdens imposed by different categories of research injuries be ranked for funding priority. If the frequency and magnitude of "social" injuries is far less significant in abridging the opportunity of subjects to pursue their life plans, then compensation for these harms may not achieve funding priority.

A third crucial component in the design of a compensation program concerns the nature and extent of benefits for compensable injuries. Benefits might include medical care for injuries, lost wages, monetary awards for pain and suffering, and death benefits for survivors. At present, some academic institutions and pharmaceutical companies provide short-term medical care for injuries incurred by research subjects. With few exceptions, other benefits are not provided.

Our understanding of the obligation of compensatory justice also provides guidance in determining relevant types of benefits. The function of compensation benefits is to restore any deficit in the ability of injured subjects to pursue their life plans compared to citizens who are not injured in research (20). This suggests that benefits should be provided to redress any negative consequences that flow from a compensable injury and that impair the ability of injured subjects to pursue their life plans. These negative consequences obviously include the costs of short-term and long-term medical care, earnings lost due to time in treatment and residual disabilities, as well as loss of income support for families in the event of the premature death of subjects.

However, monetary awards for pain and suffering are more controversial. On one hand, it is clear that pain and suffering may result from compensable injuries and may seriously erode the capacity of persons to pursue their goals. On the other hand, there are practical difficulties in devising a benefits schedule for pain and suffering. One problem is that degrees of pain and suffering can differ substantially for different persons with similar compensable injuries. Another is measuring the pain and suffering experienced by specific individuals. A third problem is that, when pain and suffering is genuinely disabling, much of its harmful impact will be ameliorated by benefits for items such as the cost of medical care and lost wages. Isolating the residual impact of pain and suffering is exceedingly difficult. These problems have led some commentators to suggest that providing compensation for pain and suffering poses insuperable administrative burdens. Nevertheless, one simple solution would be to provide a fixed percentage of the other total benefits as a payment for pain and suffering. This would avoid the administrative complexities of individual determinations, while acknowledging the obligation to rectify its harmful consequences (27).

The status of compensatory justice as a prima facie societal obligation also affects selection of the types and amount of benefits. Other obligations of social justice may limit total government funds that are properly allocated to the compensation of injured subjects. As a result there may be overriding moral reasons for limiting the types or amounts of benefits for injured subjects. Once again, such limitations would require establishing priorities among the types of benefits to be provided.

Another relevant variable in the design of a compensation program concerns the scope of covered research activities. According to Childress, the societal obligation of compensatory justice arises only for research done at the behest of society, as sanctioned through official government agencies (18). This condition reflects the notion that society should redress sacrifices that it compels or encourages for the common good. In accepting this proposition, the HEW Secretary's Task Force recommended that a compensation program cover injuries to subjects in research conducted or supported by the federal government. It also recommended that FDA consider legislation to require compensation in research conducted under regulations for testing the efficacy and safety of new drugs and medical devices (13). The President's Commission agreed that the societal obligation to offer compensation is strongest when the nexus between the government and the research activity is closest. However, the Commission recommended that the pilot compensation program include only research conducted or supported by the federal government, while requesting voluntary participation by private research institutions conducting FDA-regulated research (14). Thus research funded by private philanthropic organizations is not covered by the societal obligation of compensatory justice in the view of either advisory body.

This conclusion can be challenged by reconsidering the factors that determine whether research is conducted with the official sanction of society (1). Privately funded, non-regulated medical research may produce generalizable medical knowledge. The results are typically published in peer-reviewed, academic medical journals. The information is available to other investigators in designing their own research studies. Insofar as the use of these results is permitted in designing research conducted, supported or regulated by the federal government, it might be concluded

that these privately funded research studies are officially sanctioned by society. If this argument is compelling, then equity requires that subjects injured in privately funded medical research be covered by society's obligation of compensatory justice. Legitimate worries about the prospective social control of such research could be addressed by extending the use of institutional review boards to all human subjects research occurring within the United States.

A final critical design variable concerns the source of funding for the compensation program. There are two basic alternatives: governmental or nongovernmental funding. A federally funded program would likely be administered by a government agency, although an alternative might involve provision of funds in research grants for institutions to secure private insurance. Nongovernmental funding would involve statutory or regulatory standards for compensation programs, with research institutions being required to provide benefits through self-insurance, private insurance, or cooperative insurance pools. Some combination of funding might also be possible, with research institutions securing basic coverage for compensation benefits and the government providing backup insurance against catastrophic losses.

There are complex issues regarding what funding mechanism would permit the most efficient and effective administration for a compensation program (28). The moral implications of these alternatives are more limited. If the moral justification for a compensation program lies in its beneficial consequences, then alternative funding mechanisms may create incentives that differentially impact on achievement of these goals. For example, if a compensation program is justified as a mechanism for controlling exposure of subjects to risks and for encouraging surveillance of their safety, then a program in which research institutions purchase private insurance is more likely to provide incentives for safe practice. The costs of coverage for research institutions will depend partly on their success in limiting the frequency and magnitude of compensable injuries to subjects. An alternative moral justification for redressing research injuries is based on a societal obligation of compensatory justice. In this case the primary goal is to ensure that benefits are provided to injured subjects. Funding by the government rather than research institutions is more likely to encourage investigators to assist subjects in identifying compensable injuries and securing adequate benefits. If investigators must worry about institutional expenditures for compensation provided through private insurance, then incentives to discourage legitimate claims are created. The modest level of claims in the few compensation programs funded by research institutions suggests that cost consciousness inhibits advocacy for subjects with compensable injuries.

SUMMARY

Although initial discussions of compensation for research injuries were stimulated by research whose conduct violated the rights or welfare of subjects, subsequent focus has been on injuries occurring without negligence. Numerous government advisory panels have accepted the moral claim that society has an obligation to compensate human subjects for injuries incurred when they assume a position of risk in officially sanctioned research activities contributing to the common good. Despite the intuitive plausibility of this proposition, implementation of a national compensation program has been stymied by a series of vexing issues. The applicability of the obligation to subjects injured by therapeutic research procedures has been widely challenged. A compelling need for a federal program has never been clearly established. Complex issues arise in defining compensable injuries, determining appropriate benefits, clarifying the scope of covered research, and designing a suitable funding mechanism. Implementation of federal regulations for the prospective review of research protocols has substantially reduced incidences in which the plight of injured subjects has received public attention. In an era of shrinking federal programs and without public perception of a palpable need, the prospects for a national compensation program are not promising.

BIBLIOGRAPHY

1. B.R. Adams and M. Shea-Stonum, *Case Western Reserve Law Rev.* **25**, 604–648 (1975).
2. G.R. Smith, in HEW Secretary's Task Force on the Compensation of Injured Research Subjects, *Report*, Appendix A, DHEW Publ. No. (OS) 77-004, U.S. Department of Health, Education, and Welfare, Washington, DC, 1977, pp. 27–42.
3. R.J. Levine, *Ethics and Regulation of Clinical Research*, 2nd ed., Urban & Schwarzenberg, Baltimore and Munich, 1986.
4. D.J. Rothman, *Strangers at the Bedside*, Basic Books, New York, 1991.
5. H.K. Beecher, *N. Engl. J. Med.* **74**, 1354–1360 (1966).
6. J. Jones, *Bad Blood: The Tuskegee Syphilis Experiment*, 2nd ed., Free Press, New York, 1993.
7. W.J. Curran, in P.A. Freund, ed., *Experimentation With Human Subjects*, George Braziller, New York, 1969, pp. 402–454.
8. U.S. Department of Health and Human Services, *Fed. Regist.* **46**(16), 8366–8392 (1981).
9. I. Ladimer, *J. Chronic Dis.* **16**, 1229–1235 (1963).
10. C.C. Havighurst, *Science* **169**, 153–157 (1970).
11. G. Calabresi, in P. Freund, ed., *Experimentation with Human Subjects*, George Braziller, New York, 1969, pp. 178–196.
12. Tuskegee Syphilis Study Ad Hoc Advisory Panel, *Final Report*, U.S. Department of Health, Education and Welfare, Washington, DC, 1973.
13. HEW Secretary's Task Force on the Compensation of Injured Research Subjects, *Report*, DHEW Publ. No. (OS) 77-003, U.S. Department of Health, Education, and Welfare, Washington, DC, 1977.
14. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Compensating for Research Injuries*, vol. 1: Report, Stock No. 040-000-00455-6, U.S. Government Printing Office, Washington, DC, 1982.
15. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *Report and Recommendations: Institutional Review Boards*, DHEW Publ. No.

(OS) 78-0008, U.S. Government Printing Office, Washington, DC, 1978.
16. U.S. Department of Health, Education and Welfare, *Fed. Regist.* **43**(214), 41449 (1978).
17. Advisory Committee on Human Radiation Experiments, *Final Report*, Stock No. 061-000-00-848-9, U.S. Government Printing Office, Washington, DC, 1995.
18. J.F. Childress, *Hastings Cent. Rep.* **6**(6), 21–27 (1976).
19. H.T. Engelhardt, Jr., in HEW Secretary's Task Force on the Compensation of Injured Research Subjects, *Report*, Appendix A, DHEW Publ. No. (OS) 77-004, U.S. Department of Health, Education and Welfare, Washington, DC, 1977, pp. 45–63.
20. B.R. Boxill, in President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Compensating for Research Injuries*, vol. 2: Appendices, U.S. Government Printing Office, Washington, DC, 1982, pp. 41–55.
21. H.M. Smith, in President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Compensating for Research Injuries*, vol. 2: Appendices, U.S. Government Printing Office, Washington, DC, 1982, pp. 19–39.
22. T.F. Ackerman and A.M. Mauer, *N. Engl. J. Med.* **305**, 760–763 (1981).
23. P.V. Cardon, F.W. Dommel, and R.R. Trumble, *N. Engl. J. Med.* **295**, 650–654 (1976).
24. D. McCann and J.R. Pettit, in President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Compensating for Research Injuries*, vol. 2: Appendices, U.S. Government Printing Office, Washington, DC, 1982, pp. 241–274.
25. J.D. Arnold, in President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Compensating for Research Injuries*, vol. 2: Appendices, U.S. Government Printing Office, Washington, DC, 1982, pp. 275–302.
26. C.J. Zarafonetis et al., *Clini. Pharmacol. Ther.* **24**, 127–132 (1978).
27. C.C. Havighurst, in HEW Secretary's Task Force on the Compensation of Injured Research Subjects, *Report*, Appendix A, DHEW Pub. No. (OS) 77-004, U.S. Department of Health, Education, and Welfare, Washington, DC, 1977, pp. 81–132.
28. R. Zeckhauser, in HEW Secretary's Task Force on the Compensation of Injured Research Subjects, *Report*, Appendix A, DHEW Publ. No. (OS) 77-004, U.S. Department of Health, Education, and Welfare, Washington, DC, 1977, pp. 155–166.

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, ETHICS, FAMILY, AND PEDIGREE STUDIES

GEORGIA L. WIESNER
SUSAN LEWIS
JENNIFER SCOTT
Case Western Reserve University
University Hospitals of Cleveland
Cleveland, Ohio

OUTLINE

INTRODUCTION

Our current understanding of human genetic disease is a direct result of family studies research and the explosion of genetic technology over the last half-century. The benefits of this research have been enormous, and range from the discovery of specific genes causing common diseases, such as cancer or heart disease, to a deeper understanding of the mechanics of how environmental factors interact with the human genome or set of genes. However, these advances in genetic technology often alarm the lay public. Indeed, one harsh view is that scientists "have given no more thought to the potential social applications of genome mapping and sequencing than Victor Frankenstein had given to the consequences of creating his monster..." (1, p. 660). Nevertheless, a revolution in biomedical technology and genetics has taken place in recent years, so that genetic studies based on family research are now the mainstays of biomedical research.

The genetic revolution now enjoys governmental support through the Human Genome Project (HGP) (2), a large international effort to sequence and map each of the 24 human chromosomes as well as to understand the underlying genetic variation and function of the genome (3,4). Indeed, HGP is a monumental undertaking that seeks to unravel the human "genetic code" discovered only 47 years ago by Watson and Crick (5). Francis Collins, the current head of the National Human Genome Research Institute (NHGRI) in the National Institutes of Health (NIH) describes HGP as the "single most important project in biology and the biomedical sciences — one that will permanently change biology and medicine" (3).

Both supporters and detractors of HGP predicted that this undertaking will have unprecedented effects on the social fabric of the human community (1). In response to these concerns, the analysis of the ethical, legal, and social implications of this new genetic knowledge has become a vital component of HGP. According to Eric Juengst, the

first chief of HGP's Ethical, Legal and Social Implications (ELSI) branch (italics added):

> . . .the Human Genome Project represents geneticists' growing ability to explore human heredity. It thus generates a wide range of charged questions about how our society's genetic explorations should proceed and how their results should be used. These questions include a set of *unresolved issues regarding the conduct of research involving human subjects*. While these issues predate the Human Genome Project and would continue to exist without it, they are becoming increasingly important as genomic tools and *genetic strategies become pervasive* in biomedical research (6, p. 401).

Juengst and others have thus identified two essential themes for human genetics research (6,7). First, researchers who conduct genetic studies will face new ethical challenges that will require the implementation of specific strategies in the study protocol addressing these issues. Second, because more common diseases and traits will be studied in the future, genetic research will increase in scope and complexity (6,8).

Many human genetic studies such as gene-searching projects sponsored by HGP are based on the analysis of the family group. These are collectively known as pedigree or family studies and are an integral component of modern genetic research. Family pedigree studies were previously considered to have little, if any, inherent ethical conflicts by researchers in the field because the studies were often based on observations of a rare condition in a small number of families. In joining a proposed research study, these families were eager to discover the genetic link to their condition, and often worked closely with the research team. The potential impact of this research on such rare conditions was limited to a few families with these unusual disorders. In contrast, present day family pedigree studies are often focused on the familial nature of common diseases and are conducted on a national and international scale. Thus these studies can potentially impact many more people in communities across the country. Contemporary family pedigree projects are a hybrid between standard epidemiological research and molecular analysis of gene function, in which large numbers of individuals and their family members are enrolled to identify genes that cause human disease.

A crucial difference between family pedigree studies and other biomedical research lies in the collection and analysis of information on family groups rather than unrelated individual volunteers (9). Epidemiologic protocols analyze clinical information on a large number of cases, and thus the risks and benefits of these studies stem from the potential impact on an individual volunteer. In contrast, in family pedigree research, the information gained from one family member can lead to consequences for other family members (6,9). Holtzman and Andrews further argue that genetic research "is different because it often involves testing, and thus *creates* genetic information about individuals and groups that did not exist before" (10). Thus the ethical issues arising from family pedigree research may be inherently more difficult for researchers in the biomedical community. This difficulty may stem from the lay public's view of genes and heredity. Henry T. Greeley, professor of law at Stanford University writes:

> Rightly or wrongly, many people are convinced that genes are special, that they contain and reveal a person's, or a people's, essence, which has enormous value, spiritual and commercial. This exaggerated emphasis on the importance of individual genetic variation makes human genomic research particularly sensitive (11, p. 625).

Participating in genetic research can have significant psychosocial consequences for research volunteers (10). By simply enrolling in a pedigree research study, participants may experience unforeseen psychosocial anxiety and depression. These feelings stem from concern over the propensity to develop a disease, potential genetic discrimination, and the consequences of medical decisions based on perceived risk for disease (6,7,10). Human geneticists who study rare monogenic disorders have recognized these psychosocial issues for many years (12). Current and future researchers aided by the enormous growth in genetic research will focus on common disease processes. In this way a greater proportion of the general population could potentially learn they are at risk for health problems because a relative participated in a family pedigree study. Thus the potential for unforeseen harm to participants and their family members will increase, as genetic family-based studies become more prevalent.

The explosion in the number of family pedigree projects has generated a lively debate in the biomedical community as to the proper ethical conduct of genetic research. Most ethicists agree that the Belmont Report defines the moral guidelines for researchers who enroll human subjects in biomedical research. The three overarching principles contained in this document are respect for persons, beneficence, and justice for human subjects (7,13). However, the authors of the Belmont Report did not specifically address issues relating to use of genetic material and the potential impact of family pedigree research. Several organizations, such as the American Society of Human Genetics (ASHG), the National Bioethics Advisory Commission (NBAC), and the Institute of Medicine have recognized the need for a standardized policy for the conduct of genetic research (14–18). The Institute of Medicine's Committee on Assessing Genetic Risks summarized the relevant ethical principles for genetic research subjects and recommended "vigorous protection be given to autonomy, privacy, confidentiality and equity" (15).

Using unconfirmed research findings in medical decision-making is also a major ethical concern for genetic researchers and has prompted several policy statements by interested groups (17,19). This issue gained national attention in the mid-1990s when BRCA1, the first susceptibility gene for early-onset breast cancer was identified and cloned (20). Most researchers and lay persons viewed the application of genetic markers to forecast the risk of future disease as a positive advance in biomedical sciences (21). However, some groups felt that health decisions based on unsubstantiated genetic research without confirmatory clinical trails could be harmful. Recognizing the uncertainty arising from the

use of the new genetic BRCA1 research results, ASHG strongly advised (italics added) "it was *premature* to offer population screening" for BRCA1 gene testing (17). This issue has been raised with other gene discoveries, such as the ApoE link with Alzheimer's disease, since the health implications of genetic research are not limited to breast cancer research (22). Most authors now recommend that genetic results should ideally be provided within the context of clinical care after the health implications are studied, although the pressure to hasten the transition of new genetic findings into the medical arena is pervasive (17,23,24). The blending of scientific discoveries with the potential health benefits of the new research represents a potential conflict for human genetic researchers, who are then cast in the dual role of healer and scientist.

The time is long past when scientists were able to conduct genetic research involving individuals and their families isolated from the ethical and social impact of the research process and eventual findings of the study. In discussing the transition of research findings to patient care, Ray White from the University of Utah states:

> Human geneticists have a problem. Finally, after years of effort, we are beginning to resolve and identify the genetic components of a number of genetically transmitted disorders and predispositions. On the eve of this scientific triumph, however, at a time when we should be delivering this new knowledge to affected individuals, we have instead discovered that this delivery is compromised by social, economic, and ethical issues (25, p. 173).

It is thus imperative to incorporate ethical strategies into all aspects of study design. These strategies should consider the ascertainment of subjects and family members, information supplied to study participants, control and databasing of study data, disclosure of results, publication of data, and communication among researchers.

This article addresses the many ethical dilemmas faced by genetic researchers who perform family pedigree research. There is a growing realization that conflicts arising from family studies will be encountered in other types of genomic research projects. These genomic studies explore a variety of genetic topics such as the translation of genetic discoveries into clinical practice or analyses of genetic variation between specific human ethnic groups and subpopulations (26). Thus equal attention to the medical and psychosocial impact of the genetic research must continue in parallel with the exciting biomedical and genomic research of the future.

GENETIC RESEARCH ON HUMAN POPULATIONS

A major goal of genetic research is to understand the hereditary factors that cause human disease. Studies based on the family unit have been essential to this research over the last half-century and will continue to form the foundation for future investigations (27). Several older methods, such as twin and family pedigree studies, were developed before the introduction of molecular technology, emphasizing the notion that important genetic information can be gleaned from an individual's family history alone. Geneticists then developed sophisticated study methods, such as linkage analysis and marker association studies, to exploit newly discovered molecular markers for the analysis of family information. While the types of analyses and research goals differ, a common feature of these study methods is the collection and statistical analysis of a trait, or phenotype, in multiple family members. Although most biomedical genetic research focus on specific disease process, the amount of clinical information gathered from each family member and the extent of the family history is variable for each study design (Table 1). For example, twin studies gather complete information about both twins but may not collect information about other relatives. In contrast, family pedigree or linkage studies collect clinical information about many members in the extended family.

Twin studies have been used extensively by researchers to support or refute the genetic nature of a specific disease process or physical attribute, such as height. This method compares the prevalence of a disease in identical or monozygotic (MZ) twins and in fraternal or dizygotic (DZ) twins. Since MZ twins share 100 percent of their genetic material, it is logical to expect that both members of an MZ twin-pair would develop the condition if an underlying gene causes the disease. However, both members of a DZ twin-pair would be less likely to be affected as they only share about 50 percent of their genes. This approach is the basis of a 1991 twin study demonstrating a strong genetic component for the development of asthma (28). MZ twins participating in the study both suffered from asthma or allergies 80 percent of the time compared to 0 percent of the DZ twin volunteers. This asthma study illustrates that genetic information can be gained from the analysis of phenotypic and family information alone, since the researchers did not analyze genetic material from the twin participants.

The risk of a family member developing a specific disease can also be estimated from family-based studies without the analysis of DNA markers. One example of the clinical usefulness of family information is from the

Table 1. Types of Human Studies Used by Researchers to Determine the Genetic Component of Human Disease

Study Design	Human Subjects	Use of Family Information	Use of Genetic Material
Twin	Monozygotic and dizygotic twins	Minimal	No
Segregation analysis	Multigenerational families	Extensive	No
Linkage analysis	Multigenerational families	Extensive	Yes
	or		
	Relative pairs (i.e., sibling pairs)	Immediate family members	Yes
Molecular epidemiologic	Individual cases and nonrelative controls	Minimal	Yes

National Polyp Study, a multicentered clinical trial that examined family history as a risk factor for colon cancer (29,30). The study showed that a family history of colon cancer and colon polyps increased the risk of developing colon cancer two- to threefold over cases with no family history of the disease (30). This study compared the trait in question in affected persons to the occurrence of the trait in biological family members. Importantly, the results from this study were then used to develop risk profiles for colon cancer, illustrating the far-reaching effect of genetic research in clinical practice.

Researchers use the twin and family history study methods to investigate whether hereditary factors play a role in a particular condition, but these studies cannot be used to establish a pattern of inheritance based on Mendel's laws. Family pedigree studies are required in order to determine whether a trait is segregating in a specific pattern of inheritance in the pedigree. Most hereditary conditions typically follow an autosomal dominant, autosomal recessive, X-linked dominant or X-linked recessive pattern, although nontraditional patterns have been described. In order to determine the inheritance pattern or segregation of a gene or trait in a family group, researchers must rely on studies of the extended family (27). An essential ingredient of such studies is the construction of a family tree from the proband or other informants that includes detailed information about affected and unaffected family members over several generations. The research group then uses this multigenerational information to examine whether a gene is responsible for the trait segregating in the affected kindreds. This type of study, called segregation analysis, is a genetic epidemiological research method designed to measure the likelihood, or chance, that a hereditary factor causes the trait or disease in question (27,31). Because entire family groups are analyzed, segregation analysis can determine the inheritance pattern of a disease. This type of study is often the first piece of evidence tying hereditary factors to a disease process. For example, segregation analyses of several hundred families with breast cancer in the female relatives first suggested the existence of a gene or genes responsible for familial breast cancer (32). On the basis of these studies, researchers were then able to estimate that the gene would be carried by 1 in 500 individuals and would be transmitted as an autosomal dominant trait to other family members (33,34). Likewise segregation analyses showed that a major gene segregating in high-risk families most likely caused Hirschsprung's disease, a form of congenital megacolon (35).

Neither twin studies or segregation analyses are used to identify the exact gene causing a particular disease. However, the chromosomal location for the disease-causing gene can be found when family pedigree information is jointly analyzed with genetic markers, or molecular signposts, from multiple members of a family (36). This powerful genetic method is called linkage analysis because statistical tests are used to "link" the disease exhibited by affected family members to known molecular markers found at regular intervals along each human chromosome. These markers are now easily analyzed in a small sample of genetic material, or DNA, from the person being tested. Since the researchers know the chromosomal location of each marker, the disease-causing gene can be "mapped" to a specific genetic region. Linkage studies, otherwise known as mapping studies, rely on the ability to distinguish between a chromosomal region inherited from one's father and the same chromosomal region inherited from one's mother. The strength of this type of study stems from the ability to track these genetic regions between the parent and child within the family. One of the important advances in molecular technology has been the development of highly informative molecular markers for genetic mapping studies that allow the researcher to distinguish the paternal and the maternal copies of a particular genetic region.

Up to 400 separate genetic markers can be used to blanket the entire genome for mapping studies searching for disease causing genes. The markers can be used for candidate gene analysis or whole scale genome scan. Candidate gene analysis is used when the researcher suspects a known gene causes the disorder. The researchers must also know the chromosomal location of the suspected gene, and will test markers that are physically close to the suspected gene. In contrast, a genome scan is used when the location of the gene or genes is not known. These scans test numerous polymorphic markers diffusely located across the entire genome, resulting in the generation of enormous numbers of genotypes derived from each individual and family group.

Several computational methods are employed for linkage analysis each differing in the amount of family information required for the study. Thus the family groups used for linkage analysis range from large multigenerational kindreds to a smaller number of relatives, such as the affected sibling pairs method (31,36). Contemporary family pedigree linkage studies are thus very complex and require experts from a wide variety of disciplines in order to be successful (9). Model-dependent linkage analysis has been very successful in identifying the genes causing many diseases, such as cystic fibrosis, breast cancer susceptibility, and Huntington's disease. In this study design, clinical and genetic information on multigenerational families are studied under assumptions of monogenic autosomal recessive or autosomal dominant inheritance. The results are provided in a log of odds (LOD) score that signifies whether the genetic region is linked or unlinked to the disorder under study. Model-independent methods have also been developed that analyze the genotypes of pairs of family relatives, such as sibling pairs, or parent–child pairs. LOD expresses the chances that a marker is associated, or linked, to the phenotype under study. LOD scores of greater than three signifies that a particular marker has 1000 to 1 odds of occurring by chance alone and is generally used as evidence for linkage. A LOD score of negative 2 or less is generally accepted as evidence against linkage, and LOD scores between 3 and negative 2 is considered unclear for linkage (36).

The components of a linkage study are diagramed in Figure 1 and consist of family recruitment, molecular laboratory, and statistical analysis groups. The recruitment

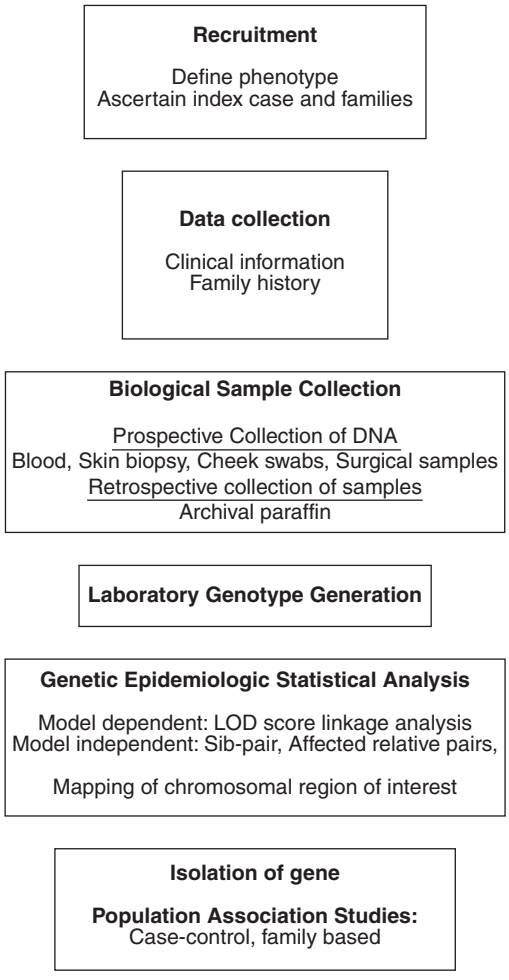

Figure 1. Components of a genetic pedigree or linkage study.

component includes experts in family contact, medical information retrieval, and sample collection. The laboratory component includes technical experts in sample processing and molecular analysis. Experts in genetic epidemiologic methods then carry out the linkage analysis in the statistical analysis component. All these components must be coordinated so that the medical, family, and genetic information can be used to identify disease-causing genes. The common feature of family pedigree linkage studies is the analysis of clinical information and blood samples from several family members in order to assess whether a genetic region is linked to the disease under study.

After a linkage study identifies a possible chromosomal location of a disease-causing gene, the gene is then isolated and identified with more refined genetic techniques and analysis, including positional cloning and sequencing (37). A full explanation of these techniques is beyond the scope of this chapter, except to note that genes are analyzed by a number of technologies. One method, called *DNA sequencing*, determines the exact order of the chemical building blocks, or nucleotides, of the gene. DNA sequencing also allows the researchers to identify genetic changes in an individual compared to the usual sequence found in the general population. Some of these changes are true deleterious mutations in the genetic code, in that the normal function of the gene is altered in the person carrying the mutation (37). However, other gene changes, called *polymorphisms*, may not have real functional significance. Polymorphisms in DNA sequences are very common and are thought to have little impact on human diseases.

Determining whether a gene change is a deleterious mutation or a polymorphism is a common problem for genetic researchers. If a potential polymorphism or mutation is found in a putative disease-causing gene, researchers use several techniques to determine the true effect on gene function. One of these techniques is to study multigenerational families with the disease to test whether the gene change correlates with the phenotype of affected individuals. A second method for researchers to study the phenotypic effect of a potential mutation is to perform molecular association studies examining the occurrence of the genetic marker in larger populations. The goal of these studies is to correlate the potential mutation in groups of unrelated affected and unaffected individuals (38). While most molecular epidemiologic studies do not use information about family relatives, these studies collect and analyze biologic samples for DNA analysis. Molecular epidemiologic studies illustrate the expanding role of genetic techniques in biomedical research, which can be used to provide information about disease-related risks in clinical practice. As knowledge about the human genome increases, future studies will focus on the interaction of several genes as well as the interaction of specific genes and environmental influences as necessary steps for disease development. Thus the methodology and technology of genetic research will become more complex in the future.

Accurate family data and information is the first necessary step in understanding human hereditary. As HGP nears its goal of determining the exact DNA sequence of the human genome, gene identification will be streamlined so that the causal genes will be identified at a faster rate in the future. Indeed, many genes will be identified without knowledge of the biological function or role so that researchers will continue to rely on family participation to understand the gene function in the population. Thus, with the advent of DNA markers for linkage analysis and genomic research, family studies have become the workhorse of modern genetic discovery. While future genetic research will continue to study rare monogenic disorders, the research focus will shift to include the study of genes that contribute to more common diseases, such as cancer, diabetes, cardiovascular disease, and aging. In this way family pedigree studies will then become more pervasive in the biomedical community (6).

POTENTIAL HARMS FOR SUBJECTS ENROLLED IN FAMILY PEDIGREE STUDIES

The wealth of information derived from genetic research is unprecedented and has given the biomedical community many new tools to understand and treat human disease. Genetic studies are no different from other forms of biomedical research, in which the benefits should outweigh

the potential harms to each volunteer. However, as Phillip Reilly notes, "that gene-discovery studies posed the threat of genetic discrimination — that there is a risk of *informational harm* associated with participating in studies that elicit genetic information from which one might infer health status" (39, p. 683). This "informational harm" is different from the usual physical harms that may result from participation in other types of biomedical research, such as clinical trials of new treatments or drugs. Informational harm from genetic research may cause emotional or social difficulties for research volunteers, including the impact of new genetic knowledge on health risks, family dynamics, and possible social stigmatization. Participants in genetic family pedigree studies are also at risk for unintended consequences of the research process. These unintended consequences can include misidentified parentage, learning private information about a family relative, having a research sample entered in additional studies without their knowledge or consent, or finding risk for additional diseases not part of the initial study (9,10,40).

Genetic Discrimination

Genetic discrimination is viewed by the general public as a major threat for individuals who enter genetic research studies (39). In particular, there is concern that genetic information obtained through the analysis of human tissue samples could be used to discriminate against individuals by insurers or employers (41). While the fear of stigmatization has generated a national debate on genetic privacy, it should be noted that few studies have scientifically examined this issue (14,42,43). In one of the first studies to document the occurrence of employment stigmatization and insurance abuse, Billings and colleagues reviewed 41 separate instances of discrimination submitted by genetic professional and patient advocacy groups (44). These authors defined genetic discrimination as "discrimination directed against an individual or family based solely on an apparent or perceived genetic variation from the 'normal' human genotype" and concluded that genetic discrimination is found in many social institutions (44). This report was controversial, as representatives of the health insurance industry argued that the number of discrimination reports were relatively small compared to the thousands of policies issued every year on a national basis (45). Other groups felt the reported instances of discrimination were primarily anecdotal and unsubstantiated by the authors, or that the cited examples did not conform to the authors proposed definition of genetic discrimination (46).

Nevertheless, the level of concern about potential discrimination is very high as shown by a 1997 study of over 1000 geneticist and primary care physicians, who reported over 550 instances of employment or life insurance refusal (43). This report echoes the sentiments of 332 members of the genetic support groups affiliated with the Alliance of Genetic Support Groups (47). This survey documented that up to 43 percent of respondents felt they experienced some form of discrimination by health insurers, life insurers, and employers, including refusal of life or health insurance or employment denial. There is also concern over genetic discrimination on an international level, as evidenced by a similar survey of genetic support groups in the United Kingdom. This survey found that one-third of the study respondents had difficulty when applying for life insurance compared with 5 percent of the control participants (48). The respondents who perceived themselves as suffering from discrimination reported that they experienced rate increases or outright refusals for insurance.

It must be remembered that the public's fear of genetic discrimination is not unfounded, as past abuses and stigmatization based on eugenics and physical disabilities are a matter of public record (1). In the United States these abuses ranged from the forced sterilization programs for persons with physical disabilities to the ill-advised sickle cell anemia screening program for African-Americans (15,49). Holtzman and Rothstein point out that the sentiments of the eugenic movement in the early part of this century still resonate in today's social institutions (50). In fact a 1998 U.S. government report estimated that 15 percent of employers plan to inquire into the genetic status of employment applicants (14). Thus the fear of marginalizing individuals on the basis of their unique genotype prompted policy review and recommendations on a federal level. Several governmental working groups and special commissions have been established to formulate specific policy agenda items relating to the social impact of genetic information (14,51). One of the first working groups was the Task Force on Genetic Information and Insurance sponsored by NIH and Department of Energy (42,51,52). NIH charged this group with studying the social implications of genetic discrimination by health insurance companies. Their 1993 report warned "people will be asked to provide information about their genetic risks to insurers" (51). The Task Force also noted that the risk of losing health insurance coverage for "preexisting" conditions may prevent people from obtaining predictive genetic information that could be used to improve the health and welfare of the person and family.

The Task Force concerns were based on the health insurance risks for people with a genetic condition or a family history of the disease, and most instances of documented discrimination have followed genetic diagnoses made in the clinical setting. However, some authors have suggested that study volunteers might also be required to disclose results from genetic studies in the research setting (53,54). Thus genetic information gained through voluntary participation in a family pedigree study might place the subject at risk for economic harm. This is an issue of *distributive justice* according to Thomas H. Murray, who writes that:

> Human genetics is, from this perspective, a science of human inequality. The principal ethical problem created by such scientific pronouncements of human difference is the task of reconciling such differences with our central moral, political and legal commitments to treating people as equals. That is, we must reconcile the ever-increasing evidences of human inequality with our vital commitment to moral equality (55, p. 80D).

For the research community it is essential to recognize the potential harms of genetic discrimination due to the scientific process. Researchers should develop appropriate study protocols to alert research participants to the possibility of informational harm. However, the magnitude of this risk is currently unknown, and may be part of an "urban myth" (56). A recent abstract presented to the plenary session of the 1999 ASHG national meeting found little evidence for genetic discrimination in a review of applications to 143 health insurance agents (57,58). In addition the study found no indication that insurers were using genetic information for health prediction and risk stratification. However, there have been no systematic surveys of the underwriting practices of life or disability insurance providers, hiring practices of employers, or services provided by social agencies such as housing or adoption. Thus the extent of insurance and employment discrimination based on genetic grounds is currently unknown. The research community must remain vigilant as to the potential economic and social harms to their study participants from inadvertent or premature disclosure of results from genetic studies.

Impact of Susceptibility Gene Identification

Pedigree linkage studies have been very successful in identifying genes responsible for human disease. Physicians are now able to use newly developed DNA tests to diagnose a suspected genetic condition, illustrating one of the benefits of genomic research. One important example in which a gene-based test has supplanted older biochemical testing is RET gene testing for multiple endocrine neoplasia type 2 (MEN2) (59). Gene carriers of this autosomal dominant condition have a 90 percent chance of developing medullary thyroid cancer, a potentially lethal cancer that can strike in childhood and early adulthood. Ninety-five percent of all cases of MEN2 are caused by deleterious mutations in the RET gene. A highly accurate DNA test for MEN2 is commercially available and has replaced the older biochemical calcitonin test used to diagnose this disease. DNA testing can be now offered to healthy at risk family members to determine if they also have a RET gene mutation. RET gene testing also illustrates the benefits of family pedigree research, in that genetic scientists used information from family pedigrees to identify a gene causing lethal thyroid cancer, paving the way for a simple medical test for the entire community. It should be remembered, however, that in order for such studies to be successful, many human volunteers donated their clinical information and blood samples to scientists studying MEN2. In addition volunteers helped the researchers to develop the clinical standards of care before the test could be provided to the general public.

While RET gene discovery is a clear example of the benefits of gene discovery, it is important to recognize that genetic tests are not absolute predictors of health or disease (40). This notion is in stark contrast to the commonly held view of genetic determinism, which suggests a person's genetic makeup is an unalterable blueprint for the future. Genetic determinism ignores the interaction of multiple genes, or environmental factors in developing disease (6,40,60). In addition proponents of this view neglect the underlying uncertainty about the health risks associated with susceptibility genes, in that the risks are more probabilistic in nature rather than an unequivocal link to the development of disease (6). The probabilistic health risks are true even for highly accurate gene tests, such as RET testing for MEN2 where most, but not all, gene carriers will develop thyroid cancer.

An additional reason that a person's genetic code can not be equated with a predetermined outcome comes from an understanding of the mechanism of action for deleterious genetic mutations. A deleterious mutation alters the expression of a single protein product produced by the cell. However, the cell may have several pathways that provide the same function, so that the effect of the deleterious mutation can be masked by the normal proteins encoded by other genes. Thus a one-to-one correspondence between gene mutation and phenotype, which is termed the *genotype–phenotype correlation*, is rarely seen for common conditions due to the complexity of the biological cellular pathways encoded by the genome. As such, disease-causing mutations will usually produce a recognizable phenotype. However, other gene changes, termed *susceptibility genes*, act to increase the propensity for the development of disease. Environmental factors are thought to interact with susceptibility genes to cause a specific disease. Thus the phenotypic effects of most genetic alterations falls somewhere on a continuum of risk for disease rather than an absolute cause of disease.

While categorizing mutations as disease-causing or susceptibility-causing is helpful in describing the potential effect of gene mutations, these labels are overly simplistic and thus are poor predictors for clinical disease. For example, individuals with neurofibromatosis type 1 (NF1) harbor a deleterious "disease-causing" copy of the NF1 gene on chromosome 17 (61). However, when the clinical phenotype of individuals with NF1 are compared, some exhibit the typical skin neurofibromas as teenagers, while others never develop this type of skin manifestation (62). This variation of clinical phenotype is known as *variable expression* of the gene and is thought to be a result of the interaction of the NF1 gene with unknown genetic or environmental factors. Nevertheless, the NF1 gene is fully penetrant. In other words, 100 percent of individuals who carry a deleterious NF1 gene mutation will exhibit symptoms of this disorder to some degree. However, because of variable expression, persons carrying a deleterious NF1 gene may exhibit very mild to severe health problems.

In contrast, other susceptibility gene mutations, such as mutations in BRCA1 and 2 breast cancer genes, will increase the propensity to develop breast, ovarian, or other cancers. Researchers estimate that between 50 and 85 percent of female BRCA1/2 gene carriers will develop breast cancer at some time in their lives, compared to general population risk of approximately 10 percent (63). In other words, since BRCA1/2 gene mutations are not fully penetrant, up to half of female gene carriers will remain cancer free and will be unaffected by their constitutional genotype. Other changes in genes, such as metabolic polymorphisms, are associated with minor functional effects of the protein product and can have

mild to moderate effect on the risk for disease. One well-known example is the CYP gene family, which is associated with differing ability to metabolize drugs and medications and other ingested compounds. Several studies have shown metabolic polymorphisms are linked with an increased rate of cancers due to the differing metabolic rates encoded by the different CYP genes (38,64). Most experts caution against the use of such polymorphisms for risk prediction, but the pharmaceutical industry has recognized that these polymorphisms will be important in drug development and have invested heavily in this area of pharmocogenetics (65).

Identifying disease-causing genes will continue to be enormously beneficial in the clinical setting. However, in the research setting, there is a growing recognition that gene identification can have harmful effects on participants and their families. Researchers may find themselves in a situation in which the very success of the project can have a negative impact on the participating individuals and family members. The major liability of pedigree studies stems from the inference of health states from the untested genetic discoveries. Volunteers may receive preliminary and unproven health information about a specific genetic marker, since the necessary follow-up studies on the function and clinical impact of the newly discovered gene have yet to be conducted. Thus there is a potential for research volunteers to receive preliminary, incomplete, and potentially harmful information from participating in pedigree research projects. In addition participants may never learn about important future developments in clinical trials or may learn about important health risks months to years after entering a study. It should be noted, however, that there is little evidence documenting harm to volunteers from genetic research at this time.

Volunteers for genetic pedigree studies are thus participating in the first step of scientific discovery in which there is a potential for finding disease-causing genes. However, since genetic mutations cause a wide range of phenotypic effects, researchers must use discretion when deciding whether to disclose preliminary study results to volunteers. In some cases the participant may wish to be informed about their genetic result prior to the completion of clinical trials. The researchers may also wish to share the experimental results with certain participants believing that more harm would come from withholding research results (6,66). Disclosing research results is a difficult dilemma, placing the researcher's duty to nonmaleficence in conflict with the subject's autonomy. The potential uses of genetic testing for disease prediction was the overriding concern of the researchers who released premature BRCA1 genetic results to a young woman intending to undergo a prophylactic mastectomy based on her family history of breast cancer (66). After being informed of her genetic research results, she found that she did not carry the family's high-risk gene mutation and was able to avoid prophylactic surgery. Before providing her with these results, the researchers weighed the potential psychosocial harms of releasing untested preliminary information with the harm of undergoing an unneeded medical procedure. The researchers were also concerned about alleviating the anxiety associated with genetic risks for disease. Anxiety and concern about health risks are well-recognized among family members at risk for Huntington's disease, Alzheimer's disease, and familial cancer syndromes (67).

Other researchers have cautioned that experimental findings from genetic studies should not be disclosed to the research subjects. The Children Cancer Group declined to disclose individual genetic research results to the physicians and parents of children enrolled in a genetic study of the p53 gene and childhood osteosarcoma (68). The p53 gene had been identified as the cause of cancer in families with the Li-Fraumeni syndrome in which multiple family members with osteosarcoma, leukemia, adrenal carcinomas, and other tumors (69). After careful consideration, this group chose to publish the genetic test results only in aggregate, as originally planned. They cited the lack of proven clinical utility of constitutional p53 gene mutations in nonfamilial cases, the potential for stigmatization and discrimination, and the subject's age as minors as factors in their decision. This group and other authors are concerned that predictive knowledge can be emotionally burdensome and can stem from knowledge of the family history as well as known carrier status from genetic testing (6,70). Thus the potential harm to research subjects from genetic knowledge cannot be neglected, and must be weighed when researchers are confronted with requests to divulge genetic research information.

Unintended Consequences of the Research Process

The problem of unexpected detection of new or secret information is also a concern for researchers who conduct family pedigree studies. The discovery of such knowledge by the researcher is an "unintended consequence" of the scientific process since this information is not related to the research goals of the study. This unexpected information can surface as a result of molecular analysis of the family blood samples and generally involves the inadvertent discovery of information that may not be known by all members in the family. For example, a serious, yet unintended, finding in genetic research is the discovery of mistaken parentage for a study volunteer, such as in nonpaternity or secret adoption. Since linkage analysis is dependent on distinguishing the maternal and paternal copies of each genetic marker, researchers can easily identify inconsistencies in the family when tracking the marker from parent to child. In this way the transmission of genetic markers is followed through subsequent generations and a family secret could potentially be discovered. While discrepancies in genetic markers can result from laboratory sample mix-up or other technical mistakes, a significant proportion is due to nonpaternity.

The rate of nonpaternity is estimated to be between 2.8 and 28 percent depending on the population group under study (71–73). While accurate figures are not known, it is not uncommon for researchers to confront this situation in family pedigree studies. Juengst outlines the dilemma of such a nonpaternity discovery for a research team studying a rare skin disorder (70,74). Subsequent publication of the pedigree revealed inconsistencies in the parental

genetic markers showing nonpaternity in two children in the kindred. Reilly also described a situation in which nonpaternity was discovered in a gene–mapping family study (75). Following discussion with outside consultants, the researchers decided that they would not disclose this sensitive information to the family. In addition this group chose to omit the details of parental inconsistencies in a subsequent publication (75).

Another unintended research finding is the discovery of a genetic mutation associated with a risk or susceptibility for a disorder that is distinct from the original focus of the study. In this way a person may be surprised to find that they are at risk for unforeseen health problems. In fact this type of incidental finding has occurred in genetic studies of common health problems, such as cardiovascular disease, as well as studies involving rare disorders, such as Hirschsprung's disease. Greely summarized this complicated aspect of genetic research by stating:

> One gene may be associated with multiple diseases. Therefore, a person who takes a genetic test to learn something about one disease may end up with information, possibly unwanted or harmful, about another disease (22).

Cardiovascular researchers have known for years that carriers of some forms of ApoE, a lipid transporting protein in the circulating blood stream, moderately increases the risk for heart disease. Many people underwent ApoE screening on a research and clinical basis to determine if they were at a greater risk for cardiovascular problems. In 1993, a surprising association was made between the APOE gene and the risk for Alzheimer's disease, where APOE4 carriers were two to three times as likely to develop dementia later in life. Individuals who accepted ApoE testing in the context of their cardiovascular health then discovered, perhaps unwillingly, information regarding their risk for Alzheimer's disease later in life (reviewed in Ref. 22).

A similar dilemma occurred when researchers attempted to isolate the genes involved in Hirschsprung's disease, or congenital megacolon. As early as 1982 families with Hirschsprung's disease and a rare form of medullary thyroid cancer were identified (76). Medullary thyroid cancer is one of the cardinal features of MEN2 and can develop in young children or in early adulthood (59). As previously discussed, the gene responsible for MEN2 is the RET gene located on human chromosome 10. In 1994 linkage studies based on families with multiple cases Hirschsprung's disease found that one of the genes causing this disease was localized to chromosome 10 at the same location as the RET gene (77,78). Further studies confirmed that mutations in the RET gene were also responsible for some familial cases of Hirschsprung's disease (79,80). Volunteers in Hirschsprung family studies were identified to be at risk for a lethal form of thyroid cancer in addition to the childhood form of megacolon. Thus the APOE and RET gene discoveries were complicated by the unexpected detection of an increased risk for more than one disease process, each with different health implications.

Entire subpopulations can also be identified to have unforeseen health risks as an unintended consequence of the research process. One example of this is the identification of three founder mutations in the BRCA1 and BRCA2 genes in the Ashkenazi Jewish families with breast and ovarian cancer. Population studies subsequently found that three specific mutations, 185delAG and 5382insC in BRCA1 and 6174delT in BRCA2, are carried by about 2 percent of all Ashkenazi Jews regardless of a history of cancer in the family (81). Female carriers of any one of these mutations have a 50 percent chance to develop breast cancer as well as a 20 percent chance to develop ovarian cancer over their lifetimes. The effect of this gene discovery on people with Ashkenazi Jewish ancestry has been immense, since an entire ethnic subpopulation learned that certain members are at a higher risk for cancer (42,82). Because of the high prevalence of these founder mutations, some feel that strong consideration should be given for genetic testing for breast cancer risk based on Ashkenazi Jewish background. Others feel stigmatized by this discovery. Thus, because of family pedigree research and gene discovery, the autonomy and decision making for individual members of this subgroup was subverted. The unexpected detection of disease-causing genes will become a prominent issue for genetic research, as gene discovery in identifiable populations will become more prevalent in the coming years (26).

Breeches of Confidentiality

Researchers using clinical and genetic material from human subjects have a duty to respect autonomy and privacy of each participant (14). As previously discussed, this becomes problematic when the context of the study is the family unit in that "family secrets" such nonpaternity or private difficulties can be inadvertently disclosed to other family members, co-investigators, personal physicians, or employers (6,83). Since genetic linkage studies are generally multifaceted (Fig. 1) and composed of several discrete areas of expertise, the possibility of inadvertent disclosure through the process of the study is a distinct possibility. For example, recruiters for genetic studies may ask, with the participant's permission, to gather medically related materials from the subject's personal physician and medical file. Researchers conducting a linkage study for colon cancer susceptibility genes learned that a subject's personal physician noted in the subject's medical chart that the patient was "in a genetic study for colon cancer" (84). This illustrates the dissemination of information about study volunteers due to the interaction with healthcare workers rather than from the results or knowledge directly gained from the study. Likewise the research team may include identifying information about the participants in reports and data analysis generated by the study, leading to the suggestion that only some investigators in the research project team should have access to sensitive materials.

Once several family members enter a family pedigree study, considerable care must be taken to ensure that medical or genetic information of one family member is not accidentally revealed to other relatives. Researchers must keep in mind that individual family members expect their information to remain confidential, since there may be " family secrets" shared with only a few members in the

kindred (9,83). Psychiatric research is one area of genetic research in which inadvertent disclosure is a concern as the clinical data collected on family members may include potentially stigmatizing information. Clinical information is not always freely shared among family members and can include the severity of psychiatric symptoms, alcohol or drug abuse, or criminal activity. Indeed, researchers may have to share background information about the proband when enrolling other family members in a family pedigree study. Thus some information may not remain private. Juengst notes, for example, that it is particularly difficult to approach a distantly related family member about enrolling in a psychiatric linkage study without sharing the fact that someone in the family is affected with a psychiatric disorder (6). In these cases Shore and colleagues feel that "Subjects will need to be informed, when agreeing to participate in genetic research, that their relatives may also be asked to participate as subjects. It should be made clear to psychiatric patients whose relatives will be contacted exactly what information about a subject will be provided to those relatives" (83).

Family registries and databases of genetic material are rapidly increasing in number. NBAC estimates that over 282 million specimens of human biological materials are currently stored in the United States, illustrating the magnitude of potential genetic information available to researchers (14). These databases contain sensitive genetic information that could be accessed by nonresearch team individuals, such as computer hackers, or social and governmental agencies requesting information. Simply removing names or social security numbers as identifiers may not guarantee that the data will remain anonymous or unidentifiable, since electronic databases are proliferating at a fast pace (85). Schulte and Sweeney point out: "Although the records of government-sponsored or funded studies will be maintained according to the Privacy Act of 1974 (P.L. 93-579), this does not ensure that records will never be disclosed" (86). For investigators accepting federal funds for specific research projects, the Privacy Act permits the release of identifiable research information in some circumstances. Particularly relevant for family pedigree studies, researchers may be obliged to respond to a court order seeking information that would be used to protect the health and safety of other persons. Thus the law can require that the investigator disclose confidential information (19,53,86).

In light of this potential ethical conflict between the duty to ensure the privacy and confidentiality of volunteers and a federally mandated court order to divulge sensitive information, Earley and Strong suggest that genetic researchers use a little known Certificate of Confidentiality (53,87). NIH established the certificate as a means to provide protection for federal research projects investigating the extent of illegal drug use in the 1970s. The 1974 and 1988 amendments expanded the certificate protection to cover other research including mental health research and genetics (87). It is important to note that the Certificate does not provide protection to individual participants in biomedical research. Rather, the Certificate protects the investigators from being compelled to disclose results to outside interests. The Certificate is issued to the principal investigator and provides protection for the life of the study. In most cases the Certificate can be extended beyond the funding cycle for the project. To date, there are no published studies addressing the efficacy of the Certificate of Confidentiality for protecting the privacy of participants in genetic pedigree, linkage, or biomedical studies. Thus it is unclear whether these Certificates will provide adequate protection for investigators or their research subjects.

Breeches of confidentiality and privacy can also occur through the publication of family pedigrees when the researchers report their results to peer-reviewed journals or to other investigators. A unique aspect of family pedigree studies is that each subject's clinical and genetic information is analyzed within the context of his or her family. The pedigree diagram is a valuable tool allowing an investigator to convey phenotypic and genotypic information while maintaining the biological relationship between each research subject in the family. The pedigree also includes such personal information such as age, gender, and birth order of the family members. Importantly, the pedigree diagram is a visual aid enabling the reader to quickly assess the mode of transmission of a disease gene, DNA marker, or trait (88). Recently concern for maintaining the anonymity of families in publications containing pedigree diagrams has been raised because of the comprehensive nature of the information (89,90). This concern has increased within the last few years partially due to the expanding interest and education of the lay public in the field of genetics. Increasing access to research articles via the Internet may also be a contributing factor (91).

In the past, journal editors have treated the publication of pedigrees similarly to the publication of traditional case studies, by simply withholding the names of the subjects depicted in the pedigree diagram (88). However, concerns have been raised that a pedigree diagram depicting the family structure with a description of the disease could pose a risk to the privacy and confidentiality of the participating family (88,90). Publication of such information may adversely impact the members of the kindred in several ways. Pedigree diagrams often contain medical and social information that is highly personal and may be not have been shared with other family members. As discussed earlier, instances of "paternal genotype inconsistencies" included in published pedigrees have caused repercussions within the family under study (6). Additionally pedigree diagrams can indicate medical illnesses, reproductive history, and adoption status of the family members (90). Publication can thus result in the disclosure of private information to the proband, other family members, or outside acquaintances. Furthermore, if genotypes are included in the pedigree diagram, information regarding disease status, carrier status, or disease susceptibility could be unintentionally communicated to the study participant and family members. In addition Byers and Ashkenas concluded that there is a remote possibility of discrimination through the inadvertent disclosure to third parties such as insurers or employers (89). Such disclosure directly violates the study volunteer's right to privacy, regardless of the level

of harm. However, it is unclear whether such disclosure to a third party is a substantial risk at this time, since no published studies have directly examined this question.

Future Use of Research Samples

By collecting medical and family information along with a blood sample, most family pedigree and linkage studies also serve as a repository of genetic material, in which most of the individuals are related in family units. The genetic material used in family pedigree research is usually DNA taken from the volunteer's white blood cells, but it can be DNA from immortalized cell lines, buccal swabs, or paraffin embedded tissue from surgical samples. In most cases the DNA samples are from living individuals who provided their consent to the investigators to use their sample when enrolled in the study. Researchers may also use genetic material from tissue from deceased family members, since adequate amounts of DNA can be removed from surgical samples stored in hospital pathology departments. Since DNA is very stable when properly stored, it is possible for scientist to use the DNA samples for ancillary studies, or for unrelated studies long after the primary study is completed.

Using research samples from previous genetic studies is a common practice, since family pedigree projects will often have unused genetic material at the completion of the initial study. Some investigators stress that such repositories are extremely valuable for future scientific discovery, in which the cost of replicating the collection would be prohibitively expensive (92). In addition others have vigorously protested limitations on studying DNA from archival and family DNA banks, fearing that scientific progress will be impeded if previously collected samples cannot be used in future projects (93–95). However, ethicists and consumer groups suggest that participants may not realize that a sample of their genetic material could be used in future research projects that study other disease processes. Participants may not understand that their sample will remain part of a larger collection of genetic material, since research samples are not routinely destroyed at the completion of the project. Thus the future use of biologic materials has become a controversial and contentious topic among researchers, clinicians, ethicists, and patient advocacy groups (14,16,94,96–98).

The fact that tissue collected for genetic research may be used in future studies complicates the researcher's ability to protect the participant's right to autonomy and privacy. For example, study participants in one survey raised concerns that the previously collected sample could later be included in a research project that the participant would not have supported, such as fetal research or cloning experiments (99). In addition the subsequent research projects may yield clinically relevant information from material from DNA banks or other tissue sources. In this case the participant would be unprepared for information about risk for other severe but preventable disease, since he or she would not have been aware of this possibility when enrolling in the initial study (6,83). Concerns have also been raised about the potential for misuse of genetic information and the possibility of lost health insurance or employment opportunities resulting from unanticipated genetic research on previously collected DNA (38,50,98).

In light of these concerns, it is important to note that two surveys have shown a high proportion of participants enrolled in genetic research studies are willing to have their sample used in future research projects. Lewis and colleagues reported a survey of 416 subjects enrolled in a colon cancer linkage study demonstrating a very small percentage (<3 percent) of respondents refused to allow their sample to be used for future unrelated genetic studies (99). These individuals were primarily concerned about privacy and confidentiality of their personal information if their sample was used in this fashion. Fifty-one percent of the remaining participants allowed their sample to be used if their confidentiality was maintained, while 46 percent indicated that they wanted the researchers to contact them to learn more about the subsequent research before permitting future use. A second survey of 263 subjects enrolled in three separate genetic protocols sponsored by NHGRI found similar results (100). Four percent refused further use of their DNA sample, while the overwhelming majority allowed their sample to be used either after being recontacted (73 percent) or after stripping personal identifiers from the sample (26 percent) (101). The results of these studies indicate that the vast majority of research participants are willing to allow their DNA sample to be used in future research, although potential discrimination is a concern for most participants.

The pace and complexity of human genomic research will continue to grow in the near future and benefit the entire community. However, numerous volunteers in family pedigree studies will be potentially exposed to genetic discrimination from the results or conduct of the study. These harms can simultaneously affect the individual volunteer and their family during the life of the study as well as in the future. The potential harms include confidentiality, genetic discrimination, clinical impact and unintended consequences of gene discovery, and the future use of DNA samples. Thus researchers need to address these issues as they conduct family pedigree studies.

PROFESSIONAL CHALLENGES FOR THE CONDUCT OF RESEARCH

While the exciting advances in molecular technology promise a better understanding of many common diseases plaguing humankind, a consensus among researchers, bioethicists, and legal experts on the ethical dilemmas posed by pedigree studies is noticeably absent (6,8,10,22). Several prominent genome scientists have openly called for the development of standardized policies for pedigree and genome research (3). Some of the proposed policies have been controversial, and may place additional burdens on the research team in the time and resources required to conduct the research project (101). Others have responded that the duty of the researcher is clear and that "ethical research is good quality research" (102). Since one of the inherent ethical conflicts in family pedigree research stems from the blending of basic bench science with the potential use of unvalidated tests for clinical health decisions, each research group should establish protocols

Table 2. Policy Areas for the Conduct of Genetic Family Studies

1. Recruitment and ascertainment
2. Privacy and confidentiality of medical, family, and genetic information
3. Disclosure of experimental results
4. Future use of DNA
5. Informed consent

that are specific for the genetic condition under study. There are five basic areas that the research community should address in developing such policies for genome research (Table 2) including recruitment, confidentiality, disclosure of results, future use of DNA samples, and informed consent.

Recruitment and Ascertainment

Identifying eligible participants who have the correct family structure for a specific study is one of the most important aspects of family pedigree research. Many recruitment issues are no different than other human subjects research, in that subjects should be informed about the goals of the study, what participation entails, and the risks and benefits of entering the study. The NIH Office for Protection from Research Risks (OPRR) published an Institutional Review Board Guidebook in 1993 listing the guidelines for protecting human research subjects (103). While the Common Rule as outlined in the Federal statutes clearly state that human subjects research must be approved by each institution's Institutional Review Board (IRB), it has only been recently that additional guidelines have been put in place for genetic pedigree studies (103, p. A58; 104). Recognizing the unique position of recruiters for genetic studies, the OPPR guidebook states that "The familial nature of the research cohorts involved in pedigree studies can pose challenges for ensuring that recruitment procedures are free of elements that unduly influence decisions to participate" (103).

Recruitment protocols for family-based studies are more complex that those used for standard epidemiologic research projects, since each family member, although biologically related, must be separately enrolled in the study. Thus it is important to recognize that the recruitment process can exert undue pressure on family members to enroll. Pressure can come from the research team, since the study actively searches for kindreds with multiple members affected with the disease under study as these families provide more genetic information to the project (6,9). Coercion can occur within the family unit, especially if a volunteer expects that his or her family will directly benefit from the results of a study. "Thus the pressure of compelling familial relationships may simply replace the researcher's influence in recruiting potential subjects" (6, p. 407). Some researchers suggest approaching families in large groups or through organizations like support groups and allowing interested members to contact the researcher if they wish to join the study. This may serve to increase the autonomy of some people's decision. However, as Juengst points out "...family members may actually feel less free to demur in large group settings, and lay-led support groups vary in expertise, understanding, and objectivity" (6, p. 407).

Consensus has not been reached on how IRBs should require researchers to protect the privacy interests of family members. OPPR suggests that researchers might collect only publicly available facts about family members, such as names and addresses, from the proband (103). Although Juengst acknowledges that obtaining detailed medical and family pedigree information from probands "...is a practice so traditional as to be ethically invisible within the community," he suggests that researchers should follow OPRR's suggestion to collect only publicly available information about family members from probands, and then "...convert this tree into a genetic pedigree by soliciting relevant health data from each relative directly" (6, p. 405). However, this recruitment practice will place additional burdens on the research team to gather this information within the research budget allocated to the project.

Pedigree studies will use many sources for referral, such as support groups, health care providers, clinics hospital databases and family members previously enrolled by the study group. Cohen and Wolpert describe several methods of family enrollment and caution that people may pressure their kin to enroll in the study (9). They also suggest that one family member must first inform other members of the family about the study and provide permission for the recruiters to separately reach each member. The researcher then contacts the family member to further discuss the project and offer participation. Other recruitment strategies in the future will be to use data from family registries that have been developed for research. One such Quebec registry for familial Alzheimer's disease developed a recruitment strategy that relied upon local health care professionals as well as leaflet advertisements in hospitals and clinics (105). This group also developed ethical guidelines for the conduct of the entire research project, incorporating protections for family members and incapacitated adults. Researchers must also determine whether minor children or incapacitated adults should be enrolled in a family study (6,9). Parents enrolling their child must use substituted judgment for the child and not be influenced by other family concerns. Thus specific protections should be in place to include the minor child's assent to join the study.

Ensuring and Maintaining Confidentiality

As previously discussed, researchers have a duty to uphold the privacy and confidentiality of participants in family pedigree research, including medical information and DNA genetic results. Pedigree studies pose additional challenges for the researcher to ensure that the data remains confidential, to prevent private information about some family members from being inadvertently disseminated within the extended pedigree, and to limit potential breeches in confidentiality through the process of publication (7,9,70). Standard approaches have been developed for coding and tracking health and genotype information for study enrollees to aid in managing confidential information (9). While there are

no federal laws that guarantee protection of genetic data, the Certificate of Confidentiality has been proposed to protect researchers from being compelled to submit genetic data to outside agencies (16,53). As previously noted, this protection is afforded to the researcher and not to the research participant. Again, there are no specific guidelines on this issue from OPPR or another regulatory agency, although patient advocacy groups have proposed that similar protections be developed for research subjects (19,96). The National Action Plan on Breast Cancer (NAPBC) focused on developing specific strategies to ensure privacy for participants in genetic studies (19). Stating that "privacy protections for experimental research data in which health care is not delivered should exceed the protections established for medical records" and recommended that identifiable genetic research data should not be included in a person's medical file.

Researchers should guard against providing identifiable data to the public at large through the publication of pedigrees. Representing the family medical history in a pedigree format is an essential part of data for publication but could possibly disclose the familial condition in an identifiable format. Thus a practice of altering pedigree information has been developed to provide anonymity for the family members, although it is debatable whether it affords true protection and may undermine the scientific validity of the study (70,89). OPRR recommends that written consent be obtained from participants as to the release of personal information (103). However, there may be no reason to assume that all family members depicted in the diagram had enrolled in the study (89). The International Committee of Medical Journal Editors (ICMJE) issued guidelines in 1995 for protecting the privacy of research subjects in scientific publications, and recommended that "Identifying details should be omitted if they are not essential, but patient data should never be altered or falsified in an attempt to attain anonymity" (88,106).

Disclosure

The genetic revolution has enabled researchers to locate disease-causing genes which as paved the way to the development of new genetic tests for clinical care. While these advances benefit the entire community, the individuals who donated their clinical and genetic material to family pedigree researchers may wish to know their personal study results. As previously discussed, these research results are experimental, and may not be clinically valid for health concerns. Indeed, the research genetic test result may be technically inaccurate, as these tests are performed in the research laboratory rather than in the clinical laboratory (15). NAPBC also points out that research laboratories have a higher tolerance level for inaccurate experimental results than do clinical laboratories (19). Thus falsely negative or falsely positive test results could be provided to a research volunteer.

Federal regulations in the form of the Clinical Laboratory Improvements Amendments (CLIA) of 1988 were enacted to codify the specific requirements for clinical laboratories providing test results that will be used in clinical management of patients (107). CLIA standards help ensure sample integrity and clinical validity of the test results. While the CLIA statute was developed for all type of laboratory tests, there are limitations in the requirements for testing and monitoring genetic testing (108). Nevertheless, these standards apply to all laboratories that test samples for clinical decision making including research laboratories that supply genetic tests for rare disorders at no cost (15). Thus researchers may be in violation of the CLIA statute when disclosing test results to their participants or to clinicians caring for the subject.

The Institute of Medicine report on genetic testing noted that research laboratories may offer the only available genetic tests for rare disorders, since it is impractical for general clinical laboratories to develop tests that would be infrequently used. The authors of this report recommended the establishment of a central repository and genetic CLIA approved laboratory to offer these tests to patients and family members. Other research groups, such as the newly formed International Gastric Cancer Linkage Consortium, have instituted specific protocols for offering genetic testing to research participants (109). This group recommended that clinical genetic counseling be offered to family members in which a mutation in the E-cadherin gene is found in research subjects. E-cadherin is cellular adhesion molecule and persons with constitutional deleterious mutations are at risk for an aggressive form of gastric cancer. In addition the consortium arranged to have the experimental research findings validated by a CLIA approved molecular laboratory. Thus individuals and their family members will be able to learn their research results while the risk of inaccurate test results are minimized. Several groups have developed guidelines as to the proper avenue of disclosure of research information to volunteers. For example, the ASHG strongly recommends that research results only be communicated "by persons able to provide genetic counseling" (13).

NBAC recognized that disclosing results to research subjects is controversial (14). In a 1999 report the commission recommended that disclosure should occur only where the findings are scientifically valid, have significant health implications for the subject and a treatment is available for the disorder in question (14). The authors of this report also assumed that disclosure would be a rare circumstance for the researcher, although this may not be true for researchers who identify highly prevalent genetic changes for a common disorder. Patient advocacy groups, such as NAPBC, feel that research participants should have access to experimental findings, except when the results have unproven clinical validity. They also recommended withholding research data when the results could harm the subject, interfere with the study, or cause harm to another individual (19). Mac Kay suggests that results should not be disclosed to research volunteers as "a more equitable way of dealing with the possibly conflicting views of family members as well as avoiding the problems of information whose reliability is not yet established" (7, p. 489). However, it is important to recognize that many subjects will be interested in their personal genetic results. Thus, researchers must specifically address whether experimental information

will be provided to their study volunteers and, if so, how disclosure will take place.

Future Use of Samples

One of the most difficult issues for genetic researchers is the development of an ethical framework for the future use of the DNA samples collected for the research study. Members of the research community have hotly debated specific guidelines for research using previously collected or archival tissues, since the future use of genetic material was not considered when many of these repositories were established (14). Most authors recommend contacting and obtaining the subject's consent for research on projects that require the use of specific personal identifiers (67). An alternative approach would be to use samples from retrospective DNA repositories as long as the sample is anonymized and stripped of all identifiers (102).

Several organizations and consumer groups have made recommendations for the future use of genetic samples. In a 1996 policy statement ASHG classified biological samples into one of four groups (16). First, " anonymous" biological samples are defined as samples that were originally collected without any specific identifiers from the person who donated the sample. Thus, linking an anonymous sample to the original source is impossible. Second, "anonymized" samples are defined as samples which were initially collected with specific identifiers, were subsequently stripped of all these identifiers. Anonymized samples are thus irreversibly removed from any link to their source, except that the samples can remain linked with clinical, pathological, and demographic information as long as the amount and type of this linked information does not breech anonymity. Third, "identifiable" samples are linked to sources by a confidential code developed by the original investigator. While a member of the original research team can decode these samples, the person's identity can not be revealed to persons outside of the study. Fourth, "identified" samples are those tissue samples associated with the participant's name, hospital number, or pedigree location and are available to the researchers. Hospital pathology departments are an example of repositories of identifiable tissue samples. A similar classification for research samples has been proposed by NBAC, where samples are categorized into unidentified, unlinked, coded and identified samples (14).

The ASHG report on informed consent for genetic research made several recommendations regarding future use of biological samples. Regarding permission to use the sample in other unspecified studies, this group stated "...It is inappropriate to ask a subject to grant blanket consent for all future unspecified genetic research projects on any disease or in any area if the samples are identifiable in those subsequent studies." This group also recommends that researchers ask the volunteer to "indicate if unused portions of the samples may be shared with other researchers." The report also recommended that the subject should indicate whether subsequent researchers should "receive their samples as anonymous or identifiable specimens" (16).

Several other groups have made recommendations for future use of research samples. In general, most groups agree with the OPRR guidelines calling for researchers to "obtaining consent from the participants for any use of the data (and samples) that is not strictly within the original uses to which the participants agree" (103). Clayton and colleagues made further recommendations, by suggesting that researchers inform the subjects "about the scope and potential consequences of the projects" (98). These authors also suggest that subjects should be asked if they are wish to have their sample anonymized, as well as if they would allow their sample to be used by investigators outside the institution or outside the original research project. The American College of Medical Genetics (ACMG) guidelines propose that researchers request permission for future use of the sample from the volunteer at the time the sample is collected (97). This group also recommended that researchers develop a specific policy about whether subjects will be recontacted if permission to use the sample in the future was not obtained at the time the sample was collected.

Informed Consent

The ethical complexities of genetic research studies prompted one ethicist to write that the "steps to obtaining consent from members of a pedigree can be tortuous" (7). However, the evolving duty of the genetic researchers has been recently described:

> In sum, the people whose genetic and clinical data will be essential for the next phase of human genomics research need to be treated not merely as "subjects" but more as (somewhat limited) partners. Researchers must realize that these people have interests beyond safety; ethicists must recognize that, when well informed, they have the right to participate even in broadly defined research. The goal of this approach is not to prevent research but to prevent research subjects from feeling cheated, powerless, misled, or betrayed (22).

Most researchers recognize that a trusting relationship with the study volunteers is essential for the success of the entire project. Indeed, it is through the informed consent process that such a relationship is first developed. In this way the informed consent process is more significant than the signed document detailing the proposed research and can be a blueprint for the ethical conduct of the research to be performed. Hence the study protocol should include detailed plan for informing the volunteer about the study and obtaining consent prior to enrollment.

Since family pedigree research involves the collection of clinical information and biological samples on multiple family numbers, these studies fall under the Common Rule, which requires that human subjects research supported by federal agencies be reviewed by an IRB (104). NBAC extended this recommendation for IRB oversight for all human subjects research, regardless of federal support (14). The ASHG statement on informed consent for genetic research encouraged researchers to develop procedures to obtain informed consent for both prospective and retrospective studies (16). In addition this report suggested that specific protocols be developed for maintaining confidentiality, disclosure of expected

Table 3. Elements of Informed Consent Document for Family Pedigree Studies

1. Purpose of study
2. Participation is voluntary
3. Costs and/or reimbursement
4. Benefits of participation
5. Disclosure of experimental results
7. Risks or informational harm from participating
8. Ensuring confidentiality: "Certificates of Confidentiality"
9. Future use of DNA sample

and unexpected experimental data, and deposition of samples.

The elements to be included in the investigator's informed consent document should follow standard formats with special additions relevant to genetic studies (Table 3). As with all human subject research, informed consent requires that the consent document be written in language understandable to average readers. It must also include a description of the project and the purpose of the research. The potential participant must be given the option of withdrawing from the research at anytime without penalty. The researcher must also identify any costs related to participating in the project and include an estimate of the amount of time. In addition to these points, informed consent documents must also contain information on benefits, disclosure of results, risks, confidentiality, and future use of samples. Researchers who are composing consent documents for genetic family pedigree studies must pay particular attention to these components to insure that participants are informed of the potential harms specific to genetic studies (103).

Researchers must clearly state that there may be no direct benefits for the participant or family members. However, as discussed previously, researchers must also indicate whether experimental results will be communicated to study volunteers. If the research team is planning to withhold results, the informed consent document should indicate that no research results will be given to the participant. On the other hand, if the research team is planning to share experimental results, the disclosure protocol should be explained to the participant in advance. In this case additional costs that may be incurred by the participant, such as CLIA laboratory confirmation or genetic counseling, should be included in the consent document. In addition to the physical risks that may be incurred in the genetic study the risks associated with "informational harms" must be disclosed (16,39). As previously discussed, these include potential psychological harms from learning preliminary health information directly from the study, or from new and unintended information which may not be related to the initial focus of the study. IRB guidelines state: "Prospective subjects should be informed during the consent process that the discovery of such information is possible" (103). Identification of nonpaternity or undisclosed adoption in a family is another unintended consequence of genetic studies that may cause psychological harm to participants. ASHG recommends that researchers consider including a statement in the informed consent document that mistaken parentage will not be disclosed (16).

A description of steps that will be taken by the researcher to protect the privacy of the study participants should be included in the informed consent document. Most volunteers will want to know that personal information or study results will not be disclosed to third parties, including employers, insurers, and family members, unless there is written consent from the participant. If the researchers follow the ASHG recommendations and obtain a Certificate of Confidentiality, a brief description of the protections it affords should be included in the consent form (16). The researcher should describe how confidentiality will be maintained if the research results or the family pedigree will be published. In regards to publication of family pedigrees, the ICMJE guidelines state: "Identifying information should not be published in written descriptions, photographs, or pedigrees unless the information is essential for scientific purposes and the patient (or parent or guardian) gives written informed consent for publication" (106). This requires the consent of all family members depicted in the pedigree, which can be a daunting task for researchers. Finally, researchers must clarify use of the samples in future research. Many groups have made recommendations about the use of genetic samples for future genetic studies, and most groups agree with the recommendation of the 1995 ASHG report that it is inappropriate to ask a participant to provide unrestricted consent to the future use of a sample when the risks of the future project are unknown (16).

Successful family pedigree studies are dependent upon the generous donation of clinical information, family information and biologic samples by volunteer participants. Recruitment of potential subjects for genetic studies requires a trusting relationship between the volunteer and researcher. Thus the protocol that a researcher develops for the informed consent process is a major component of the study; it will require policies addressing the potential benefits and risks specific to genetic studies. In order to continue the enormous success of the genetic revolution, researchers must develop guidelines that will ensure the ethical code of conduct for their genetic studies.

BIBLIOGRAPHY

1. G.J. Annas, *Emory Law J.* **39**, 630–664 (1990).
2. National Human Genome Research Institute in NIH, 1998, Available at: *http://www.nhgri.nih.gov/HGP/*
3. F.S. Collins et al., *Science* **282**(23), 682–689 (1998).
4. J. Watson, *Science* **248**(4951), 44–49 (1990).
5. J.D. Watson and F.H.C. Crick, *Nature* **171**(4356), 737–738 (1953).
6. E.T. Juengst, in H.Y. Vanderpool, ed., *The Ethics of Research Involving Human Subjects: Facing the Twenty-first Century*, University Publishing Group, Frederick, MD, 1996.
7. C.R. Mac Kay, *Hum. Genet. Theory* **4**(4), 477–495 (1993).
8. E.T. Juengst, *Am. J. Hum. Genet.* **54**(1), 121–128 (1994).
9. P. Cohen and C. Wolpert, in J.L. Ed Haines and M.A. Pericak-Vance, eds., *Approaches to Gene Mapping in*

Complex Human Disease, Wiley-Liss, New York, 1998, pp. 131–159.

10. N.A. Holtzman and L.B. Andrews, *Epidemiol. Rev.* **19**(1), 163–174 (1997).
11. H.T. Greeley, *Science* **282**(5389), 625 (1998).
12. A. Baum, A.L. Friedman, and S.G. Zakowski, *Health Psychol.* **16**(1), 8–19 (1997).
13. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines in the Conduct of Research Involving Human Subjects*, Department of Heath, Education and Welfare, U.S. Government Printing Office, Washington, DC, 1979.
14. National Bioethics Advisory Commission, *Research Involving Human Biological Materials: Ethical Issues and Policy Guidance*, vol. 1, U.S. Government Printing Office, Rockville, MD, 1999.
15. L.B. Andrews et al., eds., *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994.
16. American Society of Human Genetics, *Am. J. Hum. Genet.* **59**(2), 471–474 (1996).
17. American Society of Human Genetics Ad hoc Committee on Breast and Ovarian Cancer Screening, *Am. J. Hum. Genet.* **55**(5), i–iv (1994).
18. American Society of Human Genetics social Issues Subcommittee on Familial Disclosure, *Am. J. Hum. Genet.* **62**(2), 474–483 (1998).
19. B.P. Fuller et al., *Science* **285**(5432), 1359–1361 (1999).
20. Y. Miki et al., *Science* **266**(5182), 66–71 (1994).
21. R. Nowak, *Science* **266**(5190), 1470 (1994).
22. H.T. Greeley, *Genet. Test.* **3**(1), 115–119 (1999).
23. E. Kodish et al., *J. Am. Med. Assoc.* **279**(3), 179–181 (1998).
24. American Society of Clinical Oncology, *J. Clin. Oncol.* **14**(5), 1730–1736 (1996).
25. R. White and C.T. Caskey, in G.J. Annas and S. Elias, eds., *Gene Mapping: Using Law and Ethics and Guides*, Oxford University Press, New York, 1992, pp. 173–183.
26. M.W. Foster and W.L. Freeman, *Genome Res.* **8**(8), 755–757 (1998).
27. J.S. Risch, in D.L. Rimoin, M.J. Connor, and R.E. Pyeritz, eds., *Emery and Rimoin's Principles and Practice of Medical Genetics*, Churchill Livingstone, New York, 1997, pp. 371–382.
28. B. Hanson et al., *Am. J. Hum. Genet.* **48**(5), 873–879 (1991).
29. S.J. Winawer et al., *Cancer* **70**(5), 1236–1245 (1992).
30. S.J. Winawer et al., *N. Engl. J. Med.* **334**(2), 82–87 (1996).
31. R.C. Elston and J. Stewart, *Hum. Hered.* **21**, 523–542 (1971).
32. R.C. Go et al., *J. Nat. Cancer Inst.* **71**(3), 455–461 (1983).
33. M.C. King et al., *J. Nat. Cancer Inst.* **71**(3), 463–467 (1983).
34. B. Newman et al., *Proc. Nat. Acad. Sci.* **85**(9), 3044–3048 (1988).
35. J. Badner et al., *Am. J. Hum. Genet.* **46**(3), 568–580 (1990).
36. E.S. Lander and N.J. Schork, *Science* **265**(5181), 2037–2048 (1994).
37. A.L. Beaudet et al., in C.R. Scriver et al., eds., *The Metabolic and Molecular Bases of Inherited Disease*, McGraw-Hill, New York, 1995, pp. 53–228.
38. P.A. Schulte, G.P. Lomax, and E.M. Ward, *J. Occup. Environ. Med.* **41**(8), 639–646 (1999).
39. P.R. Reilly, *Am. J. Hum. Genet.* **63**(3), 682–685 (1998).
40. N.A. Holtzman and D. Shapiro, *Br. Med. J.* **316**(7134), 852–856 (1998).
41. Ad hoc Committee on Genetic Testing/Insurance Issues, *Am. J. Hum. Genet.* **55**(1), 327–331 (1995).
42. K.H. Rothenberg, *Health Matrix* **7**(1), 97–124 (1997).
43. J.C. Fletcher and D.C. Wertz, *Am. J. Hum. Genet.* **61**(4), A56 (1997).
44. P.R. Billings et al., *Am. J. Hum. Genet.* **50**(3), 476–482 (1992).
45. J.A. Lowden, *Am. J. Hum. Genet.* **51**(4), 901–903 (1992).
46. E.B. Hook, *Am. J. Hum. Genet.* **51**(4), 899–901 (1992).
47. E.V. Lapham, C. Koxma, and J.O. Weiss, *Science* **274**(5287), 621–624 (1996).
48. L. Low, S. King, and T. Wilkie, *Br. Med. J.* **317**(7173), 1632–1635 (1998).
49. P.S. Harper, *Am. J. Hum. Genet.* **50**(3), 460–464 (1992).
50. N.A. Holtzman and M.A. Rothstein, *Am. J. Hum. Genet.* **50**(3), 457–459 (1992).
51. NIH/DOE Task Force on Genetic Information and Insurance, *Genetic Information and Health Insurance. Report of the Task Force on Genetic Information and Insurance*, U.S. Government Printing Office, Washington, DC, 1993.
52. M.R. Natowicz, J.K. Alper, and J.S. Alper, *Am. J. Hum. Genet.* **50**(3), 465–475 (1992).
53. C.L. Earley and L.C. Strong, *Am. J. Hum. Genet.* **57**(3), 727–731 (1995).
54. T.H. Murray, *Hastings Center Rep.* **22**(6), 12–17 (1992).
55. T.H. Murray, *Am. J. Cardiol.* **72**(10), 80D–84D (1993).
56. J. Stephenson, *J. Am. Med. Assoc.* **282**(23), 2197–2198 (1999).
57. M.A. Hall and S.S. Rich, *Am. J. Hum. Genet.* **65A**(4), A3 (1999).
58. M.A. Hall and S.S. Rich, *Am. J. Hum. Genet.* **66**(1), 293–307 (2000).
59. G.L. Wiesner, *The Genetics of Multiple Endocrine Neoplasia, Type II in GeneClinics: Medical Genetics Knowledge Base*, University of Washington School of Medicine and National Institutes of Health, Washington, DC, 1999.
60. G.J. Annas, *N. Engl. J. Med.* **331**(5), 1027–1030 (1994).
61. R.M. Cawthon et al., *Cell* **62**(1), 193–201 (1990).
62. J.M. Friedman and P.H. Birch, *Am. J. Med. Genet.* **70**(2), 138–143 (1997).
63. D.F. Easton et al., *Am. J. Hum. Genet.* **56**, 356–371 (1995).
64. H. Raunio et al., *Gene* **159**(1), 113–121 (1995).
65. S. Lehrman, *Nature* **389**(6647), 107 (1997).
66. B.B. Biesecker et al., *J. Am. Med. Assoc.* **269**(15), 1970–1974 (1993).
67. P.S. Harper, *Br. Med. J.* **306**(6889), 1391–1394 (1993).
68. E. Kodish, T.H. Murray, and S. Shurin, *Clin. Res.* **42**(3), 396–402 (1994).
69. F.P. Li et al., *Cancer Res.* **48**(18), 5358–5362 (1988).
70. E.T. Juengst, *Genome Sci. Technol.* **1**(1), 21–36 (1995).
71. M.-G.L. Roux et al., *Lancet* **340**, 607 (1992).
72. S. Macintyre and A. Sooman, *Lancet* **338**, 869–871 (1991).
73. Parentage Testing Standards Program Unit, *Annual Report Summary for 1998*, vol. 4 (1998).
74. J.G. Compton et al., *Nat. Genet.* **1**(4), 301–305 (1992).
75. P.R. Reilly, M.F. Boshar, and S.H. Holtzman, *Nat. Genet.* **15**(1), 16–20 (1997).
76. M. Verdy et al., *J. Pediatr. Gastroenterol. Nutr.* **1**(4), 603–607 (1982).

77. S. Lyonnet et al., *Nat. Genet.* **4**(4), 346–350 (1993).
78. M. Angrist et al., *Nat. Genet.* **4**(4), 351–356 (1993).
79. P. Edery et al., *Nature* **367**(6461), 378–380 (1994).
80. G. Romeo et al., *Nature* **367**(6461), 377–378 (1994).
81. B.B. Roa et al., *Nat. Genet.* **14**(2), 185–187 (1996).
82. G.L. Wiesner, *Health Matrix* **7**(1), 3–30 (1997).
83. D. Shore et al., *Am. J. Med. Genet.* **48**, 17–21 (1993).
84. G.L. Wiesner, Personal communication, 1999.
85. J.J. McQueen, *Clin. Chem. Lab Med.* **36**(8), 545–549 (1998).
86. P.A. Schulte and M.H. Sweeney, *Environ. Health Perspect.* **103**(3), 69–74 (1995).
87. National Institutes of Health, *Privacy Protection for Research Subjects, Certificates of Confidentiality*, 1999. Available at: *http://grants.nih.gov/grants/oprr/human subjects/guidance/certconpriv.htm*
88. P.H. Byers and J. Ashkenas, *Am. J. Hum. Genet.* **63**(3), 678–681 (1998).
89. M. Powers, *Inst. Rev. Board* **15**(9), 7–11 (1993).
90. J.R. Botkin et al., *J. Am. Med. Assoc.* **279**(22), 1808–1812 (1998).
91. Editorial, *Nat. Genet.* **19**(3), 207–208 (1998).
92. E. Marshall, *Science* **282**, 2165 (1998).
93. W.W. Grody, *Diag. Mol. Pathol.* **4**(3), 155–157 (1995).
94. W.W. Grody and M.E. Sobel, *Diag. Mol. Pathol.* **5**(2), 79–80 (1996).
95. E. Marshall, *Science* **271**, 440 (1996).
96. Alliance of Genetic Support Groups, *Informed Consent: Participation in Genetic Research Studies*, Chevy Chase, MD, 1993.
97. American College of Medical Genetics Storage of Genetics Materials Committee, *Am. J. Hum. Genet.* **57**(6), 1499–1500 (1995).
98. E.W. Clayton et al., *J. Am. Med. Assoc.* **274**(22), 1786–1792 (1995).
99. S. Lewis et al., *Am. J. Hum. Genet.* **63**(4), 28 (1998).
100. L.A. Middelton et al., *Am. J. Hum. Genet.* **61**(4), A58 (1997).
101. M. Bondy and C. Mastromarino, *Ann. Epidemiol.* **7**(5), 363–366 (1997).
102. P. Hainaut and K. Vahakangas, in W. Ryder, ed., *Metabolic Polymorphisms and Susceptibility to Cancer*, IARC Science Publication, Lyon, France, 1999, pp. 395–402.
103. (Human Genetic Research) OPRR, *Protecting Human Research Subjects: Institute Review Board Guide Book*, U.S. Government Printing Office, Bethesda, MD, 1993, pp. 5–49.
104. 56 Federal Register, *Code of Federal Regulations (CFR) 45* (1999).
105. B. Godard et al., *Alzheimer Dis. Assoc. Disorders* **8**(2), 79–93 (1994).
106. International Committee of Medical Journal Editors, *Br. Med. J.* **311**, 1272 (1995).
107. M.K. Schwartz, *Clin. Chem.* **45**(5), 739–745 (1999).
108. N.A. Holtzman and M.S. Watson, eds., *Promoting Safe and Effective Genetic Testing in the United States. Final Report of the Task Force on Genetic Testing*, U.S. Government Printing Office, Washington, DC, 1997.
109. C. Caldas et al., *J. Med. Genet.* **36**, 873–880 (1999).

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, ETHICS, INFORMED CONSENT IN RESEARCH

NANCY E. KASS
John Hopkins University
Baltimore, Maryland

OUTLINE

INTRODUCTION

It is only in the latter half of this century that significant intellectual and regulatory attention has been devoted to human subjects protection. In response to isolated, but horrific, examples of unethical research studies, codes of ethics for human subjects research were developed, regulations were passed, and standards of informed consent were established. And yet as often is the case, while this branch of ethics has made tremendous progress in just a few decades, new challenges continue to emerge. It is the purpose of this entry to provide a brief history of human subjects protections, to describe specific elements of informed consent as they apply to research, and to discuss specific examples in which upholding standards of informed consent remains particularly challenging.

HISTORY OF HUMAN SUBJECTS PROTECTIONS

The Nuremberg Code of 1948 usually is considered the first code of research ethics (1). This Code grew out of the Nazi war crime tribunals, during which descriptions had been revealed of experiments conducted on concentration camp prisoners. These experiments, conducted through force and coercion, studied such questions as how long humans can be immersed in ice water before dying of hypothermia, the effects of ingesting poisons, and the effects of being injected with viruses. The Nuremberg Code, intended to guide all future research with humans, was developed as part of the judgment in *United States v. Karl Brandt* et al. (1a). Given the context out of which the Nuremberg Code emerged, it is not surprising that its first stipulation

is that "the voluntary consent of the human subject is absolutely essential" (1, p. 181). It elaborated that the subject should have the legal capacity to consent, should be able to exercise free choice without any coercion or deceit, and should have sufficient knowledge and understanding of the experiment to enable an "enlightened decision." Ensuring that the consent is voluntary and informed is the "personal responsibility" of the investigator. The Code further states that the experiment should be expected to yield important results for society that cannot be obtained through other methods, should be based on previous animal research and a knowledge of the problem, should avoid all unnecessary physical and mental suffering and injury, should not be conducted when there is a priori reason to believe death or disabling injury will result, should not involve a level of risk that exceeds the importance of the problem, should be conducted only by qualified persons, should guarantee that the subject has the right to stop participating at any point, and should be terminated early should there be reason to believe that the experiment is unduly risky.

Shortly afterward, in 1953, the clinical center of the National Institutes of Health (NIH) opened. The NIH clinical center is a research hospital, funded by the federal government, where all "care" provided to patients is part of a research protocol. When the clinical center opened, the NIH decided to require informed consent of healthy volunteers who entered studies, but not of patient-subjects, who were presumed to have reason to want to participate in clinical research (2). Nonetheless, this was one of the first times that an entity of the U.S. government required informed consent for any type of human subjects research.

The first code regarding the ethical conduct of research put forth by a professional medical body was from the World Medical Assembly in 1964 (3). The Declaration of Helsinki, as it was called, echoed many of the tenets of the Nuremberg Code but included additional elements of relevance to doctors who conduct research. It reminds doctors that despite the importance of conducting research, "the health of my patient will be my first consideration." It adds that the responsibility for the welfare of the subject always rests with the investigator and not with the subject, despite the subject having given voluntary consent. Further, "concern for the interests of the subject must always prevail over the interest of science and society." The Declaration further states that experiments not conducted in accordance with the proposed ethical requirements should not be considered publishable regardless of their scientific findings, and that the doctor should be particularly "cautious" if the subject is in a dependent relationship with him or her, in which case a different member of the research staff should obtain consent from the subject.

In the 1960s and 1970s several events in this country brought attention to human subjects research and the abuses that potentially can be associated with it. The Willowbrook hepatitis study was conducted from 1956 to 1970. The Willowbrook School, where this study occurred, is an institution for mentally retarded children. There were poor sanitary conditions at the school, and most children contracted hepatitis A at some point after being sent to live there. Researchers wanted to study the natural history of hepatitis A and the possibility of creating a vaccine for the disease. They decided to inject children who were newly admitted to Willowbrook with the strain of hepatitis that was rampant there. The "study" was justified by saying that the children probably would have become infected anyway and that, scientifically, more can be learned about the natural history of the disease if it is known precisely when the child became infected.

In 1963 the public became aware of experiments conducted at the Jewish Chronic Disease Hospital in New York. The Jewish Chronic Disease Hospital was an institution for elderly, chronically ill adults. In the experiments, residents of the hospital were injected with live cancer cells without their knowledge or consent. The experiments were justified by arguing that these patients would have died soon anyway.

In 1966 Henry Beecher, a well-respected Harvard physician, published in *The New England Journal of Medicine* an article that has become one of the classic pieces in research ethics (4). Beecher, who earlier had called for "a long, straight look at our current practices" (5), now conducted a review of articles published in top medical research journals. In the article he described 22 articles gleaned from medical literature of the time in which ethically questionable practices had been involved. Among his examples were placebos being substituted for an established treatment without patient-subjects' knowledge, studies of vulnerable subjects, and/or studies with a high degree or risk relative to benefit. Dr. Beecher's article received considerable attention, in great part, because his examples were drawn from *published* and therefore well-sanctioned research, and also because he implied that the research described was not necessarily unrepresentative nor unusual.

In response to this series of events, the United States Public Health Service established guidelines for research in 1966 (6). These guidelines required that each institution conducting human subjects research funded by the U.S. Public Health Service establish an Institutional Review Board (IRB) that would review projects in advance. The IRB would determine whether (*1*) the rights and welfare of study subjects are protected, (*2*) the methods to obtain informed consent are appropriate, and (*3*) the risks and potential benefits of the investigation are clear, and the potential benefits outweigh the risks (6). Therefore research review was to be prospective decentralized, based at the researchers' institution, and required to include informed consent.

In 1971 the public's attention turned to yet another horrific example in the history of research ethics. The Tuskegee syphilis experiment had been conducted by the American government from 1932 to 1972 (7). In this example, the Public Health Service was studying the natural history of syphilis. They chose as their subjects poor, black men from the rural south, 400 of whom had syphilis and 200 of whom served as controls. None of the men were told that a study was being conducted. Rather, they were led to believe that the government doctors were providing medical care for them, something

that poor, rural men were eager to find. When certain diagnostic procedures (e.g., spinal taps) were conducted for research purposes, the men were told that they were receiving treatment, and when efficacious antibiotic treatment became available in the 1940s, these men were denied therapy. The justification given for the study was that researchers simply were observing the disease that men already had, and since most of these men had no access to care, not providing care was no worse than what they already would have experienced. This study violated all ethical requirements of research, in that the men were not told they were participating in research and consequently could not provide meaningful consent. The research itself was unreasonably risky, particularly after penicillin became available; and the research singled out certain segments of the populations who were the poorest, were from a racial minority, and experienced none of the study's benefits. After press reports in the early 1970s exposed the horrors of the Tuskegee study, the Department of Health, Educations, and Welfare (DHEW) appointed the Tuskegee Study Ad hoc Panel to review the study, as well as to review the Department's policies for the conduct of human subjects research (8). The panel noted that, although DHEW guidelines had been in place since 1966, it was a journalist, rather than a review committee, that brought the conduct of this study to light. (for further discussion, see Ref. 9). The panel recommended that the Tuskegee study be stopped immediately, and also that a permanent body to regulate human subjects research be established by Congress. While such a proposal was introduced before Congress (9). it was not successful. However, two other responses to this decade of research exposes were successful: New regulations were promulgated by DHEW designed to build upon and strengthen the 1966 guidelines, and the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (National Commission) was established (10).

The National Commission, in existence from 1974 to 1978, was charged by Congress to investigate the ethics of human subjects research, particularly research with vulnerable populations, such as prisoners, children, and mentally disabled adults. In addition to producing 17 separate reports on research with each of these populations, the National Commission created The Belmont Report (11). *The Belmont Report* laid out three principles of bioethics to help examine the ethics of any research endeavor that involves human subjects, principles that remain extremely influential in contemporary bioethics: beneficence, respect for persons, and justice (for further discussion of these principles, see Ref. 12). Briefly, beneficence requires us to look out for the welfare of others. In the context of research, this means that harms to potential subjects must be minimized, and balancing of harms and benefits must occur. Respect for persons requires us to treat individuals as autonomous agents and, when persons have diminished capacity, requires us to protect them from harm. Therefore this principle requires us to respect the decisions and judgments of others, even if we may disagree with them. It is out of the principle of respect for persons that we are required to engage in a process of informed consent with research subjects. Justice requires us to be fair in the distribution of research burdens and benefits. Because of justice, we cannot disproportionately target vulnerable populations for enrollment in research to bear its risks, nor can we allow only those who are well-to-do and sophisticated to reap the benefits of research participation.

The 1974 DHEW regulations in many ways formalized the 1966 PHS guidelines that had required IRB review at local research institutions. The regulations went further, however, by delineating the elements that must be included in the informed consent process conducted with research subjects. These elements will be described below. All institutions that receive federal funding remain subject to these regulations. In 1975 the original regulations (Subpart A) were supplemented by another set of regulations (Subpart B) pertaining to research with pregnant women and fetuses. In 1978 Subpart C was added, providing regulations for the conduct of research with prisoners, and in 1983 Subpart D was added to provide oversight for research with children. In 1991 fifteen other federal agencies adopted Subpart A to apply also to their own research, calling it "The Common Rule." *The Common Rule* is in existence today, providing a common set of regulations for almost all federal agencies that either sponsor or conduct human subjects research in the United States.

Clearly, attention to the adequacy of human subjects protections usually has occurred in the context of a specific example that raised concern. The most recent such instance prompted the creation of the President's Advisory Committee on Human Radiation Experiments (ACHRE). ACHRE was formed in 1994 in response to allegations that radiation-related research was conducted on Americans between the 1940s and 1970s (during the cold war) without the participants' knowledge or consent. The Advisory Committee investigated these allegations and their extensive surrounding history. Moreover ACHRE did work examining the ethics of contemporary human subjects research, understanding that accusations of past abuses would raise questions in the minds of Americans about how much trust ought be put in current research practices and to establish a basis in fact on which to make recommendations for change in the future. Among ACHRE's recommendations were that research ethics training should be required of all research students and trainees, and that competency in research ethics should be required of all individual and institutional federal research grant recipients (9). Further ACHRE recommended that IRBs develop mechanisms to allocate their time more appropriately to riskier and more complex research, that information provided to patient-subjects clearly distinguish research from treatment, and not overestimate potential benefits. ACHRE also recommended that oversight of research be improved and that sanctions be created for those who do not comply with federal regulations (9, pp. 524–526).

Mention also should be given to the code that addresses specifically the conduct of human subjects research in the international setting. Put forth by the Council for International Organizations in the Medical Sciences

(CIOMS), the CIOMS guidelines were established in 1982. Among its provisions were that community consent, while often appropriate and necessary to obtain, cannot serve as a substitute for individual consent, and that research must be responsive to the health needs of the community in which the research occurs.

ELEMENTS OF INFORMED CONSENT

This section will describe both the specific elements of informed consent as delineated in the Common Rule and also will provide broader discussion of concepts inherent to the theory of informed consent. Informed consent has a history in medical practice that precedes but clearly influences its history in research. The clinical history starts with a series of cases at the beginning of the twentieth Century brought by patients who had not given their *consent* to certain procedures. Later in the century are cases brought by patients for not having been adequately *informed* about the procedures to which they were providing consent. In 1906, in *Pratt v. Davis*, a doctor performed a hysterectomy on a woman without her consent (13). The defense had been that when a patient enlists a doctor or surgeon's services, the doctor is given "implied license to do whatever in the exercise of his judgment may be necessary." The defense was rejected, and the case was decided in the patient's favor. Perhaps the most famous case on consent was *Schloendorff v. Society of New York Hospitals* in 1914 (14). In this case the patient had given consent for exploratory abdominal surgery but had specifically requested no further surgery. The surgeon, upon finding a fibroid tumor during the surgery, had gone ahead and removed it. One of the justices ruling on the case, Judge Benjamin Cordozo, wrote in his historic opinion, "Every human being of adult years and sound mind has a right to determine what shall be done with his own body; and a surgeon who performs an operation without his patient's consent commits an assault, for which he is liable in damages." The next fifty years brought a series of cases that led to a new requirement of patients also being informed. A landmark case, perhaps because it coined the phrase "informed consent," was in 1957, *Salgo v. Leland Stanford Jr. University Board of Trustees* (15). In this case Martin Salgo had undergone translumbar aortography which resulted in permanent paralysis, a known potential risk of the surgery. Mr. Salgo sued physicians for failing to warn him that a potential risk of the procedure was paralysis. The court found, in the patient's favor, that physicians had a duty to disclose "any facts which are necessary to form the basis of an intelligent consent by the patient to proposed treatment." This evolved into a requirement that all issues that would be pertinent to a patient when making a decision — such as the nature, consequences, risks, benefits, and alternatives to a proposed treatment — be disclosed before a patient makes a decision.

Extrapolating to the research context, informed consent requires both informing the research participant about the research and obtaining the participant's consent. *Informing* a participant requires disclosing pertinent information and ensuring at least some threshold level of understanding. Of course, fulfilling the former is considerably easier than fulfilling the latter, and consequently significantly more attention in the literature and the regulations exist concerning disclosure. For *consent* to occur, the participant must be competent and must make the decision voluntarily. The concepts of disclosure, understanding, competence, and voluntariness will be discussed below.

Disclosure

While understanding is what ultimately is required for valid informed consent, it is typically through disclosure that a research subject learns enough about the research project to understand it. If, indeed, a research subject were to understand the research other than through disclosure (e.g., through a video tape or from prior familiarity with the research), then substantially less disclosure would be necessary. Most discussions of informed consent, however, rightly assume that subjects know little about the research before they enter into a research relationship. Consequently the regulations governing human subjects research (the Common Rule) lay out in detail the elements of informed consent that must be disclosed (16):

1. A statement that the study involves research, an explanation of the purposes of the research and the expected duration of the subject's participation, a description of the procedures to be followed, and identification of any procedures that are experimental.
2. A description of any reasonably foreseeable risks or discomforts to the subject.
3. A description of any benefits to the subject or to others which may reasonably be expected from the research.
4. A disclosure of appropriate alternative procedures or courses of treatment, if any, that might be advantageous to the subject.
5. A statement describing the extent, if any, to which confidentiality of records identifying the subject will be maintained.
6. For research involving more than minimal risk, an explanation as to whether any compensation and an explanation as to whether any medical treatments are available if injury occurs, and if so, what they consist of, or where further information may be obtained.
7. An explanation of whom to contact for answers to pertinent questions about the research and research subjects' rights, and whom to contact in the event of a research-related injury to the subject.
8. A statement that participation is voluntary, refusal to participate will involve no penalty or loss of benefits to which the subject is otherwise entitled, and the subject may discontinue participation at any time without penalty or loss of benefits to which the subject is otherwise entitled.

Ruth Faden and Tom Beauchamp, in their book on informed consent (17), discuss that potential subjects want

to know information that is *material* to them. The federal regulations are based on assumptions about what would be material to most persons when considering participation. Faden and Beauchamp remind us, however, that certain additional details about a project may be important for a given individual, whereas they may be irrelevant to others. For this reason it is important that in addition to providing the basic elements of disclosure, there also be the opportunity for an informed consent discussion in which potential subjects can raise additional questions. Faden and Beauchamp note that information that is material is not necessarily information that is required for the *decision*, but rather it may be important for the potential subject to feel s/he has a good understanding of the research. For example, it may be important to a potential subject to know whether study hours can be arranged at night, even if the potential subject knows that s/he will participate regardless of the answer.

Understanding

It has been written that "assent to [research] given by a [subject] who actually has not understood disclosed information is not valid authorization" (18, p. 59). This reminds us that while adequate disclosure in most instances is necessary for there to be understanding, it in no way guarantees it. Many factors can contribute to inadequate understanding of research. Investigators may use jargon or language that is difficult for potential subjects to understand (19–21). Subjects who also are patients may be anxious about their medical condition and unable to focus on the specific information provided. Information also might be provided quickly and in large amounts, with little or no time for discussion, such that subjects simply cannot remember or integrate all they were told. Sometimes the words used in research are qualitative and vague. Nakao and Axelrod (22) as well as Fetting et al. (23) found imprecision in many terms used to describe research risks and benefits (e.g., "rare," "infrequent"), and recommend that numeric estimates be used when possible. In their review Silva and Sorrell report many factors that influence comprehension of research information (24). Several studies report that often too much information is given, and that subjects better understand the research when smaller amounts of information are provided. Who delivers information can also be relevant. Muss et al. found that more information about chemotherapy side effects was retained when risks and benefits were described by personnel other than the doctor (25). Not surprising, higher education is associated with greater comprehension of informed consent information (25). Understanding also is unavoidably compromised by lack of experience with a situation. That is, it can be difficult for subjects to imagine how they would react to a certain side effect until they have experienced it. While not a guarantee of understanding by any means, most commentators recommend using some measure to assess subjects' understanding of what they have been told before embarking on the research. This includes not simply seeing whether they can recall what are the study procedures but, at least as relevant, if they can recall what the *purpose* of the research is.

One area where understanding about research may be particularly challenging is clinical research. Here patient-subjects often enroll because of a hope that the research will be of medical benefit to them personally. This can result in research subjects losing sight of the fact that the study is investigational, the intervention has not been shown to be a valid medical treatment, and the purpose of the activity is not primarily to treat their own medical condition. Given that many clinical research investigators also are physicians, it can be easy for patient-subjects confuse clinical research participation with medical care. Studies have demonstrated that some proportion of patients enrolled in research do not understand, or at least, do not remember, that they are enrolled in research. Riecken and Ravich found that 28 percent of patients enrolled in research studies through Veterans Administration hospitals were not aware of their participation in research, despite having signed consent documents, remembering that the intervention had been explained to them and believing that they had been given sufficient information (26). The rate of unawareness dropped the longer patients were enrolled in studies, and patients were more likely to be unaware of their participation if a staff member other than the investigator had explained the study. Penman et al. similarly found that nearly one-fourth of patients receiving investigational chemotherapy did not recall that it was investigational (27).

Yet more subtle clouding of the boundary between treatment and research is demonstrated by many patients who are perfectly aware that they are participating in research but nonetheless view the research as a treatment intervention to improve their underlying disease state. This clouding is critical to issues of informed consent, since a key tenet of informed consent is understanding by the patient (17,18), and a key element of understanding in clinical research is appreciating what is meant by investigational.

Some years ago Appelbaum et al. described this threat to understanding as the "therapeutic misconception" (28). Appelbaum et al. report a psychiatrist approaching a patient to consider participating in a research study. The patient responds, "Yes, I'm willing to do anything that might help me." The patient further says that he understands what is meant by the trial's randomized placebo-controlled design but then goes on to say that he believes *he* will receive the study medication most likely to help him.

Competence

In order for informed consent to be valid, the person providing consent must be considered competent to consent. When a person is *not* competent to consent, a surrogate must consent on the subject's behalf, and IRBs usually engage in a higher level of oversight. That is, there is even greater scrutiny of the risk–benefit ratio, with an assumption that individuals who cannot understand the research in which they will be participating cannot be subject to as much risk as could a person who fully understands. A key principle in discussion of competence here is that individuals are not necessarily uniformly

competent or uniformly incompetent. Rather, the relevant question is whether the individual is competent to understand what is required of him or her by participating, and whether the individual is competent to exercise autonomous decision-making capacity in consenting to enroll. Individuals, for example, may be able to understand that they are being asked to be in research and what would be required of them, while having no recollection of what day of the week it is or who is President. Conversely, some individuals may function well in day-to-day activities but have no understanding of who doctors or researchers are, and cannot comprehend an informed consent discussion. Discussion concerning three specific populations whose full consent often is impaired — children, persons in emergency situations, and persons of limited decision-making capacity — will be provided below.

Voluntariness

Faden and Beauchamp write that "a fundamental condition of personal autonomy is that actions ... are free of ... controls on the person" (17, p. 256). Persons may be *influenced* in their thinking (e.g., by family members, information, or doctors), which ethically is quite consistent with valid informed consent, in contrast to being *controlled* by others (e.g., they are being forced or coerced), whereby they are not acting autonomously. In research, potential subjects may be most likely to feel "controlled" by others if they believe that other opportunities are dependent on their participation. For example, if a patient believes that her doctor will not treat her as well if she refuses to participate, then her decision cannot be considered voluntary. Similarly, if parole comes more quickly to prisoners who agree to participate in research, their decisions about participation may not be fully voluntary. All research consent discussions and forms must therefore emphasize that access to other opportunities will not be affected by potential subjects' decisions about participation. IRBs may decide that certain recruitment strategies are unacceptable because they would raise questions of compromised voluntariness. For example, professors may be told that they cannot solicit research participation from their own graduate students, or physicians from their own patients; in the latter case another physician can seek consent from patients, but the patient's own doctor may be perceived as exerting too strong an influence. Generally, the law "has long recognized a consent or refusal coerced by threats or manipulated by misrepresentation [to be] invalid" (12, p. 163).

CHALLENGES TO INFORMED CONSENT: SPECIAL POPULATIONS

Children

Tension exists when conducting research with children. On the one hand, children are assumed to be unable to fully appreciate the consequences of their actions, and therefore are not considered to be fully autonomous for the purposes of consenting to research. On the other hand, there are conditions that uniquely affect children, or affect children in ways that are different from how they affect adults, and therefore the medical or psychological care of children cannot be improved without research. Concerns based in justice dictate both that children not be used in research when fully competent adults could provide the same answers, yet also that children as a class not be denied the benefits of research knowledge. That is, children as a class ultimately are harmed if they are treated with drugs that never have been tested properly for safety, efficacy, or dosing in children. Consequently children can be used in research ethically only when the research question relates uniquely to them or if it is not greater than minimal risk research. Further, given that children cannot fully consent, we require a higher standard of beneficence. It is assumed that more protection for their welfare should be provided, with IRBs being more paternalistic than they are with competent adults. When children are approached for participation, it is required both that their parent or guardian sign written permission and that the child "assent" at an age-appropriate level. That is, researchers are required to explain certain relevant pieces of the research to the child in language the child will understand to see if the child is willing to participate. This assent undoubtedly will not include all information about the research. It may be as simple as saying to the child, "to learn more about your health, we would like to take some blood from your arm and ask you some questions. We're asking all the kids who come into the clinic today to do this. It's OK if you decide you don't want us to do this. Is it OK with you for us to do this?" In certain instances where the research in expected to be beneficial for the child, parents may overrule a child's lack of assent.

The Departments of Health and Human Services (DHHS) regulations regarding human subjects research were amended in 1983 to include Subpart D, "Additional Protections for Children Involved as Subjects in Research." The regulations define assent as a child's "affirmative agreement to participate in research. Mere failure to object should not, absent affirmative agreement, be construed as an assent" (29). The regulations say that minimal risk research with children is ethically acceptable, assuming that permission from parents and assent of child have been obtained. Research involving greater than minimal risk is allowed only if the risk is justified by anticipated benefit, if the benefit anticipated is at least as great as that offered by alternatives to participation, and if assent and permission are obtained. Research offering more than minimal risk and *not* offering direct benefit to the individual subjects is allowable only if the risk anticipated is minimally more than the risk the child would have experienced through his or her illness or through ordinary medical treatment, and if the expected generalizable knowledge is clear. Research involving more than minimal risk over what the child otherwise would experience, with no anticipated individual benefit, is unlikely ever to be approved. To be approved, the Secretary of DHHS would need to determine, after consultation with "a panel of experts in pertinent disciplines, for example, science, medicine, education, ethics, law, and following opportunity for public review and comment" (30) that the research is

expected to yield great understanding for the treatment of prevention of a serious problem affecting the health or welfare of children, that sound ethical principles are otherwise followed, and that permission and assent are obtained. Two other changes have occurred recently with regard to inclusion of children in research. The NIH issued "Policy and Guidelines on the Inclusion of Children as Participants in Research Involving Human Subjects" in 1998 (31). The policy mandates that research conducted or funded by the NIH include children in all studies unless there are scientific or ethical reasons to exclude children (for further discussion, see Ref. 32). The Food and Drug Administration (FDA) passed similar regulations in December 1998 requiring manufacturers of new and marketed drugs to evaluate the safety of those products in pediatric patients if the product is likely to be used in children (33).

Emergency Consent. In October 1996 the Federal Register published for the first time a waiver to informed consent requirements when conducting research in the emergency setting and when certain conditions apply (34). The regulations apply to Subparts A and D of the federal human subjects regulations (i.e., research with adults and with children) but not to Subparts B and C (research with pregnant women/fetuses and research with prisoners). The rationale for the new regulations was similar to rationales for wanting to be more inclusive of other populations of persons with limited decisional capacity: While it is imperative that the welfare and interests of individual vulnerable persons always be protected, it also is important that interventions that potentially might help such persons be tested and identified. Consequently, where research questions uniquely affect persons in emergency situations, it is sometimes appropriate, in balancing risks and benefits, that research be conducted. The conditions laid out in the new requirements include that eligible patients must be in a life-threatening condition, the available (nonexperimental) treatments are unsatisfactory, there is a need to collect scientific evidence to test new interventions, and an authorized representative for the patient cannot be found in the short time frame required. The investigators are required under the new waiver to document their efforts to find the patient's representative before enrolling the patient without consent. Further there must be procedures in place to inform the subject or his/her legally authorized representative, at the "earliest feasible opportunity," of the subject's inclusion in research and relevant details. If the trial remains ongoing at the time when the subject becomes aware, or the family is contacted, they have the right to terminate participation immediately. It also is required that the relevant IRB approve both the research activity and the waiver of consent. The waiver states explicitly that when appropriate, placebo-controlled trials are allowable under the new policy.

While in general, there has been considerable support for the emergency waiver, some have taken issue with the specific language included that could, again, lead to a misunderstanding about the distinctions between research and treatment. Jay Katz, for example, objected to the insinuation in the new regulations that emergency research has therapeutic intent for the individual patient: "In its emphasis on therapeutic benefits, the FDA obscures the fact that some of the permissible research activities either hold out no promise for therapeutic benefit or are so vaguely defined that potential therapeutic benefit can be inferred when research is the predominant intent. . . . Research must be stripped of the therapeutic illusion which misleads patient-subjects into believing that they are receiving the most advanced and beneficial treatments available, when instead they are being asked to serve the interests of science. . . . Research is not treatment" (35).

Other Persons With Limited Decision-Making Capacity

It is in many ways surprising that except for research with children and with persons in emergency situations, there are no federal regulations governing research with persons with questionable capacity to consent. Instead, it is left to IRBs to determine whether the research ought to go forward and whether consent and other safeguards are sufficient. Again, the challenge when evaluating research of this nature is wanting to enable more knowledge to accrue in this area, which clearly will benefit persons with mental disabilities, yet not wanting to compromise the interests of the individuals who participate in the studies. In order to determine if the participant can provide consent for him/herself, investigators should devise a method for evaluating whether participants have the ability to "understand, appreciate, and reason about the experimental situation" (36). Where they cannot, someone else must give permission on their behalf. Moreover, again, assent to whatever degree is possible, must be sought. If the potential subject will not assent, generally speaking, s/he should not be included in the research. Exceptions sometimes are made when there is clear likelihood that individual benefit would come to the participant as a result of enrolling, but given that research is by definition testing the efficacy of interventions, this may be difficult to prove. When deciding on behalf of someone else, two different standards can be used. If the person with limited decisional capacity previously had capacity, and had made relevant preferences known, then a substituted judgment standard can be used, meaning that the surrogate is simply voicing what s/he believes are (or previously have been) the wishes of potential subject. In contrast, if the subject has never had relevant capacity, then a best interest standard must be used. Essentially a caring person charged with the responsibility of guardianship for the subject decides what is best for the subject, based on how much anxiety it could provoke, its safety and invasiveness, and expected benefits and burdens (36,37). The obligation remains throughout the study to monitor ongoing effects in order to determine whether the subject's participation should be stopped at any time.

Persons With Serious Illnesses

As described earlier, a challenge for researchers is conducting research with persons with serious illnesses. These persons often are vulnerable by virtue of their illness. They may be so eager to participate in anything

that they, rightly or wrongly, believe may help them, that their judgment may be clouded. Moreover physician-investigators who sincerely care about the well-being of their patients, may encourage their patients to participate in research of unknown or little value because they too do not want to admit that few other options remain. While participation of such persons also may occur for altruistic reasons and while research in certain circumstances requires the participation of such persons, IRBs and investigators must be sensitive to these vulnerabilities, vulnerabilities that are not recognized the way being a prisoner, a child, or a person with limited decisional capacity routinely are.

Studies have been done that highlight how persons with serious illnesses often overestimate the benefit they could get from the research and/or forget or ignore altogether, as described earlier, that the research intervention is investigational. For example, Penman et al. found that the primary reasons 144 cancer patients accepted investigational chemotherapy were trust in their physician and belief that the treatment would help (27), and that one-quarter of the patient-subjects interviewed did not recall that the chemotherapy was investigational. Cassileth et al. documented that respondents describe research in different terms depending on whether they are speaking about research generally or their own participation. Patients and members of the general public reported that people generally should participate in research, in order to benefit others and increase scientific knowledge, but that they themselves would participate primarily in order to help "get the best medical care" (38). In a study conducted by the Advisory Committee on Human Radiation Experiments, cardiology and oncology outpatients who had been research participants said in closed-ended interviews that they viewed research as a way to help others (76 percent) and joined to get better treatment (67 percent) and because the research gave them hope (61 percent) (39). When interviewed in greater depth using more open-ended questions, however, these patients with serious illnesses reported that they joined research studies either because their doctors had recommended it, or because they believed they would gain additional medical benefit. For example, one patient-subject said, "When you reach that stage . . . and somebody offered that something that could probably save you, you sort of make a grab of it, and that's what I did" (40).

Concern about patient-subjects' vulnerabilities ought be most acute in the context of Phase I trials where the chance of personal medical benefit is minimal at best (41–43). Studies with patient-subjects enrolled in Phase I research, however, echo other findings. Rodenhuis et al. interviewed 44 patients who had agreed to participate in a Phase I cancer trial. They report that for all patients who participated in the study, "the hope for stabilization, improvement, or even cure of their diseases was the major motivation" (44). Further they report that doing *something* seemed to be of psychological benefit to patients in and of itself: "By continuing to receive medical attention and some form of treatment, they were able to cope with their incurable diseases and deny or postpone more easily the realization of impending death." Daugherty et al. asked patient-subjects enrolled in Phase I cancer research specifically whether they *expected* therapeutic benefit as a result of participation (45). Twenty-two percent of patients said they believed they would receive therapeutic benefit from their participation.

Genetics Research Informed Consent

Increasingly, researchers are conducting genetic research that raises new issues related to genetics research informed consent. Among these are issues surrounding use of stored tissue samples, the challenge of learning uninterpretable information, the potential for learning information about others who did not consent to the research, the potential for learning potentially harmful or damaging information, and the potential for risk to communities as well as to individuals.

Stored Tissue Samples. Tissue samples that are obtained from individuals, either as part of a research study or as part of clinical care, can be stored indefinitely. As such, previously collected samples of blood, particularly those that can be linked to certain demographic and clinical characteristics of the source individuals, are of great interest to other researchers. The ethics question becomes when and under what circumstances those samples may be used by future researchers for purposes quite unrelated to those for which they originally were collected and for purposes never disclosed to source individuals. The concern is that material risks to individuals can occur when genetic information about them is shared (see the discussion below); moreover, gathering information about someone that does *not* result in material harm still can wrong them if it is done without their knowledge and consent. In deciding individual cases, it is important to examine for what purposes consent originally was obtained from the source individuals and whether future researchers want the samples to remain identifiable. Many persons believe that it is inappropriate for individuals ever to be asked to provide blanket consent for all unknown future purposes if the samples remain identifiable (46). Rather, individuals may be asked willingness to provide consent for focused future purposes of their identifiable samples (e.g., future studies also related to Alzheimer's research). Bartha Knoppers has developed a core list of elements to include in a consent discussion when DNA samples will be stored (47). She suggests that individuals be required to agree or disagree with specific uses of data, including, for example, whether to undergo diagnostic tests, whether to permit consultation of their medical records, and whether to be contacted if genetic disorders are identified. Alternatively, samples may be stripped of identifiers. When samples are made anonymous, most ethics concerns disappear, since most risks and harms to individuals can never occur. The use of anonymous samples requires that the samples already existed when the new research plan was proposed and that it is impossible to go back and link samples to identifiable persons (48). Generally, when genetic information is anonymous, the potential benefits of scientific research

are thought to outweigh the risks to individual integrity even without permission from the source individuals.

Uninterpretable Information Learned through Genetic Research. Genetic testing, particularly in its early stages, often is probabilistic, rather than predictive, in nature. Through genetic research, markers may be identified that are associated with certain conditions, but their presence does not guarantee that the individual will become clinically affected, nor will all clinically affected persons possess the marker. Rather, additional markers and/or environmental factors will need to be identified that increase the predictive value of the genetic tests. This fact poses the challenge of, first, explaining complex probabilities to research participants who are more accustomed to extremely sensitive and specific diagnostic tests, and, in addition, determining when the level of uncertainty is so great that it is inappropriate to provide genetic testing results conducted in research to research participants. Indeed, some research at the initial stages of identifying a genetic marker may be considered far too premature to provide information to subjects. Researchers should consider the implications of ambiguous information before the study is initiated and determine *in advance* whether or not research test results will be made available to study subjects. If researchers determine that the results will not be made available, this must be communicated very clearly to subjects during the informed consent process. If results are to be made available, it is best to have a genetic counselor provide and interpret the information to subjects (46). Occasionally researchers may find themselves in the position of having informed subjects in advance that no information gained through the research would be disclosed, but then coming to believe that the information has greater clinical relevance than anticipated. Deciding whether to change disclosure procedures from that which was originally described is a difficult decision and should be made on a case-by-case basis with careful consideration and consultation from others, such as the IRB (49).

In addition to information being of questionable predictive value, there may be little to do as a result of learning it. There are many genetic conditions for which tests now are available yet for which there are no good treatments. It must be highlighted to research participants in such circumstances that the individual benefit to gaining the information (should it be disclosed) is psychological, rather than clinical. Further, when information will not be disclosed, it must be made clear that the purpose of the study is *not* to provide clinical benefit to the individuals who enroll. That is, the study may be conducted in order to compare the prevalence of a marker among different populations, or to determine the sensitivity and specificity of a recently developed test. Ellen Wright Clayton et al. (48) discuss possible legal liability from knowing information about a person that is not disclosed; she suggests that the risk of this is small if truly clinical decisions would not have been made differently had access to the information been available. It must be remembered that in some instances, however, the information *is* of clinical value to the individual or his/her relatives. In such instances, if researchers have a threshold level of confidence in the validity of their data, research participants should be given the choice of whether to learn the information discovered about them through their participation. Further, given that subjects often expect to be told information learned about them, or expect that hearing nothing means that no abnormalities were found, researchers must be very clear in circumstances in which research findings are not disclosed that this does not in any way reveal whether or not a marker was identified for that person.

Potential for Learning Information About Others without their Consent. Inherent to genetic information is that it usually is hereditary, and usually reveals at least some amount of information about other members of one's family. The type of study often conducted to identify a genetic marker is called a pedigree study. In pedigree studies researchers do genetic analyses of blood samples of many members of an extended family known either to be affected or unaffected by a hereditary condition in order to identify a genetic marker that exists uniquely among the affected individuals. Although researchers seek the consent of as many family members as possible, some cannot be found, and others refuse to participate. The awkwardness arises that through the testing of those who *do* provide consent, researchers may de facto learn genetic information about individuals who were not involved in the study. For example, an adult child may refuse to be in a study in which both of her parents agreed to participate, and the child's makeup is inevitably knowable; or one parent may refuse to participate while the other parent and children consent, again inevitably revealing part of the genotype of the parent who refused participation. It is up to IRBs and researchers to determine whether this ability to identify individuals, which ultimately is an invasion of their privacy, is acceptable, taking into consideration the sensitivity of the information revealed. In rare instances, pedigree studies may only be allowable when all members of an extended family agree to participate. Again, it is the role of genetic counselors, acting on behalf of investigators, to make this knowledge clear to participants in the informed consent process before a pedigree or family study is initiated.

Issues of family identification and confidentiality also become relevant in the context of research presentation and publication. While it always is true that research findings must be presented in ways that do not reveal the identities of the individuals who participated, researchers may forget that simply deleting names of individual research subjects does not guarantee anonymity. Pedigree studies may reveal the identities of individuals if, for example, the disease is rare and the family has other unusual characteristics (e.g., it includes triplets), and/or if the town in which many of the family live is identified in the report. It also is possible that individuals reading a published pedigree in a medical journal will recognize the family described as their own and will learn genetic information about other members of their extended family. There is great debate about whether it is ethically and scientifically appropriate for researchers to modify the

pedigree slightly in published reports in order to make the family no longer identifiable, for example, by randomly adding a few family members to the tree who never existed.

Gathering Potentially Harmful or Damaging Information, and Issues of Confidentiality. The information gathered through genetic testing may put individuals at risk of psychological or material harm. Individuals differ greatly in how they react to information that they often can do nothing about it. While some individuals find the knowledge beneficial, others may be made more anxious and upset about something over which they have no control. To the degree possible, counselors should try to walk individuals through different scenarios and try to determine how they think they would react to different types of information before they decide to be in studies in which disclosure could occur.

Moreover, impaired access to certain opportunities, such as health or life insurance and employment, may result as a consequence of participation in genetic research (50). While this rarely is due to researchers' negligence in maintaining confidentiality — and indeed some researchers have Certificates of Confidentiality from the federal government to provide additional confidentiality protections (51,52) — information can become available to insurers in other ways. For example, an individual may mention to her personal physician that she is enrolled in the pedigree study. If the physician makes a note of this in the patient's medical record, the information — even if the test results are not documented — could cause a red flag for insurers who might for other reasons be examining the patient's medical record. In such an instance, the insurer might insist on obtaining a copy of the research records before determining whether the individual was eligible for health, life, or disability insurance. Also persons applying for individual health, life, or disability insurance policies might be asked on a general health question about any medical testing or adverse findings of which they are aware. There is not clear consensus concerning whether information that individuals did not obtain through the clinical setting and was not obtained out of clinical concern must be disclosed to insurance companies in this context.

In some studies researchers bill health insurance companies for genetic testing conducted as part of the research. Obviously this alerts companies to the fact that the testing was done and, usually, that the individual is from a family of higher than average risk of disease. It may be appropriate for researchers, out of respect for the welfare of their participants, not to bill insurance companies in this manner.

A further potential harm arises from learning unanticipated information. While uncommon, it sometimes happens when conducting genetic testing that information that was not sought or anticipated becomes apparent. For example, researchers may become aware of mistaken paternity or of anomalies in the sex chromosomes (e.g., XXY genotypes). This often is troubling to researchers who are caught completely off guard and wonder if they have a responsibility to share the information with participants. In general, it is best for researchers to make it clear in advance that such incidental information will *not* be shared with participants. The exception may be information that researchers, *in consultation with others*, believe holds clinical relevance to the individual would lead the individual to act differently in some relevant way were he aware of it.

Information of Relevance to Communities, in Addition to Individuals. Often specific communities are targeted for genetic research. This may be a community of family members, or it may be an ethnic or religious group (e.g., Ashkenazi Jews), or it may be a group defined by where it lives, such as research conducted as part of the Human Genome Diversity Project, which seeks "information on human genetic diversity, the origins and migration of human populations, and genetic factors related either to resistance or susceptibility to disease" (53, p. 7). In such contexts, risks and potential benefits to communities must be considered, and a person or persons able to provide consent on behalf of the community must be identified. It has been suggested that a fourth principle of bioethics is needed to supplement those described earlier, a principle of "respect for communities" (53). Risks to communities from genetic research are real, since even research that strips samples of individual identifiers often still identifies the sample by "kindred, locality, or ethnicity" (54). In the past, material harms came to persons with sickle cell trait (rather than disease) as a result of being identified (54), and harmful stereotyping or self-perception can result from sweeping conclusions that sometimes are made in research. Foster et al. suggest a process of "communal discourse" to supplement and inform individual informed consent (54). This would occur through public meetings with representatives of the community in which opinions and suggestions are sought and communicated to investigators. Moreover, consulting an advisory board through the duration of the study can be helpful in ensuring that two-way communication remains ongoing.

In terms of recommendations, researchers should engage in an extensive consent discussion and dialogue with potential participants before and, often, during the research. All practical and procedural issues must be covered (amount of time required, purpose of results, etc.). Moreover, a genetic counselor should try to "walk" potential participants through different scenarios to try to discern how they would react to hearing ambiguous or probabilistic information. As a practical matter, given the risks to confidentiality that exist, investigators should apply for a certificate of confidentiality in genetic studies and also should minimize the likelihood that information they learn through studies can become part of the medical record. This means maintaining separate research and clinical records, not filing claims with health insurance companies for tests conducted solely for research purposes, and counseling participants about circumstances that might lead to information about them becoming part of their medical record.

CONCLUSION

Research as an enterprise has grown exponentially in the last several decades, and the field of research ethics has

grown with it. It is unfortunate that it takes examples of unethical practice to prompt us to develop appropriate guidelines and regulations, but the standards we currently have certainly are far more likely to protect the interests of potential subjects than was true decades before. As more research is conducted in the areas of genetics and incapacity, surely the field of research ethics similarly will move forward. In the meantime there is no better protection for subjects than a conscientious and humble researcher, who is aware that he or she may be easily misunderstood, and who is aware that the welfare of the subject always is of greater importance than any particular research question.

BIBLIOGRAPHY

1. *Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law No. 10*, vol. 2, U.S. Government Printing Office, Washington, DC, 1949, pp. 181–182.

1a. *United States v. Karl Brandt* et al., *The Medical Case, Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law No. 10*, U.S. Government Printing Office, Washington, D.C., 1949.

2. National Institutes of Health, Policy statement of November 17, 1953: *Group Consideration of Clinical Research Procedures Deviating from Accepted Medical Practice or Involving Unusual Hazard*, NIH, Washington, DC, 1953.

3. World Medical Association, *Declaration of Helsinki: Recommendations Guiding Medical Doctors in Biomedical Research Involving Human Subjects*, Adopted by the 18th World Medical Assembly, Helsinki, Finland, 1964 and as revised by the World Medical Assembly to Tokyo, Japan in 1975, in Venice, Italy in 1983, and in Hong Kong in 1989.

4. H.K. Beecher, *N. Engl. J. Med.* **274**, 1354–1360 (1966).

5. H.K. Beecher, *JAMA, J. Am. Med. Assoc.* **169**, 461–478 (1959).

6. Surgeon General, Public Health Service, to the Heads of the Institutions Conducting Research with Public Health Service Grants, February 8, 1966.

7. J.H. Jones, *Bad Blood*, Free Press, New York, 1993.

8. U.S. Department of Health, Education, and Welfare, *Final Report of the Tuskegee Syphilis Study Ad hoc Panel*, U.S. Government Printing Office, Washington, DC, 1973.

9. Advisory Committee on Human Radiation Experiments, *Final Report*, Oxford University Press, New York, 1996, p. 103.

10. National Research Act of 1974. P.L. 348, 93rd Cong., 2d Sess. (July 12, 1974).

11. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, April 18, 1979, U.S. Government Printing Office, Washington, DC.

12. T.L. Beauchamp and J.F. Childress, *Principles of Biomedical Ethics*, 4th ed., Oxford University Press, New York, 1994.

13. *Pratt v. Davis*, 118 Ill. App. 161 (1905) *aff'd* 79 N.E. 562 (Ill. 1906).

14. *Schloendorff v. Society of New York Hospitals*, 105 N.E. 92 (N.Y. 1914).

15. *Salgo v. Leland Stanford Jr. University Board of Trustees*, 317 P.2d 170 (Cal. Ct. App. 1957).

16. Title 45, Code of Federal Regulations, Part 46, §46.116, General requirements for informed consent.

17. R.R. Faden and T.L. Beauchamp, *A History and Theory of Informed Consent*, Oxford University Press, New York, 1986.

18. P.S. Appelbaum, C.W. Lidz, and A. Meisel, *Informed Consent: Legal Theory and Clinical Practice*, Oxford University Press, New York, 1987.

19. K.D. Hopper, T.R. TenHave, and J. Harzel, *Am. J. Radiol.* **164**, 493–496 (1995).

20. C.D. Meade and D.M. Howser, *Oncol. Nurs. Forum* **19**, 1523–1528 (1992).

21. M. Hochhauser, *IRB: Rev. Hum. Subj. Res.* September–October, 5–9 (1997).

22. M.A. Nakao and S. Axelrod, *Am. J. Med.* **74**, 1061–1065 (1983).

23. J.H. Fetting et al., *J. Clin. Oncol.* **8**, 1476–1482 (1990).

24. M.C. Silva and J.M. Sorrell, *Int. J. Nurs. Stud.* **21**, 233–240 (1984).

25. H.B. Muss et al., *Cancer (Philadalphia)* **43**, 1549–1556 (1979).

26. H.W. Riecken and R. Ravich, *JAMA, J. Am. Med. Assoc.* **248**, 344–348 (1982).

27. D.T. Penman et al., *J. Clin. Oncol.* **7**, 849–855 (1984).

28. P.S. Appelbaum et al., *Hastings Cent. Rep.* **17**, 20–24 (1987).

29. Title 45, Code of Federal Regulations, Part 46, §46.402.

30. Title 45, Code of Federal Regulations, Part 46, §46.407.

31. National Institute of Health (NIH), *Guide for Grants and Contracts*, vol. 27, NIH, Washington, DC, 1998.

32. J. Sugarman, A.C. Mastroianni, and J.P. Kahn, *Ethics of Research with Human Subjects*, University Publishing Group, Frederick, MD, 1998, pp. 169–175.

33. U.S. Department of Health and Human Services, Food and Drug Administration, *Fed. Regis.* **63**(231), 66631–66672 (1998).

34. U.S. Department of Health and Human Services, *Fed. Regis.* **61**, 51531–51533 (1996).

35. J. Katz, *Hastings Cent. Rep.* **27**, 9–11 (1997).

36. R. Dresser, *JAMA, J. Am. Med. Assoc.* **276**, 67–72 (1996).

37. E. DeRenzo, *Cambridge Q. Healthcare Ethics* **3**, 539–548 (1994).

38. B.R. Cassileth, E.J. Lusk, D.S. Miller, and S. Hurwitz, *JAMA, J. Am. Med. Assoc.* **248**, 968–970 (1982).

39. Advisory Committee on Human Radiation Experiments, *Final Report*, Oxford University press, New York, 1995.

40. N. Kass, J. Sugarman, R. Faden, and M. Schoch-Spana, *Hastings Cent. Rep.* **26**, 25–29 (1996).

41. D.D. von Hoff and J. Turner, *Invest. New Drugs* **9**, 114–122 (1991).

42. E. Estey et al., *Proc. Am. Soc. Clin. Oncol.* **3**, 35 (1984).

43. E. Estey et al., *Cancer Treat. Rep.* **70**, 1105–1115 (1986).

44. S. Rodenhuis et al., *Eur. J. Cancer Clin. Oncol.* **20**, 457–462 (1984).

45. C. Daugherty et al., *J. Clin. Oncol.* **13**, 1062–1072 (1995).

46. American Society of Human Genetics, *Am. J. Hum. Genet.* **59**, 471–474 (1996).

47. B.M. Knoppers and C.M. Laberge, *JAMA, J. Am. Med. Assoc.* **274**, 1806–1807 (1995).

48. E. Wright Clayton et al., *JAMA, J. Am. Med. Assoc.* **274**, 1786–1792 (1995).

49. E. Kodish, T.H. Murray, and S. Shurin, *Clin. Res.* **42**(3), 396–402 (1994).

50. N.E. Kass, *IRB: Rev. Hum. Subj. Res.* **15**, 7–10 (1993).

51. C.L. Early and L.C. Strong, *Am. J. Hum. Genet.* **57**, 727–731 (1994).
52. Public Health Service Act, S.301(d), 42 U.S.C.
53. K. Cranley Glass et al., *IRB: Rev. Hum. Subj. Res.* **19**, 1–13 (1997).
54. M.W. Foster, A.J. Eisenbraun, and T.H. Carter, *Nat. Genet.* **17**, 277–279 (1997).

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, ETHICS, PRINCIPLES GOVERNING RESEARCH WITH HUMAN SUBJECTS

ROBERT J. LEVINE
Yale University School of Medicine
New Haven, Connecticut

OUTLINE

INTRODUCTION

In 1974 the United States Congress passed the National Research Act (Public Law 93-348) which established the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (hereafter, the National Commission). This Commission was charged by Congress to "identify the basic ethical principles which should underlie the conduct of ... research involving human subjects (and) to develop guidelines which should be followed in such research to assure that it is conducted in accordance with such principles...." The principles identified by the National Commission were published in 1978 in its *Belmont Report* (1). This document proved to be highly influential; its principles provide the basis for virtually all commentary on the ethics of research involving human subjects. Moreover these principles have been adopted in major international documents that contain guidance for the ethical conduct of research involving human subjects such as the Council of International Organizations of Medical Sciences' (CIOMS) *International Ethical Guidelines for Biomedical Research Involving Human Subjects* (2).

The guidelines developed by the National Commission were published in a series of reports (3). They have been adopted by the U.S. federal government as the regulations for the protection of human subjects and have also influenced the development of policy in many other nations.

The National Commission defined the basic ethical principle as "a general judgment that serves as a basic justification for the many particular prescriptions for and evaluations of human actions" (1, p. 4). Such a principle is taken as an ultimate foundation for any second-order principles, rules, and norms; it is not derived from any other statement of ethical values. The National Commission identified three basic ethical principles as particularly relevant to the ethics of research involving human subjects: respect for persons, beneficence, and justice. The norms and procedures presented in regulations and ethical codes are derived from and are intended to uphold these fundamental principles. As the National Commission observed in an early draft of the *Belmont Report*:

> Reliance on these three fundamental underlying principles is consonant with the major traditions of western ethical, political and theological thought represented in the pluralistic society of the United States, as well as being compatible with the results of an experimentally based scientific analysis of human behavior... (3, p. 18).

Thus, in the National Commission's view, these principles pertain to human behavior in general; it is through the development of norms that they are made peculiarly relevant to specific classes of activities such as research and the practice of medicine.

Some of the language used by the National Commission may imply an endorsement of one or another foundational ethical theory. For example, the term "respect for persons" suggests a reference to Kantian theory. It is clear, however, that the National Commission did not embrace any such theory (3). As observed by Abram and Wolf:

> Aware of Kantian (deontological), utilitarian, and Aristotelian traditions, for instance, the commission nonetheless refrained from relying on any one of them for the legitimacy of its conclusions. Agreement on a fundamental moral system was not sought or needed (4).

Although these authors discussed the President's Commission, they made it clear that it was patterned after the National Commission, which similarly refrained from relying exclusively on any particular moral system.

ETHICAL PRINCIPLES

Respect for Persons

The principle of respect for persons was stated formally by Immanuel Kant: "So act as to treat humanity, whether

in thine own person or in that of any other, in every case as an end withal, never as a means only." However, what it means to treat a person as an end and not merely as a means to an end may be variously interpreted. The National Commission concluded that:

> Respect for persons incorporates at least two basic ethical convictions: First, that individuals should be treated as autonomous agents, and second, that persons with diminished autonomy and thus in need of protection are entitled to such protections (1).

An autonomous person is "...an individual capable of deliberation about personal goals and of acting under the direction of such deliberation" (1). To show respect for autonomous persons requires that we leave them alone, even to the point of allowing them to choose activities that might be harmful (e.g., mountain climbing) unless they agree or consent that we may do otherwise. We are not to touch them or to encroach upon their private spaces unless such touching or encroachment accords with their wishes. Our actions should be designed to affirm their authority and enhance their capacity to be self-determining; we are not to obstruct their actions unless they are clearly detrimental to others. We show disrespect for autonomous persons when we either repudiate their considered judgments or deny them the freedom to act on those judgments in the absence of compelling reasons to do so.

Clearly, not every human being is capable of self-determination. The capacity for self-determination matures during a person's life; some lose this capacity partially or completely owing to illness or mental disability or in situations that severely restrict liberty, such as in prisons. Respect for the immature or the incapacitated may require one to offer protection to them as they mature or while they are incapacitated.

Beneficence

The principle of beneficence is firmly embedded in the ethical tradition of medicine. It is commonly said that the first principle of medical ethics is "Do no harm." This principle is often stated in Latin, *primum non nocere*, which translated literally means "first (or above all) do no harm." Moreover this statement of principle is commonly and incorrectly attributed to Hippocrates.

Parenthetically, the closest approximation of this statement that can be found in the Hippocratic writings is in the book entitled *Epidemics*: "As to diseases, make a habit of two things — to help, or at least to do no harm" (5). If the first principle of medicine were truly "above all, do no harm," this would rule out virtually all medical therapy; almost all therapies present to the patient a risk of injury. The statement from *Epidemics* is much more compatible with the modern emphasis on trying to achieve a favorable balance of harms and benefits.

In the Hippocratic Oath, the principle of beneficence is expressed in several statements such as:

> I will apply dietetic measures for the benefit of the sick according to my ability and judgment. ...I will neither give a deadly drug to anybody if asked for it, nor will I make a suggestion to this effect.

In biomedical research the leading ethical codes such as the World Medical Association's Declaration of Helsinki enjoin the physician-investigator not only to secure the well-being of individuals (research subjects and patients) but also to develop information that will form the basis of being better able to do so in the future. And, according to the Nuremberg Code, the risks of research must be justified by "the humanitarian importance of the problem to be solved by the experiment."

The National Commission defined 'beneficence' as follows:

> The term, *beneficence*, is often understood to cover acts of kindness or charity that go beyond strict obligation. In this document, beneficence is understood in a stronger sense, as an obligation. Two general rules have been formulated as complementary expressions of beneficent actions in this sense: (*1*) Do no harm and (*2*) maximize possible benefits and minimize possible harms (1).

The first of these two "general rules" proscribes the deliberate infliction of serious injury on an identified individual for research purposes. That is, one may not impose a 100 percent probability of disabling injury on a human subject with no justification other than to solve a research problem (6). "Do no harm" as envisioned by the National Commission does not mean "above all, do no harm." Rather, it permits exposing research subjects to a statistical probability of harm if such exposure is justified in terms of the anticipated benefits, among other considerations (3).

The second general rule of beneficence calls upon investigators to design all of their work so as to maximize the probability and magnitude of benefit to individual research subjects as well as to society. It further requires investigators to minimize the probability and magnitude of injury to subjects and of harm to the interests of human collectives (3). Among the ethical norms that are grounded primarily in the second general rule are the requirements for good research design and competent investigators (3) and the requirement for a favorable relation of risks to anticipated benefits (3, pp. 37ff).

Some authors argue that separation of these two general rules into two fundamental ethical principles, beneficence (do good) and nonmaleficence (do no harm), would tend to decrease confusion (7). For the present purposes, however, it is more convenient to treat beneficence as a single principle; it is generally necessary to consider harms and benefits in relation to each other.

Frankena identifies four obligations that derive from the principle of beneficence (8); listed in decreasing order of ethical force, these are:

A. One ought not to inflict evil or harm.
B. One ought to prevent evil or harm.
C. One ought to remove evil.
D. One ought to do or promote good.

Statement A is a straightforward articulation of the National Commission's first general rule, do no harm (also known as principle of nonmaleficence). Few, if any, would argue that the injunction against inflicting harm or evil

is not at least a very strong prima facie duty. Statement D, by contrast, is not regarded generally as a duty or obligation but rather, as the National Commission pointed out, an exhortation to act kindly or charitably. Statement D becomes a duty in the strict sense most typically when one consents or contracts to be bound by it. Physicians, for example, pledge themselves to act for the benefit of patients. Similarly, researchers who accept public support for their work assume a contractual obligation to promote good by contributing to the development of new knowledge. Moreover, without regard to funding, when one invites human beings to participate in activities that expose them to risk of injury, it is generally necessary to offer them in return something they find valuable; in the context of research involving human subjects, the most generally acceptable item of value that can be offered is a promise to pursue benefits either to the individual subjects or to others with whom they feel a bond of kinship (3). Such promises are made explicit in the process of informed consent. In the light of these considerations the National Commission determined that investigators who performed research involving human subjects incurred a strict duty or obligation to do or promote good.

Justice

Justice requires that we treat persons fairly and we give each person what he or she is due or owed (3). Justice is either comparative or noncomparative. Comparative justice is concerned with determining what a person is due by weighing his or her claims against the competing claims of others. Noncomparative justice is concerned with identifying what persons are due without regard to the claims of others (e.g., never punish an innocent person). The concerns addressed by the National Commission under the rubric of justice are exclusively concerns of distributive justice, a type of comparative justice. This article follows the National Commission in using the term "justice" to mean "distributive justice."

Distributive justice is concerned with the distribution of scarce benefits where there is competition for these benefits. If there is no scarcity, there is no need to consider just systems of distribution. Distributive justice is also concerned with the distribution of burdens, specifically when it is necessary to impose burdens on fewer than all members of a seemingly similar class of persons.

Justice requires a fair sharing of burdens and benefits; however, just what constitutes a fair sharing is a matter of considerable controversy. To determine who deserves to receive which benefits and which burdens, we must identify morally relevant criteria for distinguishing unequals. Various criteria have been proposed (9). Is it fair for persons to be treated differently on the basis of their needs? Their accomplishments? Their purchasing power? Their social worth? Their past records or future potential?

There are those who argue that the fairest distribution of burdens and benefits is precisely that which creates the most benefits for society at large. This is the classical utilitarian argument, which harmonizes the principles of justice and beneficence by stipulating that there is no conflict. To create goods is to do justice; just institutions act so as to produce the greatest good for the greatest number. The National Commission rejected this formulation because it does not accord either with Western concepts of the fundamental equality of persons (e.g., before the law) or with the very strong tradition that interprets fairness to require extra protection for those who are weaker, more vulnerable, or less advantaged than others. This latter interpretation is reflected in such disparate sources as the injunction in the Judeo-Christian tradition to protect widows and orphans, the Marxist dicta "from each according to ability; to each according to need," and more recently, Rawls's contractual derivation of principles of justice (10).

The National Commission's interpretation of the requirements of justice is embodied in one of its statements on the relevance of this principle to the problem of selection of subjects:

> [T]he selection of research subjects needs to be scrutinized in order to determine whether some classes (e.g., welfare patients, particular racial and ethnic minorities, or persons confined to institutions) are being systematically selected simply because of their easy availability, their compromised position, or their manipulability, rather than for reasons directly related to the problem being studied. Finally, whenever research supported by public funds leads to the development of therapeutic devices and procedures, justice demands both that these not provide advantages only to those who can afford them and that such research should not unduly involve persons from groups unlikely to be among the beneficiaries of subsequent applications of the research (1, pp. 9–10).

Each of the three basic ethical principles identified by the National Commission is intended to have equal moral force. In this way the ethical system differs from those that have a single overarching superprinciple such as "justice" or *agape*. It further differs from those that assign priority to principles — to rank them in order of moral forcefulness — to resolve disputes engendered by conflicting requirements of two or more principles (10). The three principles give rise to norms that often create conflicting requirements. For example, the principle of justice, as articulated by the National Commission, creates requirements that are incompatible with some of those created by its principle of respect for persons (3). Similarly norms derived from the principle of beneficence inevitably engender conflicts with those arising from the principle of respect for persons. Implicitly, then, the National Commission also endorsed the notion of prima facie rules; these are rules that are binding unless they are in conflict with other stronger rules or unless in specific situations there is ethical justification for overriding the rule's requirements (3).

ETHICAL NORMS

An ethical norm is a statement that actions of a certain type ought (or ought not) to be done. If reasons are supplied for these behavioral prescriptions (or proscriptions), they are that these acts are morally right (or wrong).

Statements of ethical norms commonly include the words "should" or "ought," but in some cases there are stronger terms such as "must" or "forbidden." A typical statement of an ethical norm is: Research should be conducted only by scientifically qualified persons. The behavior-prescribing statements contained in the various codes and regulations on research involving human subjects may be regarded as variants of five general ethical norms (11). There should be (*1*) good research design, (*2*) competent investigators, (*3*) a favorable balance of harm and benefit, (*4*) informed consent, and (*5*) equitable selection of subjects. In addition a sixth general ethical norm appears in some international guidelines: (*6*) there should be compensation for research-induced injury. The purpose of these ethical norms is to indicate how the requirements of the three fundamental ethical principles may be met in the conduct of research involving human subjects.

Because statements of the ethical norms in codes and regulations tend to be rather vague, they permit a variety of interpretations; it is sometimes difficult to know exactly how to apply them to particular cases. When faced with such uncertainty, it is generally helpful to look behind the norm to examine the fundamental ethical principle or principles it is intended to uphold or embody. Accordingly, the discussion of each ethical norm will call attention to the fundamental ethical principle or principles it is designed to serve.

Good Research Design

> The experiment should be so designed and based on the results of animal experimentation and a knowledge of the natural history of the disease or other problem under study that the anticipated results will justify the performance of the experiment (Nuremberg 3).
>
> Biomedical research involving human subjects must conform to generally accepted scientific principles and should be based on adequately performed laboratory and animal experimentation and on a thorough knowledge of the scientific literature (Helsinki I. 1).
>
> [S]cientifically unsound research on human subjects is *ipso facto* unethical in that it may expose subjects to risk or inconvenience to no purpose (CIOMS 14, commentary).

These are typical expressions of the ethical requirement that research must be sufficiently well-designed to achieve its purposes; otherwise, it is not justified. The primary purpose of this norm is to uphold the principle of beneficence. If the research is not well designed, there will be no benefits; investigators who conduct badly designed research are not responsive to the obligation to do good or to develop generalizable knowledge that is sufficiently important to justify the expenditure of public funds, to impose upon human subjects risks of physical or psychological harm, and so on.

This norm is also responsive to the principle of respect for persons. Persons who agree to participate in research as subjects are entitled to assume that something of value will come of their participation. Poorly designed research wastes the time of the subjects and frustrates their desire to participate in a meaningful activity.

Competence of the Investigators

> The experiment should be conducted only by scientifically qualified persons. The highest degree of skill and care should be required through all stages of the experiment of those who conduct or engage in the experiment (Nuremberg 8).
>
> Biomedical research involving human subjects should be conducted only by scientifically qualified persons and under the supervision of a clinically competent medical person. The responsibility for the human subject must rest with a medically qualified person . . . (Helsinki I.3).

This norm requires that the investigators be competent in at least two respects. They should have adequate scientific training and skill to accomplish the purposes of the research. The purpose of this component of the norm is precisely the same as that requiring good research design; it is responsive primarily to the obligations to produce benefits to society through the development of important knowledge. It is also responsive to the obligation to show respect for research subjects by not wasting their time or frustrating their wishes to participate in meaningful activities. In addition investigators are expected to be sufficiently competent to care for the subject. The Declaration of Helsinki, as an instrument of the World Medical Association, is addressed only to medical research. Therefore, it places responsibility with ". . .a medically qualified person." The Nuremberg Code, on the other hand, is addressed more generally to research; consequently, it does not call for medical qualification.

Competence to care for the subjects of most clinical research requires that at least one member of the research team be responsible for observing the subject with a view toward early detection of adverse effects of his or her participation or other evidence that the subject should be removed from the study. The investigator should have the competence to assess the subjects' symptoms, signs, and laboratory results. There should further be the competence to intervene as necessary in the interests of minimizing any harm, such as by prompt administration of an antidote to a toxic substance.

Balance of Harms and Benefits

There are normative statements in the ethical codes and regulations that require a favorable balance between harm and benefit. Without such a favorable balance there is no justification for beginning or continuing the research.

> The degree of risk to be taken should never exceed that determined by the humanitarian importance of the problem to be solved by the experiment (Nuremberg 6).
>
> Biomedical research involving human subjects cannot legitimately be carried out unless the importance of the objective is in proportion to the inherent risk to the subject (Helsinki I.4).
>
> Risks to the subjects [must be] reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result (CFR. 46. 111a(2)).

There are additional norms in codes and regulations that call for vigilance on the part of those conducting

or supervising the research. At any point along the way, the balance of harms and benefits may become unfavorable; under these circumstances the research should be terminated.

> During the course of the experiment the scientist in charge must be prepared to terminate the experiment at any stage, if he has probable cause to believe ... that a continuation ... is likely to result in injury, disability or death to the experimental subject (Nuremberg 10).
>
> The investigator ... should discontinue the research if in his/her ... judgment it may, if continued, be harmful to the individual (Helsinki III. 3).
>
> Where appropriate, the research plan makes adequate provision for monitoring the data collected to insure the safety of the subjects (CFR. 111a (6)).
>
> An IRB shall have authority to suspend or terminate approval of research ... that has been associated with unexpected serious harm to subjects (CFR. 113).

The requirement that research be justified on the basis of a favorable balance of harms and benefits is derived primarily from the ethical principle of beneficence. In addition a thorough and accurate compilation of the risks and hoped-for benefits of a research proposal also facilitates responsiveness to the requirements of the principles of respect for persons and justice. A clear and accurate presentation of risks and benefits is necessary in the negotiations with the subject for informed consent. Similarly such a compilation of burdens and benefits facilitates discussions of how they might be distributed equitably.

Ethical codes and regulations require not only that risks be justified by being in a favorable relationship to hoped-for benefits but also that they be minimized.

> The experiment should be so conducted as to avoid all unnecessary physical and mental suffering and injury (Nuremberg 4).
>
> The IRB shall determine that ... risks to subjects are minimized: 1) By using procedures which are consistent with sound research design and which do not unnecessarily expose subjects to risk, and (ii) whenever appropriate, by using procedures already being performed on the subjects for diagnostic or treatment purposes (CFR. 111a (1)).

Informed Consent

Principle I of the Nuremberg Code provides the definition of consent from which the definitions contained in all subsequent codes and regulations are derivative:

> The *voluntary* consent of the human subject is absolutely essential.
>
> This means that the person involved should have *legal capacity* to give consent; should be so situated as to be able to exercise *free power of choice*, without the intervention of any element of force, fraud, deceit, duress, over-reaching or other ulterior form of constraint or coercion; and should have sufficient *knowledge* and *comprehension* of the elements of the subject matter involved as to enable him to make an understanding and enlightened decision. This latter element requires that before the acceptance of an affirmative decision by the experimental subject there should be made known to him the nature, duration, and purpose of the experiment; the method and means by which it is to be conducted; all inconveniences and hazards reasonably to be expected; and the effects upon his health or person which may possibly come from his participation in the experiment [*emphasis supplied*].

Thus the consent of the subject in order to be recognized as valid must have four essential attributes. It must be competent (legally), voluntary, informed, and comprehending (or understanding).

It is through informed consent that the investigator and the subject enter into a relationship, defining mutual expectations and their limits. This relationship differs from ordinary commercial transactions in which each party is responsible for informing himself or herself of the terms and implications of any of their agreements. Professionals who intervene in the lives of others are held to higher standards. They are obligated to inform the lay person of the consequences of their mutual agreements.

According to the President's Commission, "Although the informed consent doctrine has substantial foundations in law, it is essentially an ethical imperative" (12, p. 2). The President's Commission refers repeatedly to "ethically valid consent" and in this way reflects a perspective differing with that of federal regulations, which refer to "legally effective informed consent."

The requirement for informed consent is designed to uphold the ethical principle of respect for persons (3, pp. 95ff). It is through informed consent that we make operational our duty to respect the rights of others to be self-determining, namely to be left alone or to make free choices. We are not to touch others or to enter their private spaces without permission. As stated by Justice Cardozo, "Every human being of adult years and sound mind has a right to determine what will be done with his own body..." (13, p. 526).

Privacy and Confidentiality

Closely related to the norms calling for informed consent and minimization of risk are the requirements found in ethical codes and regulations that protect privacy and confidentiality. Privacy is "the freedom of the individual to pick and choose for himself the time and circumstances under which, and most importantly, the extent to which, his attitudes, beliefs, behavior and opinions are to be shared with or withheld from others" (14). In general, investigators are not permitted to intrude into individuals' privacy without their informed consent. When an informed person allows an investigator into his or her private space, there is no invasion.

"Confidentiality," a term that is often and incorrectly used interchangeably with "privacy," refers to a mode of management of private information; if a subject shares private information with (confides in) an investigator, the investigator is expected to refrain from sharing this information with others without the subject's authorization or some other justification.

The ethical grounding for the requirement to respect the privacy of persons may be found in the principle of respect for persons. The ethical requirement for the maintenance of confidentiality of private information is grounded in the norm calling for the minimization of harms. Breaches of

confidentiality may result in such social injuries as loss of personal autonomy, valued relationships, or eligibility for insurance or employment. Moreover, maintenance of confidentiality is essential to the successful practice of various professions; for example, sick people would not consult physicians unless they were confident that their private information would be kept secret (15).

Equitable Selection of Subjects

> Individuals ... to be invited to be subjects of research should be selected in such a way that the burdens and benefits of the research will be equitably distributed. Special justification is required for inviting vulnerable individuals ... (CIOMS 10).
>
> [T]he IRB shall determine that.... Selection of subjects is equitable ... (CFR. 111a (3)).

This requirement is derived from the principle of justice, which requires equitable distribution of both the burdens and the benefits of research. Until the 1970s codes of ethics and regulations were relatively silent on this matter; however, the preamble to the Nuremberg Code reflected an concern with issues of social justice. It pointed out that the "crimes against humanity" were particularly egregious in that they were perpetrated on "non-German nationals, both prisoners of war and civilians, including Jews and 'asocial' persons" Implicit in this statement is the perception that because these subjects were not considered persons in the full sense of the word, they were not accorded the respect due to fully enfranchised persons. As a consequence principle I of the Nuremberg Code established the high standards for consent discussed earlier. When thoroughly honest offers are made to fully autonomous persons, they are presumed capable of defending their own interests and of selecting themselves as research subjects. Because the Nuremberg Code does not entertain the possibility of involving less than fully autonomous subjects, no requirements for their selection are provided.

Ethical codes and regulations in the field of research involving human subjects project an attitude of protectionism. Their dominant concerns are the protection of individuals from injury and from exploitation. There are important historical reasons for this protectionistic attitude. These documents were written with the aim of ensuring that there would never be a repetition of atrocities like those committed by the Nazi physician-researchers, calamities like the thalidomide experience, or ethical violations like those of the Tuskegee syphilis study. In recent years society's perception of biomedical research has shifted dramatically. Now, largely as a consequence of the efforts of the AIDS activists, biomedical research is widely perceived as benign and beneficial (16).

In the era of protectionism, which lasted from the mid-1940s through the mid-1980s, those who wrote policy for the protection of research subjects interpreted the principle of justice to require protection of vulnerable or disadvantaged persons from bearing an unfair share of the burdens of serving as research subjects. Since the mid-1980s, as a reflection of the shift in society's attitude toward research, policies and practices are being revised; the same principle of justice is being interpreted to require assurance that the vulnerable and disadvantaged will enjoy equitable access to the benefits of participation in research.

Populations who were excluded previously because they were considered vulnerable include children, women who have the biological capacity to conceive, and members of racial and ethnic minorities; in recent years federal policies have been revised to require adequate representation of each of these populations in research (17).

Compensation for Research-Induced Injury

> Research subjects who suffer physical injury as a result of their participation are entitled to such financial or other assistance as would compensate equitably for any temporary or permanent impairment or disability. In case of death, their dependants are entitled to material compensation. The right to compensation may not be waived (CIOMS 13).

During the 1970s commentators on the ethics of research reached a consensus that subjects who are injured as a consequence of their participation in research are entitled to compensation. The ethical arguments to support this entitlement are grounded in considerations of compensatory justice (18). Compensatory justice consists in giving injured persons their due by taking account of their previous conditions and attempting to restore them. Sometimes it is possible to literally restore injured persons to their previous conditions, such as through medical therapy for the research-induced injury or illness. On other occasions, when literal restoration is not feasible, a monetary substitute is about the best we can do. Most discussions of compensation for research-induced injury focus on the provision of monetary substitutes in cases in which there is temporary or permanent disability or death.

In the commentary under its guideline number 13, CIOMS notes:

> In some societies the right to compensation for accidental injury is not acknowledged. Therefore, when giving their informed consent to participate, research subjects should be told whether there is a provision for compensation in case of physical injury, and the circumstances in which they or their dependants would receive it.

The United States is one society in which the right to such compensation is not acknowledged. Federal regulations require that prospective research subjects receive:

> an explanation as to whether any compensation and an explanation as to whether any medical treatments are available if injury occurs and, if so, what they consist of, or where further information may be obtained ... (46.116a).

PROCEDURAL NORMS

Ethical codes and regulations contain, in addition to the substantive ethical norms, descriptions of procedures (procedural norms) that are to be followed to assure that investigators comply with the requirements of the substantive norms. The most important general procedural requirement that is relevant to all research

involving human subjects conducted in the United States is the review by an Institutional Review Board (IRB).

Federal regulations require that research involving human subjects be reviewed and approved by an IRB before it may be initiated. The criteria for approval are to ensure that plans are adequate for compliance with each of the substantive norms other than those calling for good research design and competent investigators; responsibility for these determinations is assigned to other agents or agencies (19). In other countries the same assignment is issued to research ethics committees or research ethics boards (20). Article I.2 of the Declaration of Helsinki requires only that the "experimental protocol ... be ... transmitted for consideration, comment and guidance to a committee independent of the investigator and sponsor...." CIOMS guideline 14, by contrast, requires "review and approval by one or more independent ethical and scientific review committees."

Another general procedural requirement (general in that it is designed to ensure compliance with all of the substantive norms) is concerned with the publication of the results of research that appears to have been done unethically. US federal regulations are silent on this issue. There are differing positions in the international documents:

> Reports of experimentation not in accordance with the principles laid down in this Declaration should not be accepted for publication (Helsinki I.8).

CIOMS offers the following commentary under Guideline 15:

> Refusal to publish the results of research conducted unethically ... may be considered, as may refusal to accept unethically obtained data submitted in support of an application for drug registration. However, these sanctions deprive of benefit not only the errant investigator or sponsor but also that segment of society intended to benefit from the research; such possible consequences merit careful consideration.
>
> publication of reports of the results of research ... should include, when appropriate, a statement that the research was conducted in accordance with these guidelines. Departures, if any, should be explained and justified....

There are also specific procedural norms such as the requirement for documentation of informed consent which is designed to ensure compliance with the substantive norm that calls for informed consent.

VULNERABLE PERSONS AS RESEARCH SUBJECTS

When the federal government published its first proposals to develop regulations providing additional protections for especially vulnerable populations of research subjects, it designated them as persons having "limited capacities to consent" (3). The choice of this label highlights the nature of the fundamental problem in justifying their use as research subjects. Because the Nuremberg Code identifies voluntary consent as "absolutely essential," it is clearly problematic to involve subjects who lack free power of choice (e.g., prisoners), the legal capacity to consent (e.g., children), or the ability to comprehend (e.g., the mentally infirm). (Another term commonly used for "those having limited capacities to consent" is "the special populations.")

The National Commission concluded that persons having limited capacity to consent are vulnerable or disadvantaged in ways that are morally relevant to their involvement as subjects of research (3). Therefore the principle of justice is interpreted as requiring that we facilitate activities that are designed to yield direct benefit to the subjects and that we encourage research designed to develop knowledge that will be of benefit to the class of persons of which the subject is a representative. However, we should generally refrain from involving the special populations in research that is irrelevant to their conditions as individuals or at least as a class of persons. Respect for persons is interpreted as requiring that we show respect for a potential subject's capacity for self-determination to the extent that it exists. Some who cannot consent can register knowledgeable agreements (assents) or deliberate objections. In most instances the assent of an individual who cannot consent must be supplemented by the permission of that individual's parent or guardian (in the language of federal regulations, the "legally authorized representative").

To the extent that the capacity for self-determination is limited, respect is shown by protection from harm. Thus the Commission recommends that the authority accorded to members of the special populations or their legally authorized representatives to accept risk be strictly limited; any proposal to exceed the threshold of "minimal risk" requires special justification.

The reports of the National Commission on each of the "special populations" were followed by the promulgation of federal regulations providing "additional protections" for the fetus and pregnant women and for human in vitro fertilization (IVF); for prisoners, and for children (3). Regulations also were proposed for "those institutionalized as mentally disabled" but these have not been promulgated as final regulations (3).

Federal regulations also require:

> Where some or all of the subjects are likely to be vulnerable to coercion or undue influence, such as persons with acute or severe physical or mental illness, or persons who are economically or educationally disadvantaged, appropriate additional safeguards (should be) included in the study to protect the rights and welfare of these subjects (46.111b).

CIOMS, in the commentary under guideline 10, provides an extensive list of the types of persons who may in certain circumstances be considered vulnerable.

ARE THE BASIC ETHICAL PRINCIPLES UNIVERSAL?

The question of whether the basic ethical principles are universal will elicit radically different responses from adherents of ethical universalism than it will from the cultural pluralists. The tension between these two positions has existed since classical times and it is unlikely to be resolved in the foreseeable future (21).

Ethical universalists believe there is a universal set of ethical principles that are applicable to all human beings regardless of their situations in particular cultures. The task of the moral philosopher, then, is to *discover* those universal principles that apply in all times and in all places. Variations across cultures indicate that some societies are ahead of others in the degree of "moral progress" they have accomplished (22).

Ethical pluralists, by contrast, recognize that all ethical principles are developed in the course of discussions held within particular cultures and that these discussions necessarily reflect the unique histories and other circumstances of particular cultures. On this view ethical principles are *invented* rather than discovered. Pluralists further acknowledge the inevitability and recognize the legitimacy of variation across cultures of ethical norms and principles.

This debate has practical import in the increasingly common circumstances in which research protocols are designed by investigators and sponsors in technologically developed countries and then carried out in developing (or underdeveloped) countries or communities. What if there are differences in ethical values in the two cultures? Whose ethics should apply?

It has been argued that there should be a compromise between the two extremes (21). Some ethical principles seem to be universally valid. For example, there is a universal proscription against inflicting injury on a person without justification. But what counts as justification in various societies differs substantially. The principle of respect for persons, when stated at a sufficient level of abstraction, enjoining people to treat persons as ends and not merely as means, is universally applicable. However, when this principle is elaborated to require that all persons are to be treated as self-determining, it loses its relevance to some cultures in which individual self-determination is less highly valued than it is in the United States.

American universalists would argue that persons in such cultures must be educated; they must be taught to value self-determination as much as we do. Some might add that they must learn to value and protect the right to be self-determining or else they will remain vulnerable to exploitation by those who have decision-making authority. Pluralists counter this argument by pointing out that the society seems functional as it is; if we impose on it our ethical standards, it may have a destructive effect on the culture. Furthermore we should show respect for a society by allowing it to be self-determining.

The CIOMS *International Ethical Guidelines* reflect what may be a satisfactory compromise position, holding that some ethical standards are universal while recognizing the legitimacy of some degree of ethical pluralism. These guidelines set forth procedures to be followed when research is initiated and financed in one country (the external sponsoring country) and carried out by investigators from the external sponsoring country in another country (the host country) involving as subjects residents of the host country:

> In short, ethical review in the external sponsoring country may be limited to ensuring compliance with broadly stated ethical standards, on the understanding that ethical review committees in the host country will have greater competence in reviewing the detailed plans for compliance in view of their better understanding of the cultural and moral values of the population in which the research is to be conducted (Commentary under guideline 15).

ACKNOWLEDGMENT

Passages of this article have been adapted or excerpted from previous writings of the author (3,16,21). This work was funded in part by grant number PO1 MH/DA 56 826–01A1 from the National Institute of Mental Health and the National Institute on Drug Abuse. References to U.S. federal regulations are to the "Common Rule," which is applicable to all federal agencies in the executive branch that conduct or support research. It is customary to cite these in the form "_____CFR_____.101"; in this article the dashes are omitted resulting in the following form: "CFR.101."

BIBLIOGRAPHY

1. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, DHEW Publ. No. (OS) 78-0012, U.S. Government Printing Office, Washington, DC, 1978.
2. Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO), *International Ethical Guidelines for Biomedical Research Involving Human Subjects*, CIOMS, Geneva, 1993.
3. R.J. Levine, *Ethics and Regulation of Clinical Research*, 2nd ed., Urban & Schwarzenberg, Baltimore and Munich, 1986.
4. M.B. Abram and S.M. Wolf, *N. Engl. J. Med.* **310**, 627–632 (1984).
5. A.R. Jonsen, *Ann. Intern. Med.* **99**, 261–264 (1983).
6. R.J. Levine, *Hastings Cent. Rep.* **16**(4), 32 (1986).
7. T.L. Beauchamp and J.F. Childress, *Principles of Biomedical Ethics*, 4th ed., Oxford University Press, New York, 1996.
8. W.K. Frankena, *Ethics*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1973.
9. G. Outka, *Perspect. Biol. Med.* **18**, 185–203 (1975).
10. J. Rawls, *A Theory of Justice*, Harvard University Press, Cambridge, MA, 1971.
11. R.J. Levine and K. Lebacqz, *Clin. Pharmacol. Ther.* **25**, 728–741 (1970).
12. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Making Health Care Decisions: The Ethical and Legal Implications of Informed Consent in the Patient-Practitioner Relationship*, Stock No. 040-000-00459-9, U.S. Government Printing Office, Washington, DC, 1982.
13. J. Katz, *Experimentation with Human Beings*, Russell Sage Foundation, New York, 1972.
14. H.C. Kelman, *J. Soc. Issues* **33**, 169–195 (1977).
15. S. Bok, *Secrets: On the Ethics of Concealment and Revelation*, Pantheon, New York, 1982.
16. R.J. Levine, *Kennedy Inst. Ethics J.* **4**(2), 93–98 (1994).
17. U.S. Department of Health and Human Services, *Fed. Regist.* **59**, 11146–11151 (1994).
18. J.F. Childress, *Hastings Cent. Rep.* **6**(6), 21–27 (1976).

19. R.J. Levine, *Int. J. Pharm. Med.* (in press).
20. R.J. Levine, in S.S. Coughlin and T.L. Beauchamp, eds., *Ethics in Epidemiology*, Oxford University Press, New York, 1996, pp. 257–273.
21. R.J. Levine, in H.Y. Vanderpool, ed., *The Ethics of Research Involving Human Subjects: Facing the 21st Century*, University Publishing Group, Frederick, MD, 1996, pp. 235–259.
22. R. Macklin, in G.J. Annas and M.A. Grodin, eds., *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, New York and Oxford, 1992, pp. 240–257.

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, ETHICS, RESEARCH ON HUMAN EMBRYOS

RONALD GREEN
Darmouth College
Hanover, New Hampshire

OUTLINE

INTRODUCTION

Background

The development of the technology of in vitro fertilization (IVF) during the 1970s by Edwards, Steptoe, and colleagues (1,2) made possible the first systematic study of the live, developing human embryo from fertilization onward. At the same time the relatively low success rates of IVF increased demand for more systematic research on fertilization and embryo development. Together these factors have made the issue of embryo research increasingly important in law, ethics, and public policy. Embryo research is also one of the most controversial issues in biomedical ethics and law today. Wherever it has been discussed, there have been significant disagreements about the moral status of the embryo and what constitute legitimate reasons for putting it at risk.

DEFINITIONS

Embryo

The term "human embryo" is used in many ways, some opposed to one another. In ordinary speech it is used to refer to the developing human being following conception. As defined in medical literature, the term is usually applied more strictly to the product of conception from the end of the second week after fertilization to the end of the seventh or eighth week when the fetus is said to exist (3). Before two weeks, the terms zygote (the one-cell conceptus), morula (= mulberry), and blastocyst are used for the developing entity in its various stages.

In the context of ethical and legal discussions of the human embryo, the term usually refers to the product of conception (zygote, morula, and blastocyst respectively) during the first two to three weeks of development outside the womb. Normally this means an embryo that has been created by IVF, although it can also refer to a fertilized ovum that has been flushed from a uterus shortly after conception and kept alive in vitro for purposes of study. The defining features of this entity are its existence ex utero and its early stage of development, usually before or just up to the first processes of cellular differentiation, tissue formation, and the appearance of rudimentary bodily form. In both respects, embryo research differs in U.S. law from fetal research, which typically involves an embryo or later stage fetus in utero or following

abortion (4). In the mid-1980s the term "pre-embryo" was introduced for the human embryo at the earliest stage of development (5), although some have objected to it as possibly finessing the discussion of complex moral questions (6,7).

The National Institutes of Health's (NIH) Human Embryo Research Panel used the term "preimplantation embryo." This is accurate since the embryo involved in research has not yet been transferred back to a womb for implantation and, in current research, is also at the earliest or "preimplantation" stages of development. However, since implantation normally begins at 6 to 7 days in vivo, it is possible that, as our ability to sustain embryos in vitro progresses, some human embryo research, as defined here, will involve "preimplantation" embryos in the first sense (embryos that have never implanted in a womb) but not in the second sense (embryos that are less than 6 to 7 days old).

Conception/Fertilization

Further complicating the definition of the embryo is the question of what is meant by the term "conception" or "fertilization." The embryo is not regarded as coming in existence until after sex cells have joined at conception/fertilization, but conception/fertilization occurs over a period of time (8). It spans a period of at least 22 hours from the sperm's initial contact with the outer membrane (*zona pellucida*) of the egg to syngamy, the alignment on the mitotic spindle of the chromosomes derived from the male and female pronuclei (9). The first appearance of a new diploid nucleus within its own nuclear membrane occurs only after the first embryonic cell division (cleavage) three or more hours later, and the first activation of paternal genes occurs only after the second cleavage division, 12 or more hours later. A specific research program beginning with conception/fertilization therefore might involve gametes or an embryo, depending on which definition of conception/fertilization is accepted and where the cutoff point for the study is established following the admixture of sperm and egg.

In legal jurisdictions where embryo research is banned or stringently controlled, this definitional matter can be of great importance. Although the issue is often ignored in most legislation, two jurisdictions exhibit contrasting approaches. Australia's state of Victoria does not define fertilization, but its law banning human embryo research permits an exception for research "from the point of sperm penetration prior to but not including the point of syngamy" (10). Britain's Human Fertilisation and Embryology Act is quite specific, defining the embryo as present only when fertilization is complete at "the appearance of a two cell zygote" (11). These two items of legislation evidence how it is possible to select different moments in the process of fertilization. The Australian legislation also shows how important this definition can be in determining whether some types of contraceptive research aimed at blocking fertilization are permitted (12). For the balance of this discussion, fertilization will be taken to mean sperm penetration of the egg, since this is the reference point used in most chronological accounts of embryological development.

Human

The term "human" in the definition of the embryo also raises questions. Should a parthenote, an ovum artificially stimulated to begin the earliest stages of cell division, be considered a human embryo in this sense? Is an animal-human hybrid resulting from the fertilization of a human by a nonhuman gamete a *human* embryo? Are chimeric or transgenic embryos, involving the admixture of human and nonhuman embryonic cells or human and nonhuman genetic material, human embryos? Is a cloned embryo, one produced by the insertion of a nucleus from one cell into an enucleated egg cell, a human embryo (13)? Many persons who object to embryo research will predictably find research involving these other forms of embryos to be morally offensive. Current U.S. law, for example, prohibits federal funding for research on parthenotes as well as human embryos created by cloning. For this reason, research on embryos resulting from parthenogenesis, cloning, or transgenic manipulations involving nuclear human DNA should probably be considered under the heading of human embryo research.

Description

As defined, the human embryo involved in research, ranges in size from a one-cell fertilized ovum to an embryo of several thousand cells. In the words of Jones and Tefler, "The recurring motifs of embryonic development are: gradually decreasing potentiality, increasing determination and differentiation, and increasing complexity and interaction. This development occurs smoothly rather than in quantum leaps and its object is the transformation of a single fertilized ovum into a complex organism" (14, p. 45). During the earliest cleavage stages, the cells (or blastomeres) exist in a small, loosely packed mass. Each cell is undifferentiated and totipotent: taken from the cellular mass, it has the capacity to develop on its own into a full human being. Twinning can occur spontaneously at this stage, and the removal of one or more blastomeres does not interfere with the embryo's normal development. As the number of cells continues to increase, some specialization occurs. The outer layer of cells forms the trophectoderm, which at implantation give rise to the tissues that begin formation of the placenta (15). Inside the outer layer of trophectoderm cells, fluid accumulates to form a cavity, the blastocoele, and the resulting entity at 4 to 5 days following fertilization is called a blastocyst. By 6 or 7 days, when implantation normally has begun in vivo, the embryo consists of approximately 100 cells and is roughly 130 μm in diameter. Outer or trophoblast cells surround an inner group of about 20 to 30 undifferentiated cells.

One week later, by 14 days development, two fluid-filled cavities have formed, the amniotic sac and yoke sac, with a two-layered embryonic disk about 0.5 mm in diameter between them. The cells of this disk, now about 2,000 in number, remain undifferentiated. At this time the process of gastrulation begins with the appearance of the primitive streak and the establishment of left-right, head-tail orientation (16). From this point onward, the embryonic disk is committed to forming a single individual with twinning no longer possible unless two

primitive streaks have formed (17). Over the next 3 days, cells migrate to form three layers within the embryo and become pluripotent rather than totipotent: They are able to form broad categories of specialized tissues and organs but no longer an entire individual. At 17 days the primitive nervous system begins to develop with formation of the neural plate that soon develops into the neural tube. By the beginning of the fourth week of development, the neural tube closes and begins to differentiate into distinct regions of the nervous system.

In the two weeks from fertilization to the appearance of the primitive streak, therefore, the human embryo as discussed in this context is marked by very small size, the absence of bodily form, and (with the exception of placental material), the lack of differentiated tissues, organs, or a nervous system. In all these respects it differs from the postgastrulation embryo and the fetus. Currently it is not possible to sustain a human embryo in vitro until the point of gastrulation, so embryo research, as discussed here, usually refers to an undifferentiated entity of this sort. This may change in the future. Some kinds of embryo research that have been proposed, such as research to find genetic or biochemical markers for gastrulation, may extend human embryo research into the third week of development.

WHY CONDUCT HUMAN EMBRYO RESEARCH?

Limits of Animal Models

Many features of embryological development have been conserved in the course of evolution. It is now known that among vertebrates the basic principles of embryonic development, and even the genes regulating it, are similar. This means that much valuable research can be done on animals without the need to use human embryos. Nevertheless, there are significant limits to animal models. Early rubella vaccination research in monkeys indicated that the vaccination did not cross the placenta, yet subsequent research showed otherwise, rendering the vaccine unsafe for human use (18,19). Subtle matters like these can have dramatic effects on the safety of drugs or procedures when findings based on animal models are applied to human beings. From a scientific perspective some human embryo research is therefore necessary.

Assisted Reproduction

There are many areas where human embryo research can be of value. One, already suggested, is research aimed at enhancing the efficiency and safety of assisted reproductive technologies (ARTs). Despite nearly 20 years of clinical utilization and some strides forward in improving success rates, the expense and emotional toll of these procedures are enormous (20–22). Multiple safety concerns have also been raised. When used on healthy women, the drugs that stimulate development of multiple follicles only rarely have serious immediate side effects, but there is an unresolved controversy about the long-term impact of multiple cycles of superovulation on a woman's health, particularly cancer risks (23–25). Little is known about the effects of these drugs on the ova or embryos subjected to them, although it has been suggested that they may be implicated in the high rates of chromosomal anomalies (aneuploidies) found in human occytes used in IVF (26).

The need to transfer multiple embryos to enhance the chances of a pregnancy carries considerable risks in its own right. In Europe and the United States, IVF has resulted in an epidemic of higher order multiple births with its associated toll of miscarriage and prematurity (27). A better understanding of normal embryological development and improved ability to identify "implantation competent" embryos can help reduce the need for multiple embryo transfers. Because of a lack of resources for coordinated multi-center studies of assisted reproductive techniques, some newer procedures like Intracytoplasmic Sperm Injection (ICSI) have been introduced with little or no previous research into their safety for the resulting offspring. One recent study indicates that there is a slight but significant increase in the rate of spontaneous sex-chromosome anomalies among children born as a result of ICSI as compared with the general neonatal population (28). Here, as elsewhere, studies utilizing human embryos can help answer safety questions and improve the success rates of existing procedures.

Embryo research may also lead to entirely new methods of assisted reproduction. One promising set of technologies involves the cryopreservation of immature oocytes (or of ovarian tissue) followed by in vitro maturation and fertilization (29–32). Currently, multiple follicles must be matured in vivo, exposing a woman to the potent drugs used for this. Such stimulation is hazardous to women with polycystic ovarian disease, estrogen sensitive cancers of the breast, or other estrogen-sensitive disorders (31). If oocytes could be matured in vitro, these risks could be eliminated. A host of new donor sources of ova might also be made available, including women who would not wish to be exposed to stimulatory medication. Research on cryopreservation of ova is currently underway (32). Its development, along with in vitro maturation, would provide new reproductive options for women facing cancer treatments or other impediments to the future utilization of their eggs. All these promising techniques require human embryo research involving fertilization and verification in vitro of normal development.

Implantation Research

Many pregnancies fail because fertilized ova never implant in the womb. Improved understanding of the complex process of implantation can help improve pregnancy outcomes for many infertile women. Embryo research in this area can also assist our understanding of related biological processes. Implantation involves the penetration of host tissues by a rapidly growing and highly invasive foreign body. Close analysis based of the growth factors and gene expression associated with implantation can advance our understanding of similar processes that occur in cancer and tumor metastasis. Improved insight into how both fertilization and implantation occur can also lead to new techniques for preventing them. This promises development of improved contraceptive methods.

Normal/Abnormal Development

Embryo research can advance our ability to recognize key events or processes associated with normal or abnormal development. Many serious birth defects and genetic disorders that express themselves months or years following birth have their beginning in the earliest phases of embryogenesis, when cells are rapidly dividing and the basic plan of the body is being established by the operation of genes that act for only short periods of time. It is well known, for example, that viral diseases like rubella can have a devastating effect on early development. The recent discovery of the important role of folic acid deficiency early in pregnancy in contributing to neural tube defects is another example. Embryo research aimed at deepening our understanding of the patterns of normal and abnormal development may thus lead to new preventative measures or therapies.

Genetic Diagnosis

Human embryo research can help improve the newly developed technology of preimplantation genetic diagnosis (PGD). This involves coupling an IVF procedure with genetic biopsy of a single cell (blastomere) from each of the resulting embryos. It permits parents whose offspring are at risk for a genetic disease to avoid transferring embryos that exhibit the mutation or abnormality (33). Although currently much more expensive than other prenatal testing techniques like chorionic villus sampling (CVS) or amniocentesis, PGD offers parents an alternative to later term genetic abortion which is needed to avoid a birth when CVS or amniocentesis are employed.

Cell Differentiation

Human embryo research can contribute to our understanding of the complex processes of cell differentiation and can lead to new techniques for tissue transplantation and repair. Mouse research has already led to the development of techniques for producing in that species immortalized, pluripotent embryonic stem cell lines (34). In this research the inner cell mass of fertilized embryos is dissociated into single cells and dispersed into another dish with a rich culture medium. The embryonic cells continue to grow rapidly and indefinitely. Because of their low immune competence and pluripotentiality, they might be inserted back into other embryos or more mature individuals to replace defective or missing nerve, blood, skin, bone, and germ cells. Recent research by James Thomson and others suggests that this technology can be applied to the development of immortalized, canonical human stem cell lines (35,36). Somewhat differently, nuclear transfer or cloning technologies at the embryonic level have been shown to be an efficient means of introducing altered genes into every cell of the resulting organism. Employed by Wilmut and others to develop the transgenic sheep Polly, this technology also holds out the prospect of new gametic or embryonic gene therapies that might eliminate the need to discard embryos or genetic abortion (37). Human embryo research remains the essential precondition for the development and clinical deployment of all these technologies.

LAWS AND REGULATIONS

Other Countries

A review of existing national laws and regulations reveals significant disagreements about the acceptability of research involving embryos (38,39). Norway prohibits all research on fertilized ova (40); Australia's states of Victoria (41) and Western Australia (42), Austria (43), Germany (44), and Switzerland (45) forbid all embryo research other than that aimed at enhancing the survival of the embryo being studied or used in an IVF procedure. These restrictions rule out research on embryos remaining from infertility procedures that are otherwise destined to be discarded. French law specifies that research "may not be harmful to the embryo" (46), while Denmark permits embryo research only when its purpose is "to improve in vitro fertilization in order to bring about pregnancy" (47). Spain permits research employing nonviable embryos remaining from infertility procedures (48), while Sweden permits embryo research in very general terms (49). Italy and Greece have virtually no legislation in this area (39).

Britain's legislation, the Human Fertilisation and Embryology Act 1990 (50), is the most extensive and most permissive, including a provision allowing the deliberate fertilization of ova for research purposes. Similar legislation was proposed in Canada but much more stringent rules were subsequently enacted (51,52). European disagreements about embryo research are reflected in article 18 of the Convention on Human Rights and Biomedicine of the Council of Europe. This seeks to accommodate very diverse positions on embryo research by calling vaguely for "adequate protection of the embryo" when it is used in research. The convention does, however, explicity prohibits the creation of embryos for research purposes only (53). Because of its different legislation on this matter, Great Britain has entered a reservation against this article, rendering it inapplicable in that country.

Areas of Agreement

Despite the differences there are some broad areas of agreement in this legislation. These include strict requirements of informed consent for gamete or embryo donors; discouragements or prohibitions on the commercialization of gametes or embryos; and widespread prohibition of cloning, the creation of animal-human hybrids or chimeras, or the transfer of human embryos to animal wombs. Where embryo research is permitted, as in Britain and Sweden, a 14-day time limit on such research has been imposed. Several countries have also established special regulatory bodies that must approve or license human embryo research protocols. Britain's Human Fertilisation and Embryo Authority is an example. Arising out of the extensive debate occasioned by the Warnock Committee Report (54), it is charged with the extensive oversight of infertility clinics and the licensing of specific embryo research programs.

U.S. State Law

In the United States only 10 states have legislation on embryo research, nine of them very restrictive (55).

Louisiana's law defines the embryo as a "juridical person" and states that it is to be used in research "solely for the support and contribution of the complete development of human in utero implantation" (56). Federal courts have raised questions about whether this kind of restrictive legislation violates the constitutionally protected right to privacy, and there are also questions as to whether it opposes the right to freedom of expression (as this pertains to scientific research) (57), but the matter has not yet reached the Supreme Court.

Early Federal Initiatives

At the national level there have been significant regulatory and legal initiatives. In the late 1970s, shortly after the development and introduction of IVF, a special body, the Ethics Advisory Board (EAB), was formed by an act of Congress to review and provide guidance for federally funded research on the human embryo. In late 1979 the EAB issued its report containing a broad permission for such research under federal auspices subject to guidelines and limitations indicated in the report and to be implemented by the Board itself (58). However, before the EAB's recommendations could be put into effect, there was a change in adminstrations. During the Reagan and Bush years, funding for the EAB and nominations to its membership, were halted. The legal requirement of EAB approval of all such research and the absence of an EAB created a de facto moratorium on funding for embryo research in the United States.

Human Embryo Research Panel

In June 1993 Congress passed legislation nullifying the earlier requirement of EAB approval (59). To provide ethical guidance for this area, NIH established a special body, the Human Embryo Research panel. Beginning work in January 1994, the Panel held five monthly meetings in Washington open to the public and issued its report in September of that year (60). In broad terms, the report's recommendations paralleled existing legislation in Great Britain and proposals under consideration in Canada (51). Responding to its charge, the Panel divided research into three areas: (*1*) acceptable for federal funding, (*2*) unacceptable for federal support, and (*3*) warranting additional review.

In the first category the Panel placed research aimed at enhancing the safety and efficiency of infertility procedures, at preventing disease conditions arising during prenatal development, and at improving our understanding of the human embryo. Permitted research was governed by a series of stringent guidelines. They included the requirements of demonstrated scientific validity and merit, informed consent on the part of gamete or embryo donors, a prohibition on the purchase or sale of gametes or embryos, and a 14-day age limit on the use of embryos in research.

The category of research that was unacceptable for federal funding included research on embryo cloning (whether by splitting existing embryos to multiply the number of genetically identical embryos, or by somatic cell nuclear transfer to replicate an existing individual) where transfer to a womb was intended; research involving the fertilization of fetal oocytes with intent to transfer; parthenogenesis research where transfer was intended, and the creation of animal-human chimeras. Among those items placed in the category of warranting additional review were cloning by embryo splitting and the use of fetal oocytes where no transfer to a womb was intended.

Subsequent Developments

Among the most controversial of the Panel's proposals, but paralleling existing regulations in Great Britain, was a permission to fertilize oocytes for research purposes without the intention to transfer. Late on December 2, 1994, following a morning meeting in which the Advisory Council to the NIH director unanimously accepted the Panel's report, President Clinton issued a directive overruling this recommendation. Eventually this limited dissent from the Panel's recommendations was overtaken by congressional initiatives barring federal funding of any embryo research that threatened the embryo's survival (61). In May 1999, in the wake of several new research reports showing the possibility of establishing canonical embryonic stem (ES) cell lines from human embryos, the National Bioethics Advisory Commission, once again recommended establishing federal support for human embryo research (62).

Thus the United States during this period has exhibited roughly the same polarization of opinion and legislation on this issue as has occurred in Europe. In one important respect, however, the United States differs from Europe, including Great Britain, where regulations govern all research on the embryo. In the United States, federal legislation and guidelines affect only federally funded research. To date, there has been no federal legislation prohibiting or limiting research with private funds.

ETHICAL ISSUES

Range of Views

The legal turmoil over human embryo research reflects deep ethical disagreements. Foremost among these is disagreement about the moral status of the human embryo in the earliest stages of development. At one end of the spectrum of opinion are those who regard the embryo as fully a human being meriting all the protections due any other human research subject. This position has been publicly defended by the Roman Catholic Church (63,64). At the other end of the spectrum are those who see the embryo as meriting little or no protection and who believe that the primary consideration in human embryo research is whether it benefits or harms adult human beings or children. Someone holding this view, for example, might argue that if embryo research could improve contraceptive alternatives for women or reduce the risks of current methods, then it is not only morally permitted but morally required. Between these polar positions lie those who, for different reasons, accord the embryo some measure of moral respect and believe that any kind of harmful research on the embryo requires stringent moral justification. This issue of the moral status of the

preimplantation embryo is at the center of these debates. The question becomes on what grounds a being or entity achieves moral status.

Genetic Individuality and Moral Status

Some who defend the full moral status of the embryo answer this question by focusing on the embryo's possession of what they regard as the essential biological criterion of humanness in a moral and spiritual sense: a diploid human genome with the inherent and active potential to develop on its own, under normal conditions, into a child or adult human being (65,66). Those who hold this view believe that biological humanness in this sense represents a "bright-line" distinction providing maximum protection for all who are human from possible abuses that may come about if humanness is defined in less determinative ways (67).

Biological Problems

The full moral status view, although seemingly clear, is challenged by biological knowledge on several fronts. One problem has to do with the so-called moment of fertilization, which marks the cherished bright line separating embryos, as protectable, from gametes that are regarded as having no greater moral status than other expendable bodily tissue. Strictly speaking, there is no single "moment" of fertilization but a series of events — a process. This raises the question of whether references to biological phenomena really do provide the bright lines that defenders of this position seek.

A second problem has to do with the realities of twinning and recombination. In a small number of cases, single embryos spontaneously split into one or more embryos, resulting in the birth of monozygotic (identical) twins if the pregnancy proceeds to term (68). Different embryos also sometimes combine to form chimaeras containing one or more fused genomes (69). For some ethicists and theologians, these rare but naturally occurring events challenge the notion that the appearance of a single, discrete genome following fertilization is the marker for moral and spiritual individuality (8,70). The possibility of twinning and the fact that most cells of the early embryo develop into nonembryonic supporting tissue also lead some to question whether we can speak of the early embryo as a single "being" with the "potential" to develop into an adult person (71). Others question whether the high degree of human intervention necessary for IVF procedures does not efface any distinction between embryos and gametes in terms of their "inherent" potential to develop to maturity (72). The fact that cloning technology renders any cell in the body able to become an embryo further complicates the position of those who stress the presence of a diploid genome and raises further questions of what "potentiality for development" means in this context.

Finally, the clarity of this biological view is troubled by the very large proportion of embryos — estimates are as high as two thirds to three quarters of all fertilized ova (73,74) — that do not successfully implant in the womb. If the likelihood of development to term, an important consideration on some accounts of potentiality (67), takes a significant upward turn as gametes becomes an embryo, still another increment in likelihood of birth takes place at implantation. This raises the question of whether this point, or perhaps the later establishment of individuality at gastrulation (after which twinning cannot occur), is not the better "bright" line on which to base protected humanness.

Personhood Views

In addition to these biological challenges, critics of the position that locates full moral protectability at fertilization also offer many objections of a philosophical nature, some of which also appear in the context of the abortion debate. The issue, some of these critics contend, is not biological humanness but moral "personhood." By this they mean the status of being a moral subject worthy of all the rights and protections normally accorded human beings (75). They point out that some individuals who are biologically human (e.g., adults who are brainstem dead or anencephalic infants) may not be persons in a moral sense and may ethically be treated with less than equal respect. It is also possible that in the future we will encounter intelligent life forms that are not biologically human but that nevertheless are persons in a moral sense.

Those holding this view tend to espouse a variety of distinct, but sometimes overlapping, positions regarding the qualities that are needed for personhood. Some believe that highly developed qualities, such as consciousness and a sense of self are needed (76–78). Others stress the beginnings of brain activity or brain function, arguing that if the cessation of brain activity at the end of life is an appropriate marker of death, it should also be a marker for the commencement of moral personhood (79–82). Still others point to sentience, the ability to experience pain (83,84). Although those adhering to the brain activity or sentience views may come to a less permissive position on early term abortion than those holding the consciousness view, all those holding these views tend to agree about the status of the early embryo, whose lack of differentiated nervous system tissue makes it unreasonable to regard it as capable of feeling or thinking.

Symbolic Concerns

Complicating this typology of positions, are the views of those who believe that the significant moral issue here is not the status or rights of the embryo so much as the implications of its treatment for society as a whole. Some who hold this view fear that funding embryo research creates a "slippery slope" that could lead to undermining current protections for human subjects in research, erode respect for persons with disabilities, or encourage eugenic practices (85). Others emphasize the symbolic importance of the embryo, rather than its intrinsic moral claims, and ask how its treatment (or mistreatment) might affect respect for life generally (86,87).

PUBLIC POLICY

The status of the embryo and the question of what degree of respect it deserves are difficult to settle partly

because they involve intensely personal matters of moral conviction or religious belief. Where law or public policy is concerned, however, debate must focus primarily on those considerations that are appropriate for setting policy in a religiously and morally pluralistic democracy. As one writer puts it, "law is not really concerned with the enforcement of morality but rather with providing a framework of peace and order within which people may exercise their personal liberty ... and make their own personal moral choices and engage in what John Stuart Mill calls their own 'experiments in living' " (88, p. 149).

Public Reasoning

This context imposes limits on what can be brought to public debate about issues significantly affecting the lives, welfare, and basic liberties of citizens. The philosopher John Rawls argues that basic public policy in such contexts must employ "public reasoning" that is understandable in terms that are not dependent on particular religious, theological, or philosophical perspectives. It should appeal "only to presently accepted general beliefs and forms of reasoning found in common sense, and the methods and conclusions of science when these are not controversial" (89, p. 224). From this perspective, religious teachings that are unable to sustain themselves without resort to special metaphysical premises belong to the realm of private decision and should not direct public policy where basic matters of citizens' welfare and liberties are concerned.

Two considerations qualify this conclusion. One is that some religiously based views can be articulated in terms of widely shared moral values. When this is done, and the arguments are persuasive, there is no reason to set aside a view merely because it originates in a religious context or is held by a community of religious believers. Second, even when a religious position cannot commend itself to a wider society, it is good public policy to respect this view to the extent that doing so is compatible with the protection of public healthy and safety.

Policy Analysis

Because there is no clear consensus on the moral status of the embryo itself, the most relevant arguments surrounding embryo research from a public policy perspective have to do with the indirect and symbolic impacts on society and people's respect for the sanctity of human life. Will using embryos in research start us down a slippery slope and contribute to the neglect or injury of other protectable human subjects? Will use of embryos lead to a "cheapening" of the value of human life? Will embryo research damage the way we regard human parenting or diminish the protection we accord our young? These questions are all enormously difficult to answer. If we keep in mind that similar questions have been raised with the advent each new reproductive technology, we see how marked this whole area is by deep, if largely speculative, concerns.

On one side of the balance, therefore, are a set of hard-to-assess symbolic concerns to which we might add the very determined opposition of a large number of citizens, some influenced by personal religious beliefs. On the other side is the great promise human embryo research holds out for reducing illness and improving human health.

Policy Conclusions

Taking all these considerations into account, and trying to balance the avoidance of immediate harms against symbolic and other concerns, expert panels like the Warnock Committee, the NIH Human Embryo Research Panel, and Canada's Royal Commission on New Reproductive Technologies have all recommended permitting embryo research under stringent control and limitations, including a 14-day time limit. Apart from Great Britain, however, these expert public assessments have had little impact. Political opposition to embryo research, some of it influenced by religious views or traumatic national experiences with biomedical research and eugenics, as in the case of Germany, has had more effect than the kind of moral analysis and public reasoning employed by expert panels (90).

SPECIAL QUESTIONS

Those who accept the possibility of embryo research must resolve a series of additional ethical questions. Substantial consensus exists about the answers to some of these questions, while others remain very controversial.

Time Limits

As mentioned above, one area in which there has been widespread agreement is the need to set a limit, usually 14 days, on the time during which research can be conducted after fertilization. Since embryo development in vitro usually proceeds more slowly than in vivo, this ensures that embryos used in research have not entered the phase of gastrulation and have not yet developed the primitive streak. It is important to note that the 14-day limit should not be misconstrued, as it has sometimes been (91), as a statement about the definitive commencement of moral personhood at this time. Rather, those who accept this limit usually hold that the embryo cannot definitively be held to be a person before this point, although they may disagree about the time when personhood or enhanced protectability is subsequently established. It is also taken as a reasonable compromise between the moral claims surrounding the embryo and the needs of researchers.

Understanding the 14-day point as a reasonable but not necessarily absolute limit, the NIH Human Embryo Research panel urged that some research beyond this point (in vitro) be permitted in exceptional circumstances when the goal was to assist in the identification in the laboratory of the appearance of the primitive streak. Some who regard the beginning of gastrulation as a morally definitive step would not agree with this exception. Others see exceptions of this sort as propelling us down a slippery slope leading to the abuse of other vulnerable research subjects (85).

Intent to Transfer

Many believe that the question of whether or not researchers intend to transfer an embryo to a uterus is of great moral consequence in evaluating research proposals. Where transfer is in prospect, another class of human subjects becomes involved: children born as a result of these procedures. Since born children are recognized in law and ethics as having a right to protection from injuries inflicted before their birth, this means that researchers must demonstrate that manipulations of the embryo impose no greater risks than would be encountered in a normal pregnancy or, where the purpose is helping a couple have a child, of accepted assisted reproductive procedures. If also means that new assisted reproductive techniques should not initially be applied to human embryos that are to be transferred with the aim of establishing a pregnancy. Before transfer is considered, intermediate studies should be conducted on human embryos that are not intended for transfer. This approach was adopted by Steptoe and Edwards in their initial development of the clinical protocols used today in all IVF programs throughout the world (92,93). Risks to children born as a result of cloning procedures also played a large role in the U.S. National Bioethics Advisory Commission's June 1997 recommendation of a five-year legal prohibition of all attempts at cloning a human being (although the Commission did not at that time bring under review the existing ban on federal funding for the embryo research that would be needed to reduce these risks) (94).

Philosophical Complexities

There are interesting philosophical complexities associated with this high standard of safety for research involving embryos that are to be transferred. Some have argued that a child who would not otherwise exist cannot be morally wronged by procedures needed to bring it into being, even if these same procedures seriously impair its health or well-being (95,96). Without these harmful procedures, it is contended, the child would not have been and, at least up to the point where the harm is not so great as to make it reasonable to wish to not be alive, the child cannot morally complain about its treatment. Those who accept this argument might permit research on embryos for transfer that jeopardizes the well-being of a resulting child if this is the only way that the child can be brought into being.

Not everyone accepts this line of reasoning. Some doubt that being brought into existence should be reckoned as a benefit, and maintain that it is always the obligation of parents (and those who assist them in having children) to strive to ensure that the child has a healthy start in life (97,98). Fortunately these issues do not have to be resolved to establish public policy in this area. Governmental agencies supporting research are chartered to promote public health and to protect the welfare of children and other vulnerable subjects involved in research. Such agencies therefore have good reason to insist on the most stringent guarantees of safety where research anticipates transfer of embryos to a womb.

These protections are less important where transfer is not intended and the embryo is destined to be discarded in any case. It is useful to note here a difference between embryo research and fetal research. In U.S. law, fetuses intended for abortion are accorded the same level of moral protection in research (a requirement of no harm or minimal harm) as fetuses that are intended to be brought to term (99). However, it is not inconsistent to treat embryo research differently because fetal research poses a special set of problems. If researchers were permitted to perform harmful studies on the fetus of a woman intending abortion and she subsequently changes her mind and insists on continuing with the pregnancy, this would create a situation of unavoidable harm for the resulting child because it is ethically unacceptable to force her to terminate the pregnancy. A similar conflict situation cannot arise where embryo research is involved.

Donor Rights and Donor Sources

Donors of the gametes or embryos used in research are an important moral constituency whose rights and welfare must be respected. Three considerations shape thinking in this area: the requirement of informed consent, concern for the impact on family and the succession of generations in the use of certain donor sources, and the implications for society when financial incentives create a market in gametes or embryos used in research.

The requirement that researchers obtain the full, free, and informed consent of those who donate their reproductive material for research purposes means, among other things, that consent must be specific to the type of research proposed. Because donors can reasonably find one kind of embryo research morally acceptable and another repugnant, general consent to the use of gametes or embryos is not appropriate. Behind this requirement of specific consent lies the understanding that gametes and embryos are different from other bodily tissues. They contain the possibility of and association with people's dreams of offspring. Some individuals might be willing to donate reproductive material where transfer is intended, hoping to directly assist others to start a family. Other donors might be unwilling to see children of theirs raised outside the context of their family but would be willing to donate gametes or embryos for research not involving transfer. Despite their different feelings in this regard, all donors have the right to freely understand and consent to how their reproductive material is used in research.

Because oocyte donation for purposes of embryo research now requires follicle maturation in vivo, these procedures are also not without risk to the donor. Although these risks are routinely accepted by women who wish to donate oocytes to an infertile couple, some believe they are disproportionate to any benefits in research contexts where no transfer to a womb is anticipated. The NIH Human Embryo Research Panel came to this conclusion, although it permitted such donation by women about to undergo pelvic surgery who were properly informed about the additional risks of drug stimulation and superovulation. If and when embryo research goes forward in the future, Institutional Review Boards and other groups assigned to assess research risks will have to balance the values involved in permitting donation (including respect for donors' wishes) against the specific

risks to donors and their vulnerabilities to coercion or pressure from researchers or physicians. The clinical IVF context merits special attention in this regard. Although women undergoing treatment for infertility are a constituency that has a significant stake in the benefits arising from embryo research, this context also creates special pressures that run counter to the requirement of full, free, and informed consent. Among these are the hope (or promise) of price reductions or additional support from one's treating physician in return for gamete or embryo donation.

Concern for the impact on family and the succession of generation suggests limits on donor sources. For example, although the use of cadavers as sources of oocytes might be appropriate for research where no transfer is intended (assuming that the women or her appropriate surrogates consent), this source raises many questions where transfer is intended. Among other things we can ask whether it is ethical to bring a child into the world whose genetic mother is deceased. Apart from the psychological risks of this practice, the child conceived in this way is possibly cut off from access to the deceased parent's medical history. Similar questions arise in connection with the idea of using aborted female fetuses as a source of oocytes for in vitro fertilization and embryo research. No only does this undermine the idea of donor consent (a problem not entirely eased by the consent of the abortus's mother to the use of her fetus in this way), it also implicates embryo research in the abortion and fetal tissue debate. Where transfer is intended, it raises the disturbing prospect, in terms of our ideas of generational succession, of the birth of a child whose genetic mother was aborted. For all these reasons, at least where transfer is intended, this donor source is best avoided.

Commercialization

It is also wise for government agencies involved in funding embryo research to avoid creating a market in gametes or embryos used for this purpose. In the United States substantial commercialization already exists in the area of egg or embryo donation, but payment for gametes or embryos used in embryo research poses special problems. Introducing government funding for the purchase of eggs or embryos would create a large source of revenue that might be particularly attractive to poor women, a constituency largely untouched by present commercial practices. In view of the special risks associated with egg donation and the deep symbolic importance of reproductive material, the purchase and sale of eggs and embryos is a practice that could tarnish the whole field of embryo research. This does not preclude the compensation of donors for their expenses in participating in a research program.

"SPARE" VERSUS "RESEARCH EMBRYOS"

Moral Issues

Probably the most vexing ethical issue in embryo research is the question of whether it is ever appropriate deliberately to fertilize oocytes for research purposes when there is no intention of transferring the resulting embryos (100). The alternative to such "research embryos" is to utilize only "spare" or "surplus" embryos remaining from infertility procedures. Although many who regard the embryo as a fully protectable being would object to the use of even surplus embryos in research, some who place substantial moral weight on the early embryo believe that spare embryos can ethically be used, since most are destined to perish anyway. This position has been subjected to multiple criticisms (101). For example, it is not clear why the fact that surplus embryos might otherwise continue in storage or be destroyed warrants their use in research. Those who believe the embryo is fully a moral person may therefore be unable to sustain a distinction between spare and research embryos and may be required by their position to oppose all embryo research.

Some who oppose the creation and use of "research" embryos base their views less on harm done the embryo than on symbolic concerns, believing that it is objectionable to use potential persons in such an instrumental way. These arguments, too, raise many questions. Among them are the question of why it should be regarded as acceptable to create more embryos than can be transferred in an IVF procedure but impermissible deliberately to create embryos for research purposes. Whatever the reasons behind it, opposition to the creation and use of "research embryos" is widespread and cuts across familiar political lines, drawing criticism even from those who accept embryo research generally (102).

Scientific Issues

In practical terms, there are many reasons for wanting to develop embryos for research purposes only. In many types of research existing surplus embryos cannot be used. Limiting research to spare embryos, for example, precludes most research involving study of the process of fertilization, since spare embryos have already been fertilized. Similarly research on in vitro oocyte maturation and oocyte cryopreservation would require fertilization and surveillance of the resulting embryos to establish the efficacy and safety of these procedures. It is true that if fertilization is defined as occurring at some point *after* the penetration the egg by the sperm, some of this research could go forward without being regarded as "embryo research" in a technical sense. This points up the importance of these definitional matters. Nevertheless, in many cases demonstrating the safety of new techniques will require the observation of the resulting embryos for some period of time.

Embryos developed in vitro during infertility procedures, often derived from older eggs, also evidence high rates of chromosomal abnormalities that may explain their inability to implant (103,104). This means that surplus embryos are a less than ideal population for broad classes of research on normal embryo development, and their use may actually produce misleading findings in many studies. To the extent that research requires a "normal" embryo population, whether for direct study or as controls, some research embryos may be needed. For all these reasons, the stakes here are high. Those who believe that

embryos merit significant moral protection will not be persuaded that these benefits justify the deliberate creation of embryos for research purposes only. However, those who oppose this idea of "research embryos" on symbolic grounds, or merely because it strikes them as offensive, will have to reexamine their opposition in the light of its longer-term and possibly important negative impact on medical progress.

CONCLUSION

Human embryo research provides a vivid illustration of the way emerging scientific and medical capabilities are continually challenging the application of established ethical ideas and norms. Despite the elaboration of a significant body of ethical and legal rules governing the use of human subjects in research, new reproductive technologies have brought to the fore an entity, the early human embryo, whose existence raises basic questions about who is a human subject and what should be the limits of the research enterprise. The political turmoil surrounding embryo research is a sign of the difficulty and troubling nature of these questions. Few areas of scientific inquiry evoke so many emotions and differences of opinion. Although prohibited in many European nations and deprived of federal funding in the United States, research involving the human embryo continues. New scientific developments that arise from embryo research or that require it will predictably sustain public interest in this scientific and ethical frontier and stimulate continued debate.

BIBLIOGRAPHY

1. R.G. Edwards, P.C. Steptoe, and J.M. Purdy, *Nature (London)* **227**, 1307–1309 (1970).
2. P.C. Steptoe, R.G. Edwards, and J.M. Purdy, *Nature (London)* **229**, 132–133 (1971).
3. *Dorland's Illustrated Medical Dictionary*, 28th ed., Saunders, Philadelphia, PA, 1994.
4. 45 Code of Federal Regulations 46.203(c).
5. A. McClaren, in G. Bock and M. O'Connor, eds., *Human Embryo Research: Yes or No?* Tavistock, London, 1986, pp. 5–23.
6. F.E. Baylis, *Bioethics* **4**, 311–329 (1990).
7. D. Heyd, *Bioethics* **10**(4), 292–309 (1996).
8. N.M. Ford, *When Did I Begin?* Cambridge University Press, Cambridge, UK, 1988, pp. 102–118.
9. K. Moore, *The Developing Human: Clinically Oriented Embryology*, 3rd ed., Saunders, Philadelphia, PA, 1982.
10. Australia (Victoria), The Infertility (Medical Procedures) (Amendment) Act 1987, 9a(1).
11. Britain's Human Fertilisation and Embryology Act 1990, Section 1(1)(a)(b).
12. S. Buckle, K. Dawson, and P. Singer, in P. Singer et al., eds., *Embryo Experimentation: Ethical Legal and Social Issues*, Cambridge University Press, Cambridge, UK, 1990, pp. 213–225.
13. *Nature (London)* **386**, 98 (1997).
14. D.G. Jones and B. Tefler, *Bioethics* **9**(1), 32–49 (1995).
15. C.R. Austin, *Human Embryos*, Oxford University Press, Oxford, UK, 1989.
16. M.J.T. FitzGerald and M. FitzGerald, *Human Embryology*, Baillière Tindall, London, 1994.
17. Ethics Committee of the American Fertility Society, *Fertil. Steril.* **53**(6, Suppl. 2), 32SS (1990).
18. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *Report and Recommendations: Research on the Fetus*, U.S. Government Printing Office, Washington, DC, 1975, reprinted in *Fed. Regist.* **40**, 33, 530 (1976).
19. A. Vaheri et al., *N. Engl. J. Med.* **286**(20), 1071–1074 (1972).
20. P.J. Neumann, S.D. Gharib, and M.C. Weinstein, *N. Engl. J. Med.* **331**(4), 239–243 (1994).
21. H.F. Hodder, *Harv. Mag.*, November/December, pp. 54–99 (1997).
22. S.G. Stolberg, *N.Y. Times*, Sunday, December 14, pp. A1, A36 (1997).
23. B.R. Spirtas and S.J. Alexander, *Fertil. Steril.* **59**, 291–293 (1993).
24. A.S. Whittemore et al., *Am. J. Epidemiol.* **136**, 1175–1220 (1993).
25. M.A. Rossing et al., *N. Engl. J. Med.* **331**(12), 771–776 (1994).
26. L. Gras, J. McBain, A.O. Trounson, and I. Kola, *Hum. Reprod.* **7**, 1396–1401 (1992).
27. P. Belluck, *N.Y. Times*, January 3, pp. A1, A9 (1998).
28. A.C. Van Steirteghem, *N. Engl. J. Med.* **338**, 194–195 (1998).
29. K.Y. Cha et al., *Fertil. Steril.* **55**, 109–113 (1991).
30. T.L. Toth et al., *Fertil. Steril.* **61**, 1077–1082 (1994).
31. C.E. Wood, J.M. Shaw, and A.O. Trounson, *Med. J. Aus.* **166**, 366–369 (1997).
32. G. Kolata, *N.Y. Times*, October 17, pp. A1, A16 (1997).
33. A.H. Handyside et al., *N. Engl. J. Med.* **327**, 905–909 (1992).
34. T.C. Doetschman et al., *J. Embryol. Exp. Morphol.* **87**, 27–45 (1987).
35. A.J. Thomson et al., *Science* **282**, 1145–1147 (1998).
36. Shamblott et al., *Proc. Nat. Acad. USA* **95**, 13726–13731 (1998).
37. G. Kolata, *Clone: The Road to Dolly and the Path Ahead*, Wm., Morrow, New York, 1998.
38. L.B. Andrews and N. Elster, *Papers Commissioned for the Human Embryo Research Panel*, NIH Publ. No. 95-3916, National Institutes of Health, Bethesda, MD, 1994, pp. 251–264.
39. L. Nielsen, in D. Evans, ed., *Conceiving the Embryo: Ethics, Law and Practice in Human Embryology*, Martinus Nijhoff Publishers, The Hague, 1996, pp. 325–338.
40. Law No. 56 of 5 August, 1994 on the Medical Use of Biotechnology, *Nor. Lovtidend*, Section 3-1; *Int. Dig. Health Legis.* **46**(1), 51–54 (1995).
41. *The Infertility (Medical Procedures) Act, 1984* and *The Infertility (Medical Procedures) (Amendment) Act 1987*.
42. *The Human Reproductive Technology Act 1991*, Preamble, paragraphs B and C; *Int. Dig. Health Legis.* **44**(4), 630–631 (1993).
43. Federal Law of 1992 (Serial No. 275), Art. I, Section 9; *Int. Dig. Health Legis.* **44**(2), 248–249 (1993).
44. Law of 13 December for the Protection of Embryos (the Embryo Protection Law), *Bundesgesetzblatt*, Part I, pp. 2746–2748 (1990); *Int. Dig. Health Legis.* **42**(1), 60–69 (1991).

45. *Genetics and Assisted Procreation*, Amendment of Federal Constitution, dated 13 August 1992, *Rec. Off. Lois Féd.* **32**, 1579, 1580 (1992); *Int. Dig. Health Legis.* **44**(4), 745–746 (1993).
46. Law No. 94-654 of 29 July 1994 on the Donation and Use of Parts and Elements and Products of the Human Body, Medically Assisted Procreation, and Prenatal Diagnostics, *J. Off. Répub. Fr. Lois Décrets* **175**, 11060–11068 (Article I, 152, 153) (1994); *Int. Dig. Health Legis.* **45**, 473–482 (1994).
47. Law No. 503 of 24 June, 1992 on the Scientific Ethics Committee System and the Examination of Biomedical research Projects, Chapter 4, Section 14, *Lovtidende* **84** (Part A), 2017–2020 (1992); *Int. Dig. Health Legis.* **43**(4), 758–760 (1993).
48. Law No. 35/1988 of 22 November 1988 on Assisted Reproduction Procedures, *Bol. Of. Estado* **282**, 33373–33378 (1988); Law No. 42/1988 of 28 December 1988 on the Donation and Use of Human Embryos and Fetuses or Their Cells and Tissues, or Organs, *Bol. of. Estado* **314**, 36766–36767 (1988).
49. Law No. 115 of 14 March 1989 Concerning Measures for the Purpose of Research or Treatment in Connection with Fertilized Human Oocytes, *Sven. Författningssamling*, March 26, pp. 1–2 (1991); *Int. Dig. Health Legis.* **44**(1), 58–59 (1993).
50. The Human Fertilization and Embryology Act 1990, *Int. Dig. Health Legis.* **42**(1), 69–85 (1990).
51. *Proceed with Care*, Final Report of the Royal Commission on New Reproductive Technologies, vol. I, Minister of Government Services, Ottawa, 1993, pp. 607–659.
52. W. Kondro, *Lancet* **347**, 1758 (1996).
53. Council of Europe, *Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine, Convention on Human Rights and Biomedicine*, Council of Europe, Strasbourg, 1996.
54. M. Warnock, *A Question of Life: The Warnock Report on Human Fertilisation and Embryology*, Blackwell, Oxford, 1985.
55. L.B. Andrews, in *Papers Commissioned for the Human Embryo Research Panel*, NIH Publ. No. 95-3916, National Institutes of Health, Bethesda, MD, 1994, pp. 297–330.
56. La. Rev. Stat. Ann. §9 : 123, 122, *et seq.* (West 1991).
57. L.B. Andrews, *Papers Commissioned for the Human Embryo Research Panel.* NIH Publ. No. 95-3916. National Institutes of Health, Bethesda, MD, 1994, pp. 303–305.
58. *HEW Support and Research Involving Human In vitro Fertilization and Embryo Transfer*, Report and Conclusions, U.S. Department of Health Education and Welfare, Washington, DC, 1979; Ethics Advisory Board, Summary and Conclusions, *Fed. Regis.* **44**(18), 35057 (1979).
59. National Institutes of Health (NIH) Revitalization Act of 1993 (P.L. 103-43, Section 121 (c)).
60. *Report of the Human Embryo Research Panel*, National Institutes of Health, Bethesda, MD, 1994.
61. HR 3019-PL 104-34, *Congressional Quarterly*, 1874, (June 19, 1996).
62. N. Wade, *N.Y. Times*, June 29, p. A13 (1999).
63. Pope John Paul II, *On the Value and Inviolability of Human Life (Evangelium Vitae), Encyclical Letter Addressed by the Supreme Pontiff*, 1995, Section 63.
64. The United States Catholic Conference of Bishops, *Documentation on Abortion and the Right to Life*, U.S. Catholic Conference, Washington, DC, 1976, p. 39.
65. R.J. White, Testimony before the National Institutes of Health Human Embryo Research Panel, June 21, 1994; reprinted as *America*, September 14, pp. 4–5 (1996).
66. H. Watt, *J. Med. Ethics* **22**, 222–226 (1996).
67. J.T. Noonan, Jr., in J.T. Noonan, Jr., ed., *The Morality of Abortion: Legal and Historical Perspectives*, Harvard University Press, Cambridge, MA, 1970, pp. 51–59.
68. Ethics Committee of the American Fertility Society, *Fertil. Steril.* **46**(3, Suppl. 1), 26S–28S (1986).
69. L. Strain et al., *N. Engl. J. Med.* **338**, 166–169 (1998).
70. R.A. McCormick, *Kennedy Inst. Ethics J.* **1**, 1–15 (1991).
71. S. Buckle, in P. Singer et al., eds., *Embryo Experimentation: Ethical Legal and Social Issues*, Cambridge University Press, Cambridge, UK, 1990, pp. 90–108.
72. P. Singer and K. Dawson, in P. Singer and K. Dawson, eds., *Embryo Experimentation: Ethical, Legal and Social Issues*, Cambridge University Press, Cambridge, UK, 1990, pp. 76–89.
73. British Royal College of Obstetricians and Gynaecologists (RCOG), *Report of the RCOG Ethics Committee on In vitro Fertilization and Embryo Replacement or Transfer*, RCOG, London, 1983.
74. C.J. Roberts and C.R. Lowe, *Lancet* **1**, 498–499 (1975).
75. H.T. Englehardt, Jr., *The Foundations of Bioethics*, Oxford University Press, New York, 1996.
76. M. Tooley, *Abortion and Infanticide*, Oxford University Press, New York, 1983.
77. M.A. Warren, *The Monist* **57**(1), 43–61 (1973).
78. B. Steinbock, *Life Before Birth: The Moral and Legal Status of Embryos and Fetuses*, Oxford University Press, New York, 1992.
79. B. Brody, *Abortion and the Sanctity of Human Life*, MIT Press, Cambridge, MA, 1975, p. 114.
80. J.M. Goldenring, *J. Med. Ethics* **11**, 200 (1985).
81. H.M. Sass, *J. Med. Philos.* **14**, 45–59 (1989).
82. R.M. Veatch, in M.W. Shaw and A.E. Doudera, eds., *Defining Human Life: Medical, Legal, and Ethical Implications*, ALPHA Press, Ann Arbor, MI, 1983, pp. 99–113.
83. P. Singer and D. Wells, *Making Babies*, Scribner's, New York, 1985.
84. P. Singer, *Animal Liberation*, Avon Books, New York, 1990.
85. J.S. Freeman, *J. Med. Philos.* **21**, 61–81 (1986).
86. J. Robertson, *Hastings Cent. Rep.* **25**(1), 37–38 (1995).
87. R. Dworkin, *Life's Dominion*, Knopf, New York, 1993.
88. M. Charlesworth, in P. Singer et al., eds., *Embryo Experimentation: Ethical Legal and Social Issues*, Cambridge University Press, Cambridge, UK, 1990, pp. 147–152.
89. J. Rawls, *Political Liberalism*, Columbia University Press, New York, 1993.
90. I.H. Carmen, *Hum. Gene Ther.* **7**, 97–108 (1996).
91. D. Callahan, *Hastings Cent. Rep.* **25**(1), 39–40 (1995).
92. P.R. Braude, V.N. Bolton, and M.H. Johnson, in G. Bock and M. O'Connor, eds., *Human Embryo Research: Yes or No?* Tavistock, London, 1986, pp. 63–82.
93. R. Edwards, in A. Dyson and J. Harris, eds., *Experiments on Embryos*, London, Routldge, 1990, pp. 42–54.
94. Report and Recommendations of the National Bioethics Advisory Commission, *Cloning Human Beings*, Rockville, MD, 1997.
95. R. Macklin, Testimony before the National Bioethics Advisory Commission, March 14, 1997, quoted in G. Kolata (37, p. 20).

96. D. Brock, *Bioethics* **9**(3/4), 269–275 (1995).
97. R.M. Green, *J. Law, Med. Ethics* **25**(1), 5–15 (1997).
98. C.B. Cohen, in N. Fotion and J.C. Heller, eds., *Contingent Future Persons*, Kluwer Academic Dordrecht, The Netherlands, 1997, pp. 27–40.
99. The Public Health Service Act as Amended by the Health Research Extension Act of 1985, Public Law 99-158, November 20, 1985, Section 498(a)(3).
100. D.S. Davis, *Kennedy Inst. Ethics J.* **5**(4), 343–354 (1995).
101. C.A. Tauer, in N. Fotion and J.C. Heller, eds., *Contingent Future Persons*, Kluwer Academic Dordrecht, The Netherlands, 1996, pp. 171–189.
102. Editorial, *Washington Post*, Sunday, October 2, p. C8 (1994).
103. Z. Rosenwaks, O.K. Davis, and M.A. Damario, *Hum. Reprod.* **10**(Suppl. 1), 165–173 (1995).
104. C.M. Strom, *J. Assist. Reprod. Genet.* **17**(7), 592–593 (1996).

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, ETHICS, RESEARCH ON VULNERABLE POPULATIONS

PATRICIA BACKLAR
Portland State University
Oregon Health Sciences University
Portland, Oregon

OUTLINE

INTRODUCTION

Adult individuals characterized as vulnerable usually are persons who, for a variety of reasons, are incapable of protecting their own interests. Biomedical research subjects may be considered vulnerable because of illness, mental disorder, or particular circumstances (1). In general, individuals are recognized as vulnerable when they are cognitively impaired, or when their circumstances subject them to intimidation and exploitation, thereby limiting their freedom to exercise autonomous choice. Some persons with mental disorders are prototypical of individuals who are recognized as especially vulnerable. Not only is their decision making compromised by illness, but often they also are socially disadvantaged and devalued for reasons of poverty, institutionalization, or stigmatization. Historically, vulnerable and devalued populations, often unwittingly, were forced to serve as research subjects in studies not relevant to their own conditions, but which benefited the health of more privileged members of society. Yet measures designed to protect these populations have come to be seen as overly exclusionary and unjust. Indeed, populations judged vulnerable are now considered to be at risk of being relegated to a class of persons for whom little or no therapeutic benefit may be available. Exclusion may reinforce their vulnerability. However, powerless and impoverished persons may be especially vulnerable to the popular therapeutic misconception that research protocols provide beneficial treatment. Inability to adequately engage in the process of voluntary and informed consent — for whatever reasons — raises the research subject's degree of risk and vulnerability. Increased regulatory protections that provide for research anticipatory planning, surrogacy, independent health care supervision, and other safeguards may allow vulnerable research subjects to be enrolled in studies that address their conditions while shielding them from exposure to harms and wrongs.

VULNERABLE POPULATIONS IN BIOMEDICAL RESEARCH

Among persons with mental illness there are those who share at least four characteristics that are considered prototypical of other devalued and vulnerable populations. First, their disorders put them at risk for loss of decision-making capacity. Second, they are likely to be poor. Even though persons may be afflicted with mental disorders regardless of socioeconomic level, there is a significant association between poverty and serious mental illness due to the incapacitating effects of the disease itself (2). Third, they are little understood, often demeaned, and unjustifiably feared. Stigma is a significant and widespread feature of mental disorders (3,4). Fourth, like other undervalued populations who are regarded as less important than more powerful and privileged members of society, they are at increased risk of being exploited in research. Abuse of human research subjects has long been associated with members of socially devalued populations. During World War II, a euthanasia program implemented by Nazi physicians killed thousands of patients in German mental hospitals, and set the stage in the concentration camps for the infamous human

subject research on little valued groups like jews, gypsies, and homosexuals (5). Such eugenics programs also were proposed in the United States. A 1942 article in the *American Journal of Psychiatry* suggested killing retarded children (6).

Members of socially devalued groups are singularly vulnerable to exploitation in human subject research. Persons in such groups are more likely to be poor, to be welfare patients, to be in need of health care services, to be high users of institutional facilities like hospitals, group homes, jails, and prisons, to lack social support networks, and, in the United States, to have limited access to health care services. Such populations are apt to be both oppressed and exploited because society discounts them. Their well-being, their rights, and their welfare are disregarded by the society in which they live (7).

Some of the most egregious research studies have targeted specific groups of persons who were considered, in some way, to be inferior. The notorious Public Health Service Tuskegee Syphilis Study, which began in the 1930s and continued for 42 years, used poor and uneducated African-American men as study subjects. The researchers compared the health and longevity of an untreated syphilitic population with a nonsyphilitic but similar population. These men did not even know that they were in a study. They were never told that they had syphilis, nor were they given appropriate medication for their disease when it became available. In 1956 researchers took young residents at the Willowbrook State School who were severely developmentally disabled children, and deliberately infected them with isolated strains of the infectious hepatitis virus. And, another study in 1963, at the Brooklyn Jewish Chronic Disease Hospital dying, impoverished, senile, elderly were — unbeknown to them — deliberately injected with live cancer cells (8).

Genetic Research with Vulnerable Populations

In general, genetic research is minimally physically invasive and does not involve hazardous procedures. Vulnerable populations and the general public may encounter similar kinds of issues in genetic research. Nonetheless, the discrete ethical, social, and legal concerns inherent in genetic research may have greater impact on vulnerable populations. Disorders identified by genetic research are distinctive because they affect not only individuals but also groups of related persons, and groups of unrelated persons. Genetic information about an individual may reveal particular or probabilistic information about his or her living relatives, dead relatives, and future unborn offspring (3). When disclosure of genetic information occurs, individuals, their families, and their ascriptive groups may be affected by a loss of privacy, breaches of confidentiality, familial conflict, and by psychosocial harms.

Studies that make specific comparisons between racial and ethnic groups or involve behavioral genetics may be used to stigmatize and discriminate against members of such populations (9). Genetic information may affect the ways such individuals, their relatives, and their cohorts are viewed by others. On the positive side, knowledge of genetic information relevant to the biological basis of a disorder like Alzheimer's disease appears to have helped dispel myths and reduce stigma associated with the condition. However, some vulnerable and devalued populations, like persons with mental disorders, who already may suffer stigma identified with their disorders and experience unfair discrimination in housing, employment, or insurance, may endure even more intolerance and prejudice. Commentators suggest that strong protections of privacy and confidentiality should be developed in order not to add to the burdens of already disadvantaged groups (3).

Subject Selection Predicament — Exclusion or Inclusion?

A protectionist stance that shields vulnerable groups from research participation may deny them the benefits garnered from scientific research that are available to other disease populations. Paradoxically, such exclusion may reinforce their vulnerability. According to commentators, when diseases that affect women, minorities, and other undervalued populations are not addressed by research, knowledge that could rectify prevailing ineffectual or harmful routine medical care is never produced (10,11). Even though women (and particularly pregnant women), minorities, and children have been systematically excluded from biomedical research, U.S. federal policies have encouraged the research participation of all societal groups. There has been a continuing pressure to increase fair access to, and opportunity to participate in clinical trials (12). Indeed, there has been a paradigm shift in the way enrollment in research trials is viewed since acquired immunodeficiency syndrome (AIDS) was first identified in 1981 (13). Participation in research trials previously had been considered as unavoidably risky and burdensome.

Historically devalued populations, like prisoners and persons with mental disorders, served as subjects and bore the brunt of biomedical research with no forthcoming advantage to themselves or to the populations that they represented, while the more powerful members of society reaped the health benefits gleaned from the research (9,14). Thus vulnerable research subjects were seen as inherently needing protection and an exclusionary model of protection came to be employed. But with the emergence of AIDS this traditional model of protection for selection of human subjects for research trials was transformed. The energetic advocacy of AIDS activists for inclusion in research protocols, and the subsequent clamor of advanced cancer patients to gain admittance to cancer trials encouraged a tight connection between research and treatment. Research came to be seen as a pathway to better medical care for individuals and their particular disease populations. Currently a more expansive and inclusionary model is advocated and greater weight is given to sharing the benefits of research.

Perils of the Therapeutic Misconception: Making the Distinction Between Clinical Treatment and Biomedical Research

Another kind of hazard for vulnerable populations may lurk in a standard that argues for wider inclusion in research trials. Persons who suffer from disabling and

damaging diseases, for which no cure exists, understandably are desirous of relief. However, individuals whose incapacitating diseases may be inadequately treated are likely to be susceptible to the popular therapeutic misconception that research protocols provide beneficial treatment, whether or not there is any prospect of benefit. Indeed, such persons may be especially vulnerable if they are powerless and impoverished members of devalued populations (15).

The distribution of healthcare services in the United States is unequal. Participation in a research protocol for some people may be their only way to gain access to an essential therapy that addresses their particular condition. For example, many persons with mental disorders have no or only inadequate access to any mental health care service due to discriminatory and inequitable health insurance policies that anachronistically differentiate between "physical" health and "mental" health and deny parity. Persons in such circumstances may be willing to bear the risks of research when it appears to offer a prospect for beneficial treatment (12). But biomedical research is not clinical treatment. In clinical treatment the patient's welfare and "personal care" is the physician's first consideration (16). In contrast, in biomedical research the investigator's primary purpose is to develop or contribute to generalizable knowledge, not to benefit any particular research subject (17). The distinction between clinical practice and biomedical research often is misunderstood by research subjects (18). When the demarcation lines are blurred between patient and subject, between physician and investigator, and between treatment and research, patient-subjects are likely to believe, as is the case with treatment, that the research is designed to directly benefit them.

According to Jay Katz, the misconception has grown out of the recent practice of conflating the role of the physician with the role of the research investigator. Physicians who also are investigators compound the potential subject's confusion. It has become common practice for physicians to enroll their patients in their own research studies. Many patients enrolled in such studies believe that they will be the beneficiaries of effective treatments, even though there may be minimal or no likelihood of benefit. Patients misconstrue their physician's invitation to take part in research as a treatment recommendation (19). Patients trust their physicians and permit them to hold enormous power; in exchange, patients expect their physicians only to serve their therapeutic needs (20).

Because the language of many consent forms fosters the belief that a therapeutic benefit will be forthcoming, some commentators propose that the research community take measures to help potential subjects appreciate that they are unlikely to benefit from research participation (21). Other writers express concern that if the therapeutic misconception is not adequately dealt with during the informed consent process, subjects are likely to feel that they have been deceived and public trust of researchers, and the health care system in general, may be eroded. They recommend that a trained, neutral educator provide information about the study and advise subjects when it might not be in their best interests to take part in the study. Additionally these writers suggest that the therapeutic misconception may be lessened if investigators emphasize the substance of the disclosure, explain the scientific methodology, and maintain an ongoing consent process (22).

IMPACT AND INFLUENCE OF INTERNATIONAL CODES, U.S. REGULATORY POLICIES, AND BIOETHICS COMMISSIONS ON RESEARCH WITH VULNERABLE SUBJECTS

International Codes

The central tenet of the Hippocratic Oath — the primary obligation of physicians is to benefit their patients — has prevailed for more than 2400 years in Western clinical medicine. This fundamental ethic recognizes the patient's potential for vulnerability and thus requires that physicians act solely on behalf of their patients' best interests. However, because in biomedical research the scientific method requires that the researcher engender generalizable knowledge, the needs and interests of the vulnerable subject may be compromised. Thus the clinician's obligations to the patient stand in direct conflict with the researcher's duties. These incommensurable goals were rendered potentially compatible by the first set of international research principles, the Nuremberg Code. By making mandatory the voluntary informed consent of all research subjects, the Code strove to shelter the research endeavor (23). The authors of the Nuremberg Code — by employing the moral principal of respect for autonomy — established two kinds of freedom for competent research subjects; the freedom not to be interfered with by others, and the freedom to make their own choices (24).

The Code's dominant principle, "The voluntary consent of the human subject is absolutely necessary. This means that the person involved shall have the legal capacity to consent," was delineated in response to the cruel research conducted by Nazi doctors on prisoners who, because of their circumstance, were not in a position to give their consent, whether or not they were competent (25). Special protections for mentally ill persons, which allowed for proxy consent, were specified in a final memorandum by the chief medical advisor to the Nuremberg judges but were not included in the final document (26). Thus the Code's affirmation of the primacy of consent effectively excludes many persons with disorders affecting their decision-making capacity from participation in drug trials and other sorts of research.

The international codes and regulations that were developed after the Nuremberg Code have attempted to reconcile society's twin responsibilities to adequately protect vulnerable research subjects and to ensure that vulnerable populations receive the benefits of research. The various drafts of the World Medical Association's Declaration of Helsinki, first issued in 1964, endeavor to move along and ease the consent requirements via the mechanism of a legal guardian for persons who lack capacity to consent to research. The Declaration classifies research as "therapeutic" and "nontherapeutic." But the Declaration seems to forbid the enrollment of incapable

subjects in research protocols that do not offer subjects the probability of direct benefit. Research that advances generalized knowledge solely for the benefit of others only may be carried out with volunteers (27).

Another document, the *International Ethical Guidelines for Biomedical Research Involving Human Subjects*, which considers, among other concerns, research involving persons who are incapable of giving adequately informed consent, was issued in 1993 by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO) (28). In the case of incompetent subjects, the CIOMS/WHO guidelines allow for proxy consent: Informed consent may be obtained from a legal guardian or other duly authorized person. The investigator, however, is directed to obtain the consent of each subject to the extent of the subject's capabilities, and also to always respect the prospective subject's refusal to participate in "nonclinical" research.

Evolution of U.S. Policies for Protection of Vulnerable Human Research Subjects

Existing federal regulations provide special protections for research subjects regarded as vulnerable. These regulations apply to research involving children, fetuses, pregnant women, human in vitro fertilization, and prisoners (1). When the federal regulations were first promulgated, special protections for persons, then described as mentally infirm and institutionalized, were proposed but never enacted into law. Although the regulations mention that additional safeguards must be included in the study to protect the rights and welfare of mentally disabled persons, these directions do not (*1*) elucidate criteria, methods, procedures, or limitations to the research design; (*2*) specify the particularities of informed consent processes, or of appointments, powers, and education of surrogate decision makers; (*3*) describe appropriate pathways between the research study and the subject's ongoing clinical care; or (*4*) make provisions for aftercare health arrangements that may be required when the research is concluded (29–31). How did this state of affairs come about?

During and following the Second World War, this country's medical and behavioral research expanded and flourished unchecked, for the most part, by any supervision. Reports of numerous and varied research scandals were exposed (32). Initially the federal government was slow to react despite the efforts of a few members of Congress and some government scientists (33). The government's own Public Health Service's Tuskegee Syphilis Study continued unrestrained until exposed in a 1972 newspaper article. Only then did a flurry of congressional activity occur regarding human research subject protection. In May 1974, regulations protecting human subjects became effective. These regulations raised to statutory status the National Institutes of Health (NIH) Policies for the Protection of Human Subjects, which were first issued in 1966. Forty-three days later Congress created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Because of concerns about the use psychosurgical experimental procedures like pre-frontal lobotomy over which no peer group or federal regulatory body was in a position to regulate, decisionally impaired persons were among the special populations the National Commission was expected to consider (34).

The National Commission in its classic document, the Belmont Report, advanced the principle of respect for persons as autonomous agents, recognized that persons with diminished autonomy are entitled to protection, and applied this basic ethical precept to the process of voluntary and informed consent, as the first principle appropriate to the conduct of research (17). Thus the National Commission echoed the primary tenet of the Nuremberg Code. However, the National Commission allowed a less restrictive approach to involving incapable subjects in research. In its 1978 Report and Recommendations, *Research Involving Those Institutionalized as Mentally Infirm*, the National Commission recognized that indirect harms are likely to occur should research involving persons with mental disorders be forbidden:

> [S]ince some research involving the mentally infirm cannot be undertaken with any other group, and since this research may yield significant knowledge about the causes and treatment of mental disabilities, it is necessary to consider the consequences of prohibiting such research. Some argue that prohibiting such research might harm the class of mentally infirm persons as a whole by depriving them of the benefits they could have received if the research had proceeded (35, p. 58).

This opinion signaled a critical development in the conceptual framework for addressing the moral issues relevant to human subject research in general and to vulnerable human subjects in particular. The National Commission moved away from the position that procedures that hold no prospect of direct benefit to the noncompetent subject are not morally permissible (36). Instead, the National Commission was persuaded by the view that when risks are not unreasonable everyone, with or without decision-making capacity, has an obligation to benefit society, and that within the scope of that obligation research may be acceptable (37). This position allowed for the weighing or balancing of moral interests. The overlapping goals of benefiting the class of incapable persons, while at the same time safeguarding incapable individual subjects from unacceptable harm could be satisfied by balancing the risk of harm with the likelihood of benefit (38).

The National Commission recommended that researchers not involve vulnerable populations in research that was unconnected to their conditions as individuals or as a class of persons. They proposed a ranking of research classifications that instituted to a more precise degree substantive and procedural standards for research protocols involving more than minimal risk to incapable subjects. The National Commission appreciated that not all persons with mental disorders are incapable of giving voluntary and informed consent. They also recommended a process that allowed incapable subjects to assent or object with a simple yes or no when asked about their choice about being enrolled in a study. However, the National Commission was concerned about the vulnerability of this population and advised that when

research protocols involved greater than minimal risk, that institutional review boards (IRBs) be allowed to use their discretionary judgment to appoint a consent auditor to oversee and guarantee the adequacy of the research protocol's consent process. Only in greater than minimal risk research protocols where there was no prospect of direct benefit for the subjects should the presence of consent auditors be compulsory. Because incapable adults lack the legal guardian that most children have, the National Commission suggested it might be necessary for a court-appointed guardian to approve research participation.

The U.S. Code of Federal Regulations reflects many of the National Commission's other recommendations involving populations regarded as vulnerable such as children, pregnant women, and prisoners. Yet the recommendations in its report on *Research Involving Those Institutionalized as Mentally Infirm*, were never transposed into final regulations. In 1978, after the work of National Commission had been completed and the Commission was dissolved, the proposed regulations were published in the Federal Register. The Department of Health, Education and Welfare (now known as the Department of Health and Human Services) regulation writers significantly altered the original recommendations and inserted proposals that were more demanding. They suggested that a consent auditor monitor all research including that which involved no more than minimal risk (39). There was a strong negative response from those who perceived that such regulations would significantly limit valuable research. Human rights advocates had another perspective and voiced concern that clinicians, researchers, and IRBs might not adequately respect the interests of persons with mental disorders.

The next bioethics commission, the President's Commission for the Study of Ethical Problems in Medicine and Biomedical Research, which was established in 1980 by Congress, insisted in 1981 and again in 1983 that the proposed regulations be made official. The Secretary of the Department of Health and Human Services (DHHS) countered that the proposed rules produced a "lack of consensus," and furthermore that the "basic regulations on human subjects research adequately respond to the recommendations made by the National Commission to protect persons institutionalized as mentally disabled" (40). Despite the Secretary's judgment, as of this writing, it is generally agreed that the Code of Federal Regulations does not provide adequate guidance for research with adult persons who have disorders involving some degree of cognitive impairment (30,41–43).

Recent Efforts to Establish Appropriate Safeguards for Vulnerable Subjects

The Advisory Committee on Human Radiation Experiments (ACHRE) was created in 1994 by the President in response to concerns about possible past unethical research conduct by the U.S. government and institutions funded by the government. ACHRE was charged not only to "tell the full story about [human radiation research] to the American public," but also to "examine the present, to determine how the conduct of human radiation research today compares with the past and to assess whether, in light of this inquiry, changes need to be made in the policies of the federal government" for the protection of human subjects in research (44, p. 1).

Among the contemporary research protocols examined by ACHRE were four studies involving diagnostic imaging with cognitively impaired adult subjects. ACHRE's final report noted that in these studies, where subjects' movements were severely restricted, there had been no discussion in the documents or consent forms with the subjects about the implications of these potentially anxiety-provoking conditions. The ACHRE report also mentioned that there was no discussion of the subjects' capacity to consent or evidence that appropriate surrogate decisionmakers had given permission for the subjects' participation. According to the report:

> The question of whether or under what conditions adults with questionable decision-making capacity can be used as subjects of research that offers no prospect of benefit to them is unresolved in both research ethics and regulation. When such research puts potentially incompetent people at greater than minimal risk of harm, it is even more ethically problematic (44, p. 707).

Moreover ACHRE voiced concern in regard to the current system of research oversight. "Without guidance from the federal government, and perhaps regulatory relief, IRBs may not have the flexibility necessary to concentrate their efforts where subjects are in greatest need of protection — on the proposals that pose the greatest risks to subjects" (44, p. 819).

The DHHS Office of the Inspector General (OIG), which conducted a broad inquiry into current IRB practice, reached similar conclusions. According to the 1998 OIG report, IRBs have vulnerabilities that threaten their effectiveness in protecting human subjects (45). The OIG report drew attention to the fact that IRBs rarely conduct ongoing review of active research. Such lapses in oversight may have particularly serious implications when research protocols involve vulnerable subjects who have limited capacity for decisionmaking and thus a limited ability to protect their own interests.

> . . . [This] is a serious national issue because it compromises. . . [the IRBs'] protection of human subjects. It inhibits their capacity to identify and address situations where unacceptable risks emerge, or research results prove to be too favorable to continue, or protocol stray beyond approved limits. It also inhibits their capacity to ensure that the subjects have sufficient understanding of the risks they may incur in the research process (45, p. iii).

In October 1995 the National Bioethics Advisory Commission (NBAC) was established by executive order. The first meeting took place in October 1996. The immediate charge to NBAC was to respond to the ACHRE recommendations to focus on the protection of the rights and welfare of human research subjects. Primary among other tasks, NBAC was to complete the National Commission's unfinished business to consider how ethically acceptable research may be conducted with human subjects who suffer from mental disorders that may affect

their decision-making capacity. The NBAC report and recommendations, *Research Involving Persons with Mental Disorders that May Affect Decisionmaking Capacity*, was published in December 1998 (46). Specifically, the NBAC report "focused its attention on those who may be primarily considered for research because it is their particular mental disorder that is being studied" (46, p. 5). The report identified four types of limitations that may be experienced by persons with mental disorders:

> First, some individuals might have fluctuating capacity, what is often called waxing and waning ability to make decisions, as in schizophrenia, bipolar disorders, depressive disorders, and some dementias. Second, decisionmaking deficits can be predicted in some individuals due to the course of their disease or the nature of the treatment. Although these individuals may be decisionally capable in the early stages of the disease progression, such as in Alzheimer's disease, they have prospective incapacity. Third, most persons with limited capacity are in some way still able to object or assent to research, as in the case of more advanced Alzheimer's disease. Fourth, persons who have permanently lost the ability to make nearly any decision that involves any significant degree of reflection are decisionally incapable, as in the later stages of Alzheimer's disease and profound dementia (46, p. 10).

NBAC commissioners were concerned with the lack of specificity in the federal regulations in regard to safeguards that should be included in research protocols to protect the rights and welfare of adult subjects with illnesses involving some degree of cognitive impairment. Problems surrounding the informed consent process with vulnerable or potentially vulnerable subjects, and challenging moral concerns in regard to research design, distinct to this population, were addressed. Because decisional impairment, some forms of attention deficit, and incapacity may occur more often among "some people with certain mental disorders than in the general population" (46, p. 58), the commission recommended that IRBs require, in studies that pose more than minimal risk, that a qualified professional, independent of the research, assess the person's capacity to consent, even when the potential subject appears to have capacity for decision making. However, the NBAC report allowed for less formal capacity assessment approaches to be used if the researcher can establish good reasons for doing so.

The report also elaborates upon the concept of anticipatory research planning, and elucidates and expands the role of the legally authorized representative. The concept of obtaining a subject's consent to research participation in advance had been examined in 1989 by the American College of Physicians (47). And in 1996 new regulations adopted by the Food and Drug Administration (FDA) and National Institutes of Health (NIH) clarified rules in regard to research involving incapable subjects in emergent and life-threatening situations in emergency settings, and fostered the concept of research anticipatory planning (48). The new rule yields a narrow exception to federal informed consent requirements and permits research to proceed when it is not feasible to get informed consent from a potential subject or the subject's legally authorized representative. The rules allow researchers to obtain a waiver when they cannot reasonably obtain consent in advance or at the time of the subject's enrollment. The regulations also expand the definition of a family member to encompass "any individual related by blood or affinity whose close association with the subject is the equivalent of a family relationship" (49). By supporting the concept of consent in advance and opening up the meaning of "family relationship," these regulations are relevant to and provide guidance for research in general involving cognitively impaired persons (50). It should be noted that NBAC commissioners did "not endorse the idea of authorizing third parties to enroll incapable subjects in research involving greater than minimal risk without the prospect of direct medical benefit" (46, p. 63). For research protocols falling into this category, where subjects have not given specific consent in advance, NBAC proposed that the Secretary of the DHHS create a national panel to review individual protocols that cannot otherwise be approved.

During the 18 months of deliberation that preceded the NBAC report, the commissioners heard testimony from former research subjects, their families, IRB members, and researchers who described their experience with protocols involving subjects with mental disorders. From these public discussions the commissioners learned that certain types of protocols may be expected to escalate subjects' symptoms, relapse, and suffering. The NBAC commissioners reviewed a small selection of such research protocols and consent forms that recently had been conducted in the United States. According to the findings from this brief review, NBAC recommended that IRBs pay special attention and "heightened scrutiny" (46, p. 56) to protocols that incorporate an ethically controversial design. The report identified such studies as those "that are designed to provoke symptoms, to withdraw subjects rapidly from therapies, and to use placebo controls" (46, p. 64). NBAC proposed that in any such studies judged to be clearly critical to the development of scientific knowledge, IRBs, who according to current regulations have the authority to continue observation of approved studies, should exercise that prerogative in studies where subjects may be at risk of relapse.

Ethically Controversial Research Design

Use of placebo in medication trials involving persons with schizophrenia has drawn considerable attention to the ethics of research design. Persons with mental disorders involving decisional impairments, such as schizophrenia and major depression, are likely to suffer painful symptoms that can be life-threatening to themselves or others. Nowadays many of these symptoms can be well controlled by medication. Thus the ethical question arises as to whether it is ever appropriate to withdraw medication from individuals who rely on a specific medical therapy for their continued good health. Even though placebo controlled randomized clinical trials (RCT) generally are judged to be the gold standard for evaluating therapeutic efficacy, the use of placebo controls in RCTs have in certain circumstances come to be viewed as morally problematic. In particular, trials that enroll persons who have cognitive impairments that may affect their decision-making capacity raise serious ethical concerns (51). Such

persons may not adequately understand the concept of a RCT and may be especially confused about the distinction between research (which is to advance and generate generalizable research findings) and clinical treatment (which is to advance the welfare of particular patients). Indeed, as the recent ACHRE studies indicate, persons who were considered to be competent also suffered from the therapeutic misconception and had difficulties distinguishing medical research from clinical treatment (44, p. 761).

There are a variety of opinions regarding placebo use in research protocols. Some supporters of placebo use, including the U.S. Food and Drug Administration (FDA), consider that in the majority of trials with investigational drugs placebos may be necessary in order for the study's conclusions to be reliable (52). Another commentator agrees that the use of placebo controls in RCTs can be highly valuable in certain circumstances, but also points out that such RCTs may be used more often than is necessary (53). In another paper this commentator acknowledges that even when potential subjects are adequately informed, rational individuals are unlikely to agree to participate in RCTs (54). Some writers posit that despite the methodological difficulties, standard therapy should be used as the control in new drug investigations where subjects are at risk for relapse (51). The NBAC propose the employment of three criteria for excluding prospective subjects in placebo arms of studies: (*1*) when "an individualized assessment reveals that certain patients would be at high risk for relapse if a current or prospective therapeutic regimen were discontinued"; (*2*) when "a washout period would not be contemplated for these patients if they were not enrolling in the study"; or (*3*) when "standard therapy has previously proven to be effective" (46, p. 56).

Nevertheless, because unanticipated circumstances can occur, danger may remain for some persons who do not fit the exclusionary criteria and are capable of consenting to participate in drug-free research, challenge studies, or long-term protocols. Subjects who have fluctuating or prospective decision-making impairments whose symptoms are apt to increase are likely to be particularly vulnerable to the exigencies of high-risk protocols. Thus, when symptoms worsen, such individuals may no longer have the capacity for decision making that they were capable of when they initially were enrolled in the study. Such subjects are at risk of becoming vulnerable at precisely the point in a study when they most need to understand, to be aware of, and to make judgments about the use of safeguards — such as their right to withdraw from the study — that have been put place for their protection. Some commentators note that a person's autonomous and voluntary choice to enroll in a research protocol, secured by an informed consent document, by itself may not provide a sufficient safeguard against risks of harm (12,55).

For vulnerable subjects at risk for loss of decision-making capacity, there are other means by which they may provide protections for themselves. According to a number of commentators, anticipatory planning in the form of research advance directives provides not only a mechanism by which potential subjects may choose and appoint surrogate decision makers to act on their behalf should they lose their ability to make decisions for themselves, but also an important method of respecting individual choices. The prior authorization of the surrogate decision maker, and the precise delineation of appropriate subject protections, may make the employment of research advance directives desirable for subjects who may be competent to consent when a study begins, but who may lose their decision-making capacity while participating in the protocol (21,29,46,47,56,57).

RESEARCH ANTICIPATORY PLANNING

Genealogy of Research Advance Directives

In the United States there is some familiarity with anticipatory planning for end-of-life health care. The concept of anticipatory planning was embraced in order to encourage competent individuals to make autonomous choices, in the present, about the medical treatment they would or would not want should they, in the future, lose their capacity for health care decision making. It was hoped that such planning would prompt and enhance a dialogue between doctors and their patients, aid patients to be better informed, and foster the appointment of surrogate decision makers. In truth, this kind of advance planning was seen as a way to forbid heroic measures in the event of a terminal illness (58). However, the employment of such advance directives has been far less successful than numerous bioethicists had at one time anticipated (59,60). Notwithstanding this lack of success, there appears to be a burgeoning interest in psychiatric advance directives. Psychiatric advance directives and advance directives for end-of-life health care are shaped by the same concept — anticipatory planning for a time when the principal may no longer have the capacity to make treatment decisions. But the two kinds of documents differ in substance. The document attentive to end-of-life health care mainly addresses circumstances immediately preceding a singular event — the principal's death. In contrast, the document created for psychiatric treatment endeavors to secure, for a specific population of individuals, a good life (61). Psychiatric advance directives are intended for persons who already have experienced the sort of crisis that they anticipate may recur. Thus they are able to use their past experience to better plan for their needs in similar situations in the future.

The research advance directive is a direct descendant of advance directives for end-of-life care, and has been strongly influenced by the psychiatric advance directive (29). Commentators suggest that substantive and procedural research advance directives, which allow for specific instructions and the appointment of legally authorized representatives (surrogate decision makers), may afford a method to provide protection for some vulnerable research subjects. Research advance directives may be particularly suitable for potential subjects whose decision-making capacity may change during the course of the research. Such persons already may have experienced fluctuating periods of decision-making incapacity or are in

the early stages of progressive diseases like Alzheimer's, dementia, and Huntington's disease. Although research anticipatory planning has been little employed, the concept has been discussed and considered since the 1980s (62,63). In 1996 the FDA and the Office for Protection from Research Risk (OPRR) endorsed the concept of advance informed consent for emergency research (49).

There are critical distinctions between the three types of directives. The end-of-life advance directive was designed primarily to refuse treatment — in the United States the right to refuse treatment appears to be fundamental. In contrast, the purpose of the psychiatric advance directive is both to reject *and to elect* treatment — yet the right to demand treatment is not protected. The research advance directive introduces another kind of anomaly. Such documents, unless carefully regulated, could be used to authorize interventions that may not benefit the research subject. The American College of Physicians in their 1989 position paper, anticipate such a circumstance and propose that these directives may be abrogated if the research would unreasonably endanger the subject's welfare (47). The NBAC commissioners strongly urge that such documents should not be prepared as a blank check for future protocols without regard to risk and benefit:

> Prospective authorization cannot be a "blank check" for research participation.... NBAC limits valid Prospective Authorization to a "particular class of research" and then only if the potential subject, while capable, understood the "risks, potential direct and indirect benefits, and other pertinent conditions" of this particular class of research (46, p. 61).

Healthy elderly persons who participated in a Canadian study were concerned about the use of research advance directives. They believed that procedures and treatment not envisioned at the time the directive was prepared should be prohibited (57). Because the author of the document could inadvertently direct the surrogate decision maker to be an agent in harming them, other commentators also advocate that limits be set on what the principal may request. These commentators recommend that the surrogate decision maker may never — under any circumstances — overrule the principal's objections to participate in the research or any section of the research. Furthermore they suggest that should the subject lose capacity for decision making and be at risk of harm due to some aspect of the study itself, the surrogate is obligated to overturn the subject's instructions and to withdraw the subject from the protocol (64). Yet, other writers argue that because the surrogate's obligation is to implement the principle's wishes, the subject's withdrawal only may occur should the study itself has changed substantially (65).

Unanticipated Circumstances and the Appointment of a Surrogate Decision Maker

Research by its very nature is an activity designed to test a hypothesis and is characterized by uncertainty. Always there is the possibility that unanticipated circumstances will occur. The real prospect of future unknown situations prompts some commentators to argue that research advance directives should only be valid if subjects personally chose, and record, their selection of a surrogate decision maker (21). Some writers believe that if advance directives are to adequately safeguard the principal, the appointment of appropriate and reliable surrogates may prove more important for protection than the principal's ability to give detailed instructions (66,67). Findings from a study on ethical aspects of dementia research reveal that more than half of the study's cognitively impaired patients had the capacity to designate a surrogate decision maker, even though they did not have the capacity to understand a detailed protocol. According to the study's researchers, these patients were capable of identifying surrogates whom they trusted (68).

In 1987 the NIH Clinical Center developed a durable power of attorney mechanism for surrogate decision making in research with persons who were at risk for becoming decisionally impaired. Because surrogate decision makers only assume surrogacy responsibilities during the period when subjects lose their decision making capacities, some writers propose that surrogates participate with prospective subjects in the informed consent process and also co-sign the consent form. Thus surrogate decision makers would be educated about the protocol along with the potential subjects and consequently be privy to the subjects' concerns and wishes in regard to their research participation. According to these writers, surrogate decision-maker participation in the informed consent process may put to rest apprehensions on the part of researchers in regard to sharing the subject's medical information and also bypass the need to pinpoint when services of the surrogate might be required (69).

Nevertheless, there may be obstacles that impede the appointment of an appropriate surrogate decision maker. A designated surrogate, even though well trusted by the principal, may not be capable of doing the job adequately. Some surrogates may not properly pursue a subject's best interests. Who should decide if an appointed surrogate is "appropriate"? One writer suggests that this should not be a judgment made by the researchers alone. This commentator proposes that researchers and IRB committee members educate proxies on the correct ethical standards to be used when making decisions about a decisionally incapable person's research participation (70).

Capacity to Engage in Anticipatory Planning

Making judgments about another person's decision-making capacity may be problematic (70). Society appears to be unable to agree on the degree of impairment it is willing to countenance before it deems that a person lacks adequate decision-making capacity (71). Nonetheless, all potential human research subjects are presumed to be capable of making decisions for themselves, unless there are specific reasons and conditions that lead to the belief that a capacity assessment is required. A careful appraisal of each prospective subject's clinical condition, particular circumstances, and the design of the research protocol are necessary factors in capacity determination. Decision-making capacity customarily is considered to be task specific. A person may lack decision-making capacity in one area but have capacity to make decisions in other areas.

Evaluating a potential subject's capacity typically has consisted of subjective judgments. Now, however, there are beginning to be some tested approaches for assessing capacity to consent to research more objectively (72,73). Many persons with mental disorders whose capacity may fluctuate will have intermittent periods when they have decision-making capacity. It is morally correct, and also usually possible, to approach such potential subjects about participation in relevant research at a time when they are competent.

When consent to participate in research is obtained in conjunction with a research advance directive for persons with limited or fluctuating periods of capacity, or with prospective incapacity, the immediate "task at hand" for such potential research subjects is to understand the concepts involved in anticipatory planning. Commentators consider that capacity to prepare an advance directive is distinct from the capacity to consent to treatment, to research, or even to complete a testamentary will (74). These writers propose that potential subjects should be able to grasp and understand that their consent to participate in a specific research protocol, made in the present, constitutes their agreement to take part in a study that will occur over a specified and perhaps extended period of time. In other words, subjects should be able to discern that some of their choices made in the present may be acted upon in the future. Subjects also should be aware that some of their decisions, when relevant, may involve their agreement to medical procedures. Furthermore potential subjects should clearly appreciate that their appointed surrogate will make decisions for them, should they at a future time while participating in the research protocol become incapable of making decisions for themselves (74). Other commentators maintain that subjects should comprehend that whatever they may have recorded in their research advance directive, that with — or without — decision-making capacity, they may object and withdraw from the study (64).

Research Risk Assessment Ambiguities That Reinforce Need for Anticipatory Planning

Specifying criteria and developing policy that assists people to make accurate judgments about risks of harm in research protocols continues to be difficult. The National Commission's 1978 *Belmont Report* acknowledged this problem but did not resolve it: "It is commonly said that benefits and risks must be 'balanced' and shown to be 'in a favorable ratio.' The metaphorical character of these terms draws attention to the difficulty of making precise judgments" (17, p. 7). Assessments of risk of harm in research protocols must attempt to measure the harm's duration, consequences, potential damage, and how the harm might be considered from a subjective point of view. Even though certain types of risks may be precisely and objectively quantified, many risks of harm only may be qualified because they may be of a more subjective kind (75). It is this tension between objective and subjective considerations that makes it so difficult to fashion procedures and policy that do "justice to the equal importance of all persons, without making unacceptable demands on individuals" (76, p. 5).

The phrase "minimal risk" was advanced by the National Commission as an attempt to establish a baseline measure, and it is the standard used in current U.S. federal regulations. "Minimal risk" is defined as meaning "that the probability and magnitude of harm and discomfort anticipated in research is not greater than those ordinarily encountered in daily life or during the performance of routine or psychological examination" (1, §46.102i). Many commentators recognize that this definition is ambiguous. Writers question whether the "harm and discomfort" is that which may be encountered by healthy people in their everyday lives, or whether ordinary "harm and discomfort" is meant to describe that which may be endured in the daily lives of any population of research subjects who have a particular condition or disease. Some writers propose that if the definition of minimal risk is bound to a subjects's disease or condition, it may be easier to more accurately evaluate the level of risk (77). Others argue that in order to appreciate the meaning of minimal risk in the research context, it must be examined in its specific employment (78). The National Commission in its 1978 report on *Research Involving Those Institutionalized as Mentally Infirm* suggests that the "IRB may determine that prospective subjects who are institutionalized as mentally infirm are likely to react more severely than normal persons to certain routine procedures; in such instances, the procedures present more than minimal risk to the subjects" (35, pp. 8–9).

Making judgments about the risk of harm in research protocols is imprecise. A commentator recommends that when research involves vulnerable or potentially vulnerable subjects, investigators and research institutions must be held to a high standard. Not only must the scientific and ethical justifications be especially sound, but researchers should specify what extra safeguards will be put in place to protect subjects' rights and safeguard their welfare (43). In protocols involving greater than minimal risk, research advance directives may provide a practical way to specify safeguards and guarantee protections for vulnerable subjects. Commentators posit that when a legally authorized representative co-signs the consent form along with the potential subject who is competent, not only is the potential subject's autonomy respected but all parties — the subject, the researchers, and the surrogate decision makers — acknowledge the stipulated protections that must be complied with during (and in some cases after) the study period (62,69).

CONCLUSION

Biomedical research involving human subjects is a distinct kind of undertaking and is essentially different to the routine practice of medicine. In order for research with vulnerable subjects to be ethically permissible, this difference should not only be clarified, but special procedural protections should also be employed. However, the concept of protection for research subjects should no longer mean that vulnerable populations must be excluded from research participation. Rather, protection should

signify that a constellation of safeguards are provided that will guarantee the rights and welfare of all subjects enrolled in studies. The assurance that such protections will be put in place may allow vulnerable populations to volunteer and participate in research protocols designed to study their particular conditions without fear that they will be subjected to research abuse.

BIBLIOGRAPHY

1. U.S. Department of Health and Human Services, Title 45 CFR Part 46, U.S. Government Printing Office, Washington, DC, 1991.
2. H.P. Lefley, *Family Caregiving in Mental Illness*, Sage Publ., Thousand Oaks, CA, 1996.
3. Nuffield Council on Bioethics, *Mental Disorders and Genetics*, Nuffield Council on Bioethics, London, 1998.
4. G. Albrecht, V. Walker, and L. Levy, *Soc. Sci. Med.* **16**, 1319–1327 (1982).
5. R.N. Proctor, in G.J. Annas and M.A. Grodin, eds., *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, New York, 1992, pp. 17–31.
6. F. Kennedy, *Am. J. Psychiatry* **99**(13–14), 141–143 (1942).
7. F. Baylis, J. Downie, and S. Sherwin, in S. Sherwin, coordinator, *The Politics of Women's Health: Exploring Agency and Autonomy*, The Feminists Health Care Ethics Research Network, Temple University Press, Philadelphia, PA, 1998, pp. 234–259.
8. R.R. Faden and T.L. Beauchamp, in collaboration with N.M.P. King, *A History and Theory of Informed Consent*, Oxford University Press, New York, 1986.
9. National Bioethics Advisory Commission, *The Use of Human Biological Materials in Research: Ethical Issues and Policy Guidance*, National Bioethics Advisory Commission, Rockville, MD, 1999.
10. R. Dresser, *Hastings Cent. Rep.* **22**(1), 24–29 (1992).
11. A. Charo, *S. Louis Univ. Law J.* **38**, 135–167 (1993).
12. B. Freedman, in H.Y. Vanderpool, ed., *The Ethics of Research Involving Human Subjects: Facing the 21st Century*, University Publishing Group, Frederick, MD, 1996, pp. 319–338.
13. C. Levine, in H.Y. Vanderpool, ed., *The Ethics of Research Involving Human Subjects: Facing the 21st Century*, University Publishing Group, Frederick, MD, 1996, pp. 105–126.
14. A.M. Hornblum, *Acres of Skin: Human Experiments at Holmesburg Prison*, Routledge, New York, 1998.
15. Statement of Commissioner Alexander M. Capron, in Report and Recommendations of the National Bioethics Advisory Commission, *Research Involving Persons with Mental Disorders that May Affect Decisionmaking Capacity*, Appendix VI, National Bioethics Advisory Commission, Rockville, MD, 1998, pp. 86–88.
16. C. Fried, *Medical Experimentation: Personal Integrity and Social Policy*, American Elsevier, New York, 1974.
17. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, U.S. Government Printing Office, Washington, DC, 1979.
18. P.S. Appelbaum, L.H. Roth, and C.W. Lidz, *Int. J. Law Psychiatry* **5**, 319–329 (1982).
19. J. Katz, *S. Louis Univ. Law J.* **38**, 7–54 (1993).
20. M.A. Rodwin, *Medicine, Money, and Morals: Physicians' Conflicts of Interests*, Oxford University Press, New York, 1993.
21. J. Moreno, A.L. Caplan, P.R. Wolpe, and the Members of the Project on Informed Consent, Human Research Ethics Group, *JAMA, J. Am. Med. Assoc.* **280**(22), 1951–1958 (1998).
22. P.S. Appelbaum et al., *Hastings Cent. Rep.* **70**, April, pp. 20–24 (1987).
23. R.M. Veatch, in H.Y. Vanderpool, ed., *The Ethics of Research Involving Human Subjects: Facing the 21st Century*, University Publishing Group, Frederick, MD, 1996, pp. 45–58.
24. I. Berlin, in H. Hardy and R. Hausheer, eds., *The Proper Study of Mankind: An Anthology of Essays*, Farrar, Strauss & Giroux, New York, 1997, pp. 191–242.
25. *Tribunals of War Criminals before the Nuremberg Military Tribunal under Control Council Law No. 10*, Vols. I and II, U.S. Government Printing Office, Washington, DC, 1949, vol. II, p. 181.
26. M.A. Grodin, in G.J. Annas and M. Grodin, eds., *The Nazi Doctors of the Nuremberg Code*, Oxford University Press, New York, 1992, pp. 129–131.
27. World Medical Association, in A.R. Jonsen, R.M. Veatch, and L. Walters, eds., *Source Book in Bioethics: A Documentary History*, Georgetown University Press, Washington, DC, 1998, pp. 13–15.
28. Council for International Organizations of Medical Sciences (CIOMS) in Collaboration with the World Health Organization (WHO), *International Ethical Guidelines for Biomedical Research Involving Human Subjects*, in H.Y. Vanderpool, ed., *The Ethics of Research Involving Human Subjects: Facing the 21st Century*, University Publishing Group, Frederick, MD, 1996, pp. 501–510.
29. P. Backlar, *Commun. Ment. Health J.* **34**(3), 229–240 (1998).
30. R.J. Bonnie, *Arch. Gen. Psychiatry* **54**, 105–111 (1997).
31. E.G. DeRenzo, *IRB: Rev. Hum. Subj. Res.* **16**(6), 7–11 (1994).
32. H.K. Beecher, *N. Engl. J. Med.* **74**, 1354–1360 (1996).
33. J.C. Fletcher and F.G. Miller, in H.Y. Vanderpool, ed., *The Ethics of Research Involving Human Subjects*, University Publishing Group, Frederick, MD, 1996, pp. 155–184.
34. D. Rothman, *Strangers at the Bedside: A History of How Law and Bioethics Transformed Medical Decision Making*, Basic Books, New York, 1991.
35. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, Report and Recommendations, *Research Involving those Institutionalized as Mentally Infirm*, DHEW Publ. No. (OS)78-0006, U.S. Government Printing Office, Washington, DC, 1978.
36. P. Ramsey, *The Patient as Person*, Yale University Press, New Haven, CT, 1970.
37. R. McCormick, *Perspect. Biol. Med.* **18**, 2–20 (1974).
38. R. Dresser, *Research Involving Persons with Mental Disorders that may Affect Decisionmaking Capacity*, vol. II, National Bioethics Advisory Committee, Rockville, MD, 1998, pp. 5–28.
39. U.S. Department of Health, Education and Welfare, *Fed. Regist.* **43**(223), 53950–53956 (1978).
40. President's Commission for the Study of Ethical Problems in Medicine and Biomedical Research, *Implementing Human Research Regulations: The Adequacy and Uniformity of Federal Rules and Their Implementation*, Stock No. 040-000-00471-8, U.S. Government Printing Office, Washington, DC, 1983, p. 26.
41. J.W. Berg, *J. Law, Med. Ethics* **24**, 18–35 (1996).

42. R.J. Levine, *IRB: Rev. Hum. Subj. Res.* **18**(5), 1–5 (1996).
43. A. Wichman, *J. Health Care Law Policy* **1**, 88–104 (1998).
44. White House Advisory Committee on Human Radiation Experiments, The Final Report, 061-000-00-848-9, U.S. Government Printing Office, Washington, DC, 1995.
45. U.S. Department of Health and Human Services, Office of Inspector General, *Institutional Review Boards: Their Role in Reviewing Approved Research*, OEI-01-97-00190, U.S. Government Printing Office, Washington, DC, 1998.
46. National Bioethics Advisory Commission, *Research Involving Persons with Mental Disorders that may Affect Decision-making Capacity*, National Bioethics Advisory Commission, Rockville, MD, 1998.
47. American College of Physicians, *Ann. Inter. Med.* **111**(10), 843–848 (1989).
48. U.S. Department of Health and Human Services, Food and Drug Administration, *Fed. Regist.* **61**, 51498 (1996).
49. Food and Drug Administration, Rules and Regulations 21 Code of Federal Regulations 50.3(n), FDA, Washington, DC, 1996.
50. J.H.T. Karlawish and G.A. Sachs, *Am. Geriat. Soc.* **45**, 474–481 (1997).
51. K.J. Rothman and K.B. Michels, *N. Engl. J. Med.* **331**, 394–398 (1994).
52. R. Temple, *Accountability Res.* **4**, 276–275 (1996).
53. R.J. Levine, *Ethics and Regulation of Clinical Research*, 2nd ed., Yale University Press, New Haven, CT, 1988.
54. R.J. Levine, *Law, Med. Health Care* **16**(3–4), 174–182 (1998).
55. J.F. Childress, *Hastings Cent. Rep.* **20**(1), 12–17 (1990).
56. P.S. Appelbaum, *Psychiatr. Serv.* **48**(7), 873–874, 882 (1997).
57. E.W. Keyserlingk, K. Glass, S. Kogan, and S. Gauthier, *Perspect. Biol. Med.* **38**(2), 319–362 (1995).
58. D. Callahan, *Hastings Cent. Rep.* **25**(6), 25–31 (1995).
59. R. Dresser, *Hastings Cent. Rep.* **24**(6, Spec. Suppl.), S2–S5 (1994).
60. SUPPORT Principal Investigators, *JAMA, J. Am. Med. Assoc.* **274**(20), 1591–1598 (1995).
61. P. Backlar, *Commun. Ment. Health J.* **33**(4), 261–268 (1997).
62. B.I. Miller, in V.L. Melnick and N.N. Dubler, eds., *Alzheimer's Dementia: Dilemmas in Clinical Research*, Humana Press, Totowa, NJ, 1985, pp. 239–263.
63. J. Fletcher, F.W. Dommel, and D.D. Cowell, *IRB: Rev. Hum. Subj. Res.* **7**, 1–6 (1985).
64. A. Moorhouse and D.N. Weisstub, *Int. J. Law Psychiatry* **19**(2), 107–141 (1996).
65. J.W. Berg, *J. Law, Med. Ethics* **24**, 18–35 (1996).
66. D.M. High, *Gerontologist* **33**, 342–349 (1993).
67. J. Lynn, *Law, Med. Health* **19**, 101–104 (1991).
68. G.A. Sachs et al., *Clin. Res.* **42**, 403–412 (1994).
69. R. Dukoff and T. Sunderland, *Am. J. Psychiatry* **154**, 1070–1075 (1997).
70. R. Dresser, *JAMA, J. Am. Med. Assoc.* **267**, 67–72 (1996).
71. P. Backlar, *Commun. Ment. Health J.* **32**(4), 321–325 (1996).
72. T. Grisso and P.S. Appelbaum, *Law Hum. Behav.* **19**, 149–174 (1995).
73. D.A. Wirshing et al., *Am. J. Psychiatry* **155**(11), 1508–1511 (1998).
74. M. Silberfeld, C. Nash, and P. Singer, *J. Am. Geriatr. Soc.* **41**, 1141–1143 (1993).
75. E. Meslin, *IRB: Rev. Hum. Subj. Res.* **12**(1), 7–10 (1990).
76. T. Nagel, *Equality and Partiality*, Oxford University Press, New York, 1991.
77. E.G. DeRenzo, *J. Law, Med. Ethics* **25**(2&3), 139–149 (1997).
78. B. Freedman, A. Fuks, and C. Weijer, *Hastings Cent. Rep.* **23**(2), 13–19 (1993).

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, ETHICS, STOPPING RULES FOR RANDOMIZED CLINICAL TRIALS

ROBERT J. LEVINE
Yale University School of Medical
New Haven, Connecticut

SUSAN S. ELLENBERG
U.S. Food and Drug Administration
Washington, District of Columbia

OUTLINE

INTRODUCTION

"Stopping rule" is a term that refers to a statistical criterion for termination of a randomized clinical trial (RCT). Typically stopping rules are based on p values, namely the probability that an observed difference between two arms of a RCT (presumed under the null hypothesis to be equal in efficacy) with regard to an outcome measure would have occurred by chance. When the difference in outcomes appears quite large at some interim point in the trial, such that this probability is extremely small, one might conclude that the question of whether the two arms have equivalent effects can be answered definitely in the negative, and the trial can be terminated at that point. Although they are called stopping rules, they are for good reasons usually treated as "stopping guidelines," criteria for giving serious thought to ending a RCT before its planned termination date.

Stopping rules are generally developed only for primary outcome measures, which, in turn, are almost always expressions of efficacy. A typical stopping rule might be represented in the following way for a particular interim analysis: We will consider recommending that the trial be terminated if at the second interim analysis

the difference between arm A and arm B with regard to the primary outcome measure is significant at a level of $p < 0.0005$.

Responsibility for monitoring the data developed in the course of conducting a RCT is often assigned to a data and safety monitoring board (DSMB), especially when the trial is evaluating treatment effects on mortality or major morbidity. DSMB members have access to data regarding safety and, when appropriate, to efficacy; such data are kept highly confidential and neither the sponsors nor the investigators are permitted access to them. The DSMB generally conducts periodic assessments of the data regarding the outcome measures (*interim analyses*); one of its major responsibilities is to decide, based on the previously specified stopping rule, whether to recommend to the steering committee (or other group having the responsibility and authority to take such actions) that the RCT be terminated or modified.

The ethical functions served by stopping rules are related to the fundamental ethical principle of beneficence, a formal statement of which was provided by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research: (*1*) Do no harm and (*2*) maximize possible benefits and minimize possible harms (1). Skillful design and implementation of stopping rules maximizes benefits by increasing the efficiency of clinical trials. When efficacy is demonstrated before the planned termination of a RCT, the therapy that was found effective may be provided to its intended beneficiaries that much earlier. Efficiency in a different sense is also served by conserving resources that would have been wasted by continuing expensive and unnecessary clinical trials. Effective use of stopping rules also serves to minimize harms by reducing the time of exposure of research subjects in the control arm to the ineffective or less effective therapy.

STATISTICAL APPROACHES TO DATA MONITORING

Sequential Designs

When data are reviewed on multiple occasions over the course of a study, the chances of observing a statistically significant result ($p < 0.05$) on at least one occasion can be substantially greater than 5 percent. This results from having multiple opportunities to observe the event of interest, thereby increasing the overall chance of ever observing the event. To take a more everyday example, the chance of drawing the ace of spades from a complete deck of cards is 1 in 52, but the chance of drawing the ace of spades at least once if one draws one card each day for 10 days is substantially greater than 1 in 52.

McPherson, writing in the *New England Journal of Medicine* nearly 30 years ago, showed that, in a case in which there was no difference in outcomes between the treatment and control groups, the probability of ever observing a difference significant at the 0.05 level was actually about 14 percent if the data were reviewed a total of 5 times, and about 19 percent if there were 10 interim reviews (2). Thus, if the risk of a false positive finding (or type I error) is to be kept under 5 percent for the experiment as a whole, statistical designs that preserve the 5 percent level of error (or 1 percent, or whatever other level has been predetermined as appropriate) must be employed.

Study designs that provide for interim analyses of accumulating data while maintaining the overall type I error at the desired level are called *sequential designs*. The simplest approach to preserving type I error in a sequential design is to determine the number of times one wishes to examine the accumulating data during the course of the study, and to then determine the threshold significance level that, if applied at each interim analysis, would lead to a type I error of 5 percent for the experiment as a whole. This problem was studied by Pocock, who showed how to calculate these values (3). For example, he showed that if there were to be a total of five analyses, a p-value of 0.0158 would have to be used at each analysis in order to ensure that the false positive rate for the entire study did not exceed 5 percent. A problem with this approach, however, was the difficulty in interpreting the final analysis when the significance level fell between the threshold of 0.0158 and the nominal value of 0.05. However mathematically correct the threshold value was, it was disconcerting to declare a study with a final p-value of 0.03 as a nondefinitive result because of the number of times the data had been reviewed prior to the final analysis.

A few years later O'Brien and Fleming at the Mayo Clinic developed an alternative design that provided for varying threshold levels as the study proceeded (4). In this design the first interim analysis is performed using an exceedingly small threshold value. At each successive interim analysis the threshold value is increased by a modest increment. This design allows the final threshold value to be close to the nominal value of 0.05. Using the same example as above, with a total of 5 interim analyses, O'Brien and Fleming showed that the following sequence of threshold values would produce an overall type I error rate of 5 percent: 0.0000005, 0.0013, 0.0085, 0.0228, and 0.0417. This design was appealing because it decreased the chance of stopping the study very early, when most investigators would want to be especially conservative, and also decreased the chance that the final "correct" analysis (based on a p-value of 0.0417) would lead to a different conclusion than a final analysis based on the nominal 0.05 level of error.

More recently a number of variations on the O'Brien-Fleming design have appeared. Perhaps the most important is that developed by Lan and DeMets (5), who showed that type I error could be preserved even if the number of interim analyses to be performed were not specified at the beginning of the study. These statisticians recognized that ensuring patient safety sometimes required that additional interim analyses beyond those that were originally planned, and it was important to be able to maintain the validity of study conclusions about efficacy in those circumstances. They introduced the concept of an *alpha-spending function* that specifies how rapidly the type I (alpha) error is to be used up (or "spent") but does not require the number of analyses to be determined in advance. Thus the incorporation of spending functions into the sequential analysis of clinical trials has provided added flexibility without jeopardizing the control of the type I

error level. O'Brien-Fleming designs that incorporate Lan-DeMets alpha-spending functions are probably the most commonly used type of sequential statistical designs in today's RCTs. An excellent discussion of the use of alpha-spending functions can be found in DeMets and Lan (6).

Monitoring for Lack of Effect

In addition to monitoring for definitive establishment of treatment effect, it is sometimes important to monitor for lack of effect, and to have the opportunity to terminate a trial early when the accumulating data are highly inconsistent with the existence of a clinically important treatment effect. Because such monitoring may increase the number of times that a treatment could be declared ineffective, but not the number of times it could be declared effective, the concern with these designs is not increase of the false positive rate but rather an increase in the rate of false negatives (type II errors). Monitoring procedures of this type (often referred to as *stochastically curtailed testing*) have been shown to reduce substantially the number of patients treated on clinical trials of ineffective treatments, while having minimal impact on type II error (7). These procedures are based on a re-calculation of the statistical power of the trial (the complement of the type II error, i.e., the probability that a true benefit will be statistically detected), given the data that have already been observed. Because the analyses take into account the already-observed data, they are often called *conditional power analyses*. Stochastically curtailed testing can be implemented in conjunction with a standard sequential design as described above.

Monitoring for Safety

Monitoring for safety is usually very different from efficacy monitoring. With the latter, the variable identified as the primary study endpoint — such as mortality, disease recurrence, occurrence of another undesirable clinical event — is prespecified, as is the analytical method that will be used to assess it. While it is often possible to prespecify certain safety outcomes of particular concern, interim safety monitoring must cover all types of adverse events, whether or not they are anticipated. Further, in many cases safety concerns arising in RCTs, even without the definitive probabilistic framework that would be demanded for an efficacy endpoint, will lead to modification or even early termination of the trial. A strong suggestion of harm, even without definitive proof, would likely be sufficient to warrant such actions when the intended benefit of the treatment is relatively modest.

ISSUES IN THE IMPLEMENTATION OF STOPPING RULES

Need for Judgment

Sequential designs have proved very useful in the conduct of RCTs fostering careful monitoring of interim data to ensure that patients are being treated appropriately and safely without sacrificing the validity of the statistical conclusions that will ultimately be drawn. This point notwithstanding, it is essential to recognize the impossibility of developing a design that will account for all contingencies that might occur in the trial. For example, suppose that the interim data at, say, the third look, show a positive effect of treatment that exceeds the threshold for early termination, but at the same time, unexpected safety concerns have emerged. A DSMB in that situation probably should not recommend stopping at that point on the basis that efficacy had been demonstrated; rather, it probably should recommend collection of additional data to help clarify the risk–benefit considerations. In some cases information about findings from other trials or related studies might affect the DSMB's perspective on the data from the trial being monitored. DSMBs must be prepared to consider all the available information pertaining to the safety and efficacy of the treatment being studied prior to making its recommendations, and cannot rely entirely on the statistical stopping thresholds as the basis for decision making. As stated by DeMets et al.:

> Although sophisticated statistical methods have been developed to assess the quantitative strength of trial results, it is important to note that statistical methods alone are not adequate to guide early termination decisions. The collective experience and judgment of the [data monitoring committee] is necessary. Making decisions about early termination requires consideration of many additional factors such as results on secondary outcomes, safety data, degree of compliance to the protocol, possible sources of bias in outcome evaluation, completeness and currency of data, and internal and external consistency of the data as well as emerging data from other trials (8).

Concerns About Designs That Permit Early Termination of Clinical Trials

While statisticians have paid much attention to the development of methods for interim analysis of clinical trials that preserve type I error at the desired level, it must be remembered that the only occasion for concern about type I error is when one might terminate a study and draw a final conclusion based on its interim results. Early termination options are most often important for trials of therapies that may, as compared with other available treatments, improve the likelihood of survival or reduce the probability or magnitude of disability. In many clinical trials, however, there may be good reasons to continue the trial to the planned conclusion even if an interim analysis demonstrates definitive benefit with respect to the primary efficacy variable. When the new treatment is aimed at relief of symptoms, or any endpoint that is neither serious nor irreversible, the need for a fuller safety database becomes more compelling; we must have adequate data on the possible risks before being able to conclude that the observed benefit outweighs these risks. (Adverse outcomes are caused by many factors, not just investigational treatments, so safety data collected in a randomized, controlled, blinded trial are much more reliable indicators of the potential adverse consequences of treatment than safety data collected using less rigorous methods such as uncontrolled case series or retrospective data base analyses.) Thus, in trials in which the outcome measures are neither death nor disability, there may be no need for sequential designs;

monitoring for safety, however, would of course remain important.

Some investigators have argued that even in circumstances when mortality or serious morbidity is the trial endpoint, we ought not to be terminating trials early except in the rarest of situations. Their rationale is that the results of a trial with truncated enrollment may be less convincing to the medical community than results based on a larger patient population, and that a positive trial that does not lead physicians to change their practices has not accomplished any worthwhile purpose and might as well not have been performed. Concerns about the adequacy of estimates of treatment effect from trials stopped early and the consequent concern about the ability to derive reliable cost-benefit considerations for the treatment studied, may threaten the acceptance of results from such trials (9).

The counterargument is that it is unethical to continue enrolling patients in RCT when one of the treatments being compared has already been demonstrated to be superior. To do so entails a deliberate withholding of the therapy known to be superior from those subjects who are assigned to the other (inferior therapy) arm of the trial. It is hard to imagine that fully informed subjects would agree to cooperate with such a plan. Deliberate withholding of known effective therapy without the consent of the subjects is unethical. It at least appears to violate the standard set forth in Article 1.5 of the Declaration of Helsinki: "Concern for the interests of the subject must always prevail over the interests of science and society" (10).

The fundamental controversy here is perhaps more accurately focused on the level of evidence that should be required for a trial to be considered positive, rather than whether interim analysis, possibly leading to early termination, should be performed. If the threshold for statistical significance were to be set at $p < 0.01$ (or lower), the required sample size would be substantially larger (30 to 60 percent larger, for power in the typically acceptable range) than for a trial using the more conventional $p < 0.05$ threshold, and the O'Brien-Fleming stopping boundaries that would result would require much more extreme evidence of benefit for early stopping — possibly even enough to be persuasive to a highly skeptical medical community.

ACKNOWLEDGMENT

R.J. Levine's work was funded, in part, by grant number PO1 MH/DA 56 826-01A1 from the National Institute of Mental Health and the National Institute on Drug Abuse.

BIBLIOGRAPHY

1. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, DHEW Publ. No. (OS) 78–0012, U.S. Government Printing Office, Washington, DC, 1978.
2. K. McPherson, *N. Engl. J. Med.* **290**, 501–502 (1974).
3. S.J. Pocock, *Biometrika* **64**, 191–199 (1977).
4. P.C. O'Brien and T.R. Fleming, *Biometrics* **35**, 549–556 (1979).
5. K.K.G. Lan and D.L. DeMets, *Biometrika* **70**, 659–663 (1983).
6. D.L. DeMets and K.K.G. Lan, *Stat. Med.* **13**, 1341–1352 (1994).
7. K.K.G. Lan, R. Simon, and M. Halperin, *Commun. Stat. Sequent Anal.* **1**, 207–219 (1982).
8. D.L. DeMets et al., *Controlled Clin. Trials* **16**, 408–421 (1995).
9. R.L. Souhami, *Stat. Med.* **13**, 1293–1295 (1994).
10. World Medical Association (WMA), *Declaration of Helsinki: Recommendations Guiding Physicians in Biomedical Research Involving Human Subjects*, WMA, as amended by the World Medical Assembly in South Africa, 1996.

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, LAW, COMMON LAW OF HUMAN EXPERIMENTATION

WENDY K. MARINER
Boston University
Boston, Massachusetts

OUTLINE

INTRODUCTION

The Common Law

There is little common law on research with human subjects in the United States. Common law is distinct from statutes enacted by legislatures and from regulations issued by government agencies. Unlike statutes and regulations, common law rules have no single official text. They consist of principles defining general rights and obligations that are summarized from the reasoning of court decisions in lawsuits. This gives the common law flexibility to adapt itself to new circumstances without the need for rewriting a law in its entirety. It also means that the principles and their application are subject to interpretation, and often dispute, which permits a degree of uncertainty that can be discomfiting to those who seek absolute predictability in the law.

Common law in the United States developed from English common law applied in the colonies before Independence (1). It remains within the jurisdiction of the states, each of which is free to develop its own principles and rules, thereby eroding some of the commonality of the common law. Nonetheless, at the level of generality with which we are concerned here, there is enough consistency in principle and doctrine to permit useful generalizations. For detailed application of common law rules in any one state, it is essential to consult that state's specific laws.

For decades, the common law remained in the background of research policy. The fact that common law imposed duties of care on researchers and granted rights to human subjects of research did not prevent scandalous abuses of research subjects. Perhaps the most comprehensive statement of researcher's duties in what may be considered international common law is the Nuremberg Code, the 10 principles set forth in the 1947 judgment against Nazi physicians convicted of crimes against humanity (murder and torture) under the guise of medical experimentation (2). The judgment, by American judges in a military tribunal established by the United States Military Government for Germany after World War II, is also precedent for common law duties of researchers in the United States. Yet few American courts have even referred to the Nuremberg Code, much less applied it, in cases involving research subjects (3).

Revelations of unethical research and experimentation with human beings without their consent did not give rise to vigorous enforcement of common law principles, perhaps because common law rights must be enforced by lawsuits brought by those whose rights are violated. Instead, in the 1960s and the 1970s, such revelations inspired new federal guidelines and regulations (4–9). Federal laws requiring those who receive federal funding for research with human subjects to comply with specific regulations intended to protect subjects from such abuses have been harmonized into what is known as the Common Rule (10). The Common Rule applies to the Departments of Agriculture, Commerce, Defense, Education, Energy, Health and Human Services (including the Office of the Secretary, the Food and Drug Administration, the National Institutes of Health), Housing and Urban Development, Justice, Transportation, and Veterans Affairs, and the Consumer Product Safety Commission, Environmental Protection Agency, International Development Cooperation Agency (including the Agency for International Development), National Aeronautic and Space Administration, and National Science Foundation. Although the Common Rule incorporates basic elements of the common law into its regulatory provisions, the Common Rule has had more influence on the conduct of research than the common law (11).

Nevertheless, the common law remains the legal backstop to fill gaps left by regulations. More important, it provides research subjects with a legal remedy for injuries (which federal regulations do not). Thus the common law can be seen as describing the legal boundaries for lawful research and responding to the claims of injured research subjects, within which more specific statutes and regulations carve out particular additional duties.

Because common law is derived from judicial opinions deciding cases and controversies among parties to real disputes, it focuses on legal, not moral, rights and duties, and specifically on remedies for legal wrongs committed by one party against another or injuries to one party for which another party is legally responsible (12). Thus the common law does not enforce moral obligations or ethical principles, although many legal principles are based on moral theory and, in the case of research, codes of research ethics. While one's failure to act virtuously or adhere to moral principles may subject a person to moral opprobrium, the absence of virtue is not sufficient to warrant legal recourse.

Classification of Common Law Claims

Common law principles affecting research with human subjects are derived from more general principles applicable to all people and organizations. They are found primarily within the law of tort, or civil wrongs. Tort law assigns responsibility for certain duties, prescribes basic rules of conduct intended to prevent avoidable harms, and imposes penalties for unlawful conduct (13). The goals are typically described as deterring harm, compensating injury, and, sometimes, retribution for wrongs (14).

Tort law is enforced when a person brings a legal claim — called a cause of action — against another party (defendant) who has caused harm to the claimant (plaintiff) as a result of violating a legal obligation. The causes of action most relevant to research are intentional torts — including battery, fraud, intentional infliction of emotional distress, and invasion of privacy — and unintentional torts — specifically negligence, negligent infliction of emotional distress, breach of confidentiality, and products liability. Negligence is often thought to be the most likely basis for liability. However, most reported cases involve claims of battery (unauthorized touching), fraud, and misrepresentation — using people as research subjects without their knowledge or consent (15,16). Often several causes of action are brought in the same lawsuit, and these may be supplemented by actions for violations of statutes or constitutional rights. This article is limited to common law causes of action.

One can characterize the possible common law issues by (*1*) the type of legal claim (or cause of action) that might be brought by a research subject, (*2*) the type of defendant or entity claimed to be responsible for research harms, (*3*) the type of research subject, or (*4*) the nature of the research product or intervention, as summarized in Table 1. Elements in each of these categories can be combined in multiple variations to produce a staggering array of potential common law cases. There are more ways to classify the types of legal responsibility for research harms, however, than there are cases to use as examples. The full range of possibilities has not materialized in lawsuits, and there are relatively few published court decisions that address the common law duties and rights of researchers and subjects. Some consider this reassuring evidence of either the lack of harm caused by research or the negligible prospect of liability on the part of researchers. Others view the small number of cases as the tip of an iceberg of potential future litigation.

Table 1. Variables in Common Law Claims

1. Cause of action

- Intentional torts
  - Battery
  - Fraud, misrepresentation
  - Invasion of privacy
- Negligence
  - Research design
  - Research conduct
  - Informed consent
  - Failure to notify of later-discovered risks
  - Breach of confidentiality
  - Invasion of privacy
- Product liability
  - Design defects
  - Manufacturing errors
  - Failure to warn of risks
- Defenses
  - Charitable immunity
  - Sovereign immunity
  - Statute of limitations
  - Waiver, release
  - Preemption

2. Responsible party

- Individual investigator
- Research institution/employer
  - Institutional review board
- Research funder/sponsor
  - Private product sponsor
- Government

3. Research subjects

- Competent adults
  - Women
  - Fertile women
  - Pregnant women
- Vulnerable populations
  - Children
  - Fetuses, embryos
  - Incompetent adults
  - People with terminal illness
  - People in medical emergency
- Exploitable populations
  - Prisoners
  - Illiterate populations
  - Impoverished populations
  - Minority populations
  - Subordinates of investigators
  - Elderly populations

4. Product or intervention

- Pharmaceutical
- Medical device
- Biological product
- Diagnostic technique
- Surgical procedure
- Medical procedure
- Preventive intervention
- Health services
- Genetic therapy
- Tissue collection
- Psychiatric intervention
- Data collection
  - Epidemiological
  - Genetic
  - Sociodemographic

Injuries to Research Subjects

The number of research subjects to whom common law duties are owed is difficult to estimate. The U.S. General Accounting Office reports that the U.S. Department of Health and Human Services (DHHS) alone funds about 16,000 studies involving human subjects each year (for $5 billion) (17). Other federal and state agencies and private organizations that fund or conduct research significantly increase this number.

There is little empirical data on research injuries and even less on the proportion of research subjects who pursue legal remedies for their injuries. In 1976 the federal Department of Health, Education, and Welfare (HEW) Secretary's Task Force on the Compensation of Injured Research Subjects conducted a survey of researchers to estimate the number of injuries to subjects who had participated in research studies funded by the National Institutes of Health (NIH) and Alcohol, Drug Abuse, Mental Health Administration (ADAMHA) (18). That survey, which relied on telephone interviews with 331 investigators, reported 4957 injuries among 133,000 human subjects over the three preceding years (19). Research subjects who were injured represented just under 4 percent of all subjects who participated in such studies. The investigators characterized 3926 injuries as trivial, 974 as temporarily disabling, 14 as permanently disabling, and 43 deaths (fewer than 1 percent of subjects). The majority of injuries befell subjects who participated in so-called therapeutic research, which was defined as "an experimental program expected to benefit the research subject directly." Of the 39,216 subjects in these therapeutic research studies, 4246 (or 10.8 percent) were injured. In nontherapeutic studies, 711 (0.8 percent) out of 93,399 research subjects were injured. The survey was limited by reliance on reports by the principal investigators of the studies themselves and may understate actual injuries. The Harvard Medical Practice Study found that about 3.7 percent of hospitalized patients in 1984 were injured as a result of ordinary medical care (not research), which might be expected to result in fewer injuries than research (20). Without a similar study of research injuries, it is difficult to estimate the prevalence of research-related injuries that might justify claims of liability for personal injury suffered by research subjects.

There is no empirical study of the proportion of injured research subjects who seek legal redress for their injuries. As far as can be determined from published judicial decisions and anecdotal reports, nothing remotely approaching 4 percent of injured research subjects file common law claims against a third party. An Institute of Medicine committee reported that "the NIH Office of the General Counsel is only aware of three legal actions for research injuries where NIH was involved in the . . . twenty years" before 1994 (17). The Harvard Medical Practice Study found that only a tiny fraction (less than 2 percent) of patients injured as a result of negligent medical care (not research) actually file a claim for malpractice (21). The same is probably true for negligent research injuries.

There are several possible explanations for the small number of cases. Not all injuries are caused by unlawful

conduct that gives rise to a legal claim. The nature of research often precludes clear findings of negligence or other unlawful conduct. Research is done because, by definition, the safety and efficacy of what is being studied is not known (22). (This is distinct from the standard of care used in carrying out the research.) If the Secretary's Task Force's findings hold true today, the majority of injuries involve subjects who are ill so that it is often difficult to sort out whether an injury was caused by underlying illness or by the research study. Subjects may not distinguish between problems with the experimental product or intervention, and problems with the way the research is conducted. Also subjects should be aware that research necessarily entails risks and may assume that they cannot sue. Some subjects may not be aware that they were injured as a result of research. History suggests that the most common legal wrong in research has been the failure to tell research subjects that they are involved in research. However, if subjects are not injured, there are few legal remedies available. Many subjects may not be inclined to sue in any circumstances, especially in light of the time and expense required. Lawyers who represent clients on a contingency basis (receiving a percentage of the money award only if the action succeeds) are unlikely to take cases that have poor or uncertain prospects of success or low potential awards. On the other hand, some cases may be settled voluntarily before any legal claim is brought or before the case is tried or otherwise finally decided in court. Some research institutions voluntarily provide remedial medical care, insurance, or other assistance to subjects who are injured as a result of research.

Perhaps all that can be said is that no one knows how many people have been injured as a result of participation in research, but injuries are likely to befall a relatively small percentage of all research subjects, and in all probability, only a small proportion of these research-related injuries will give rise to legal claims (23).

COMMON LAW CAUSES OF ACTION AGAINST INVESTIGATORS AND RESEARCH ORGANIZATIONS

All investigators and organizations that design, supervise, carry out, or report research are accountable to research subjects for violation of their common law obligations. These duties include (*1*) determining that the proposed research is properly designed and can be conducted without posing unnecessary, avoidable, or unreasonable risks of harm to research subjects, (*2*) ensuring that each research subject who participates in the study has voluntarily agreed to participate with full knowledge of the potential risks of participation, (*3*) ensuring that all investigators are qualified and competent to carry out the research, and (*4*) ensuring that the research is in fact implemented properly and that the subjects' safety and welfare is protected (7).

These duties are derived both from the growing body of ethical literature, including "codes" or declarations of research ethics that help to create the standard of practice among researchers, as well as more deeply entrenched common law principles of self-determination and reasonable care (7,24). Many of these common law duties have been supplemented by the federal regulations that govern federally funded research (10,25). For example, the common law does not require an institutional review board (IRB) to review or approve research with human subjects, but federal regulations do. Arguably, IRB review and the threat of losing federal research funding may have improved compliance with common law duties as well as federal rules, or become the custom or standard of conduct for research institutions. In some states, violation of federal regulations can be considered negligence per se which eases a litigant's path to success. On the other hand, federal regulations do not grant research subjects a personal remedy against those who violate federal regulations. Although federal agencies may impose sanctions against violators, by withholding future funding to the researcher or organization, individual subjects are not compensated by such penalties. Thus the common law serves as the research subject's only source of rights to personal redress for harm.

The most common causes of action are described below.

Using Human Subjects Without Their Knowledge or Consent

The majority of litigated cases involve using human beings as research subjects without their knowledge or consent. The most horrific example — which generated the Nuremberg Code and sowed the seeds of codes of ethics in Western countries — was the Nazi physicians' medical experiments on concentration camp prisoners during World War II (26,27). Notorious American examples include the U.S. Public Health Service's Tuskegee study of syphilis in poor, black men, (5,28) federally sponsored cold war era studies of radiation exposure and poisoning, (29–31) and the Willowbrook study of hepatitis B in retarded children (4,32). However, less dramatic studies have also been conducted without consent (33).

More than a century ago, the U.S. Supreme Court said that "no right is more sacred, or is more carefully guarded, by the common law, than the right of every individual to the possession and control of his own person, free from all restraint or interference of others unless by clear and unquestionable authority of law" (34). The right to decide whether to permit anyone to violate one's own bodily integrity is a fundamental principle of common law (35). Informed consent is a "concept, fundamental in American jurisprudence, that the individual may control what shall be done with his own body" (36). Not even a physician who believes that medical care is necessary for a patient is permitted to act without the informed consent of the patient (37,38). Although the details of the legal cause of action have evolved over the past decades, the law has never permitted the involuntary treatment of a competent patient by a physician. Likewise no one is permitted to conduct research on any human being without that person's competent, voluntary, informed, and understanding consent (39).

Battery. At common law the failure to obtain consent is considered a battery — an unauthorized offensive touching (40). Researchers who use human beings as research subjects without telling them that they are

conducting a study or giving the subjects an opportunity to refuse to participate commit a battery. Battery is a straightforward cause of action that requires a plaintiff to prove that some form of offensive, intentional contact took place without the plaintiff's consent or against his will (41). The touching need not cause any physical harm to be unlawful. The wrong lies in the offense to personal dignity caused by invading the inviolability of the person without permission. Successful plaintiffs can recover money damages for battery. These can range from nominal amounts (e.g., $1) for minor intrusions that cause no harm to substantial awards for highly offensive acts or those that cause serious physical injury.

The offense of battery may be more common in the research context than it is in the realm of medical care. This is because physicians who provide medical care to patients are more likely to obtain consent at least to the treatment given, so that disputes tend to focus on whether the patient was advised of the risks of treatment, as discussed below. Some research studies, on the other hand, have been concealed entirely from subjects.

A notorious example occurred at the Jewish Chronic Disease Hospital in Brooklyn, New York, in 1963 (42). A researcher wished to study whether the human immune system could be used to prevent cancer. It was known that healthy people had strong defenses against injections of cancer cells from other people, while cancer patients did not, but it was not known whether the cancer patients' response was due to their cancer or their overall debilitation. Drs. Southam and Levin of the Sloan–Kettering Institute for Cancer Research injected cancer cells into 22 Jewish Chronic Disease Hospital patients who had diseases other than cancer to see whether their immune systems would reject the cancer cells. The subjects were told only that they would receive a test of their immunity; they deliberately were not told that cancer cells were injected under their skin because researchers thought this might cause them to refuse to be in the study. Three young staff physicians refused to participate and informed a member of the hospital board of directors who in turn notified the New York State Department of Education and the Supreme Court of Brooklyn. The state's highest court allowed the hospital director to inspect the subjects' medical records as part of an investigation into illegal and improper experimentation on patients (43). The licenses of the investigators were suspended for one year, but the suspension was later replaced by one year's probation (4). The patients themselves were not parties to the lawsuit, but the hospital instituted a policy of requiring informed consent in future experiments.

In *Mink v. University of Chicago*, a federal court described why the failure to tell subjects that they are being used as research subjects is a battery (44). The University of Chicago and Eli Lilly & Company conducted an experiment from 1950 to 1952 to see whether diethylstilbestrol (DES) would prevent miscarriages. They gave DES to women who came for prenatal care without telling them of the experiment or that the pills were DES. The court noted that the cause of action was battery, not lack of informed consent, because the researchers performed an action to which the women did not consent at all. The court compared their actions with performing unauthorized surgery. The administration of a drug without a person's knowledge fits within the meaning of offensive contact for purposes of battery. The fact that the women had consented to prenatal care did not mean that they had agreed to take experimental pills.

Instances of undisclosed research can give rise to other common law claims, including invasion of privacy and fraud and misrepresentation. Invasion of privacy — specifically, an unreasonable intrusion upon a person's seclusion or unreasonable publicity given to a person's private life or affairs — can occur when researchers seek out or disclose private information without the person's permission (45). Also possible is a cause of action for fraudulent misrepresentation, which includes a misrepresentation of fact or intention on the part of a researcher, the subject's reliance on the researcher's false statements, and resulting damages (46). Researchers may also be liable for failing to disclose a fact that would induce a person not to enter a study because researchers have a duty to provide all relevant information to prospective subjects. For example, in *Craft v. Vanderbilt University*, a group of women brought a class action claiming battery, fraudulent concealment, negligent misrepresentation, infliction of emotional distress, negligence, and invasion of privacy, as well as federal statutory claims, when they discovered, in 1993, that they had unknowingly been part of research studies decades earlier (47). In a study conducted from 1945 to 1947 by the university and the state of Tennessee, researchers sought to determine iron absorption in the uterus. Researchers gave 829 pregnant women patients in Vanderbilt's prenatal clinic a beverage containing the radioactive isotope Iron 59 but told them only that it was vitamins, "a cocktail" or "a sweet." The women were never told what the drink contained or that they were research subjects, even during a follow-up study in the 1960s to determine the long-term health effects of radiation exposure during pregnancy. That study found three deaths from cancer among children of the exposed mothers (and none among controls) that suggested a cause and effect relationship with the radioactive drink (48). The suit was settled for $10.3 million in 1998 (49).

Research conducted after World War II on the effects of radiation exposure has given rise to other lawsuits by surviving subjects or their families in the 1990s. Like most lawsuits, these assert multiple causes of action — slightly different legal wrongs — for the same injury. Where the subjects were not told that they were part of a research study or that they would be exposed to radiation, a claim of battery can be made. In one such experiment, the Massachusetts Institute of Technology fed Quaker Oats cereal with radioactive isotopes to children institutionalized at the Fernald School. A class action on their behalf was partially settled for $1.85 million in 1999 (50).

Where the research is misrepresented as medical care, a claim of fraud or misrepresentation is possible. For example, the plaintiff in *Stadt v. University of Rochester*, brought a claim of fraud, instead of battery, against the university hospital for injecting his 41-year-old mother

with plutonium without her consent (51). Janet Stadt was used as a subject in an Army research study of radiation effects in 1946 but allegedly was told that she was being treated for scleroderma. In 1972 she underwent additional tests, again without being told they were part of the study, and died of cancer in 1975. The failure to disclose the research was sufficient to allow a lawsuit based on fraud, although, in this case, the court converted the state common law claims into a federal claim against the federal government for violating the subject's right to bodily integrity protected by the Fourteenth Amendment to the U.S. Constitution.

A more complicated example is the Human Radiation Experiments conducted at Cincinnati General Hospital between 1960 and 1972 (52). There, at least 87 patients with inoperable cancer were given radiation in doses ranging from 25 to 300 rads — the level expected to be experienced by military personnel exposed to nuclear attack — to study the effects of radiation on human beings. The study was part of the Defense Department's cold war efforts to prepare for possible nuclear war. The patients were told that the radiation was treatment for their cancer, although it shortened their lives and caused nausea, vomiting, burns, and other suffering. The plaintiffs claimed the experiments were intentionally concealed from them, which would give rise to causes of action for battery and fraud. Although a consent form was used beginning in 1965, it said only that the patients were participating in scientific experiments without indicating the nature or purpose of the experiment or the risks of the high-dose radiation. For subjects who received this form, a cause of action for lack of informed consent (discussed below) would be possible. In May 1999 a judge approved a settlement of the class action lawsuit for $5.4 million, which provided an average of $50,000 to each family (53).

A notable exception to these general principles lies in research conducted by the U.S. military using military personnel as subjects. The armed forces exposed soldiers to radiation during a nuclear explosion without telling them they were subjects of an experiment on radiation exposure (54). The Central Intelligence Agency gave LSD to servicemen without their knowledge to study the drug's effects (55,56). Although the Federal Tort Claims Act permits the federal government to be sued for certain acts of negligence and other torts (57), military personnel who are "injured in the course of activity incident to [military] service" are prohibited from suing the federal government for damages for personal injury (56,58). The courts justify excepting military personnel from the remedies available to civilians for the reason that allowing civil damage claims would intrude on military discipline. In addition the federal government has been protected from liability for what courts construe as military acts, including the use of investigational vaccines among soldiers serving in the Gulf War (59). The fact that the Nuremberg Code was, in effect, a common law decision written by U.S. military judges to apply to wartime experiments, has been uniformly ignored by American courts in these cases (3). Thus the common law has afforded no remedy to military personnel who are subjects of research conducted by military or quasi-military officials.

Informed Consent. The common law doctrine of informed consent grew out of the principles underlying battery — bodily integrity, autonomy, and self-determination (7,39,40,60–62). While battery applies to cases in which no consent is given at all, informed consent applies to cases in which consent is given without sufficient information to render it meaningful (38,40). A person cannot make a meaningful decision in the absence of information about the benefits, risks, and consequences of the options. Ordinary lay people are not expected to have medical or scientific knowledge necessary to determine whether or not to participate in a research study. Thus the law imposes on researchers a duty to explain the research study and its potential risks and benefits to the prospective subject.

Explanation is especially important in research that uses patients as subjects. Individuals who seek medical care may fail to appreciate the uncertainty inherent in research or may assume that they are receiving proven medical care instead of an experimental technique (29,63,64). Subjects may find it difficult to keep in mind that the investigators do not assume the role of their personal physicians (65). In addition, unlike medical care, research necessarily entails unavoidable conflicts of interest for investigators who are obligated to protect the welfare of their research subjects but may also be eager to ensure the success of the study (66). This is particularly true when physicians act as investigators and use their patients as subjects (17).

The doctrine of informed consent emphasizes that every competent adult is completely free to accept or reject any medical or scientific intervention for any reason or for no reason at all. No competent adult can be forced to undergo medical treatment, even if it is certain to save his life (67,68). It should be clear, therefore, that no one can be forced to participate in research. There is almost never any justification for withholding information about a research study. (An exception may be permitted in certain behavioral experiments in which complete information will prejudice a subject's response during the experiment, but not the decision to participate, and where the subject will suffer no risk for lack of the information (69). In *Beno v. Shalala* (70), a family sued the Secretary of Health and Human Services seeking to invalidate her granting a waiver permitting California to conduct a work-incentive demonstration project to reduce welfare benefits by 1.3 percent and waive limits on income that beneficiaries could earn without losing their welfare benefits. The court found that the beneficiaries were human subjects of research but suggested that the benefit reduction was not large enough to require the informed consent of each beneficiary under specific federal law governing demonstration experiments.)

Although informed consent is required for both research and medical treatment, disclosure requirements have been more stringent for research than for medical care. As the *Belmont Report* concluded, informed consent standards applicable to patients in malpractice cases are not sufficient for research (7). At a minimum, disclosure must include (*1*) that fact that research is being conducted; (*2*) the purpose of the research, what will happen,

and why; (*3*) the requirements of participation; (*4*) what experimental agents and techniques will be used; and (*5*) the potential risks, as well as inconveniences, to the subject of participation. A cause of action for failure to obtain informed consent is treated as a negligence action because the researchers have a professional duty to provide information sufficient to permit a prospective subject to make a voluntary, informed decision. To succeed in an action for failure to obtain informed consent, a plaintiff must prove each of the following four things:

1. The defendant had a duty to disclose certain information to the subject (usually information about the risks of participating in the study).
2. The defendant did not disclose that information.
3. The undisclosed risk or problem occurred and caused physical injury to the subject.
4. The failure to disclose the risk or problem was the proximate cause of the subject's injury because a reasonable person in the subject's circumstances would not have consented to participate in the research if he had known of the undisclosed risk.

For example, in *Halushka v. University of Saskatchewan*, a student was awarded $22,500 from a university that failed to obtain his fully informed consent to research (71). The student volunteered for a study of circulatory response and was told that a new drug would be used. He was not told that the drug was an untested anesthetic nor that a catheter would be inserted through his heart to his pulmonary artery. During the experiment, the subject's suffered cardiac arrest, was resuscitated, and remained unconscious for four days. Ultimately he dropped out of his university studies because of inability to concentrate. The appeals court held that subjects of experimentation are entitled to more information than patients must receive, including the full and frank disclosure of all facts, probabilities, and opinions that a reasonable person might be expected to consider.

The courts that have considered the issue agree that the subject of research that uses an experimental agent or procedure must be told that the agent or procedure is experimental (72–76). The cases endorsing this basic rule tend to involve patients whose physician used an experimental medical device or surgical procedure during the course of medical treatment without telling the patient that the device or procedure itself was experimental. Several cases concerned the implantation of investigational intraocular lenses. For example, in *Kus v. Sherman Hospital*, Richard Kus agreed to surgery to implant an intraocular lens in his eye (77). Dr. Vancil used an investigational implant without telling Kus that the device was experimental. The hospital's IRB had required all consent forms for the surgery to specify that the lens was experimental, but Vancil removed that information from the consent form he gave Kus and other patients. There was no question that the physician was liable for the injury to Kus's eye caused by the implant because he failed to tell Kus that the device was experimental, and Vancil settled with Kus out of court. The court found that the hospital might be liable for its own failure to obtain Kus's informed consent where the hospital, as a participant in a study governed by federal regulations, assumed a duty to ensure that informed consent was obtained for all subjects in the study.

The first implantation of a totally implantable mechanical heart in a human being, Haskell Karp, resulted in a lawsuit following Karp's death shortly after the experimental surgery (78). Karp's widow claimed that her husband had not given his informed consent to the use of the artificial heart. She contended that Dr. Cooley emphasized that he would surgically repair Karp's own diseased heart and had described the implantable artificial heart as like a heart and lung machine used to sustain life during open heart surgery. However, the court found that the Karp had actually consented to the experimental surgery and that the consent form, albeit only 179 words, noted that this device "has not been used to sustain a human being and that no assurance of success can be made." It is unlikely that courts today would be as willing to accept such a vague, abbreviated description as evidence of full disclosure for such a dramatic experiment (79).

The question of what risks to disclose can be difficult in research where not all risks can be known in advance. For example, in *Whitlock v. Duke University*, Whitlock volunteered for a deep sea diving experiment and suffered permanent organic brain damage as a result, even though the research was conducted properly (80). The court held that the investigator had a duty to warn subjects of all risks that were reasonably foreseeable when the research began. The plaintiff had signed a consent form that warned of the risks of death and other unknown risks, but did not specifically mention permanent organic brain damage. Expert evidence showed that brain damage was not a foreseeable risk and had not happened in the past. Thus the investigators could not have foreseen that specific harm and therefore had no obligation to mention it. The plaintiff had no basis for recovering damages.

Risks to the subject include psychological, social, and financial risks. For example, some types of research can expose subjects to discrimination if they are publicly identified as having socially undesirable conditions or traits, such as drug abuse, sexually transmitted diseases, or HIV infection (81). Other types of information about a subject, including drug abuse and child neglect, can lead to criminal prosecution. Research involving genetic analysis may identify susceptibility to genetic diseases, which can give rise to serious psychological and emotional concerns, as well as affecting the ability to obtain insurance or employment (82–84).

Beyond the risks to the subject himself or herself, there may be other aspects of the research that warrant disclosure. For example, the Public Health Service studied uranium miners to learn whether they were at increased risk of cancer without telling the miners of the cancer risk (85). In that case the reason the subjects were chosen was an important piece of information that should have been disclosed. Genetic research can produce probabilistic information about family members who do not want to participate in a study. Genetic research such as family pedigree studies to identify patterns of gene transmission may present risks to familial relationships.

Another type of information is the use to which information or tissue collected from the subjects will be put. In *Moore v. Regents of the University of California*, the investigator used cells extracted from a subject's spleen, removed as part of standard therapy for hairy cell leukemia, to produce a lucrative cell line later patented by the university (86). Moore sued the investigators for converting his property — his tissue — to their own profitable use without his consent. The court decided that the researcher had a duty to inform Moore of its intent to develop a cell line from his tissue but, in a controversial decision, did not allow Moore to share in the profits made from his cells, even though they had been taken for that purpose without his knowledge or consent. (The court reasoned that Moore had relinquished his ownership interest in his cells because he had not expected to keep possession of them after his spleen was removed.) The court described the physician's duty to disclose broadly to include "personal interests unrelated to the patient's health, whether research or economic, that may affect the physician's professional judgment." In the case at hand, the court presumably reasoned that henceforth a subject could refuse to participate in the research if he did not want to donate his cells to a commercial operation. Theoretically, subjects could refuse to consent to participate unless they receive a share of the profits that result. Those options may be unrealistic if the research is conducted in conjunction with a person's medical treatment, as was the case with Moore. Patients may find it difficult, financially or emotionally, to obtain treatment from another physician who is not engaged in research.

Many of the first lawsuits involving informed consent to research involved experiments intended, at least in part, to help an individual patient — often called innovative therapy — rather than an organized research project with many subjects. Until the late twentieth century, courts took a dim view of such experimentation, holding physicians accountable for injuries resulting from methods or procedures that were not generally accepted in the medical profession (39,87,88). Such experiments were considered deviations from standard medical practice (89,90). However, some experiments were tolerated if no accepted therapy worked, the patient knowingly agreed, and the physician was sufficiently skilled (91).

Today few physicians conduct isolated experiments on their own initiative. Most research is conducted more formally at large institutions. Although few lawsuits have resulted, courts no longer appear to summarily reject such research as a deviation from accepted medical practice (92). However, misrepresenting an experiment resulted in a jury verdict against several California physicians who treated AIDS patients with a drug, Viroxan, made by one of the physicians at home without FDA approval (93). The injected drug caused tissue necrosis and the patients became ill without receiving standard therapy like AZT, Bactrim, or Pentamidine. The physicians were found liable for intentionally misrepresenting the drug and other unorthodox practices as "new," safe therapy that was better than conventional treatment, with intent to defraud the patients, who were awarded $925,000 in compensatory damages. The medical center was also found liable as a co-conspirator for failing to remove the physicians from its staff after learning of their unusual remedies. This suggests that highly unorthodox methods can give rise to a cause of action for fraud, as well as lack of informed consent and medical malpractice, where the "innovation" is not fully disclosed and the patient relies on the misrepresentation to his detriment.

The 1960s and 1970s saw the growth of research as a more systematic endeavor with increasing methodological sophistication as the common law developed more rigorous protection of individual subjects' self-determination with respect to research. As a result subjects are more likely to be aware that they are involved in a research study than they were several decades ago. However, the common law still has a role to play. Research projects that use subjects who might benefit from the experimental methods or procedures — so-called therapeutic research — continue to have difficulties with ensuring that subjects are fully informed (94). Numerous groups and commentators have criticized the persistent confusion of research with treatment, among researchers as well as subjects (29,39,62,66,95). Thus the importance of ensuring that all subjects are fully informed and that their consent is competent and voluntary has not diminished.

Research With Subjects Who Cannot Consent. The general principle that no one can be used as a research subject without his or her consent assumes that individuals are legally capable or "competent" to make the decision. For example, the doctrine of informed consent applies only to legally competent adults — those who are capable of understanding their circumstances, the proposed research study and its potential risks, and making and communicating a decision to participate or not (35,36,96). (All states presume that every person over 18 years of age is legally competent, unless a court has adjudicated the person to be incompetent.) Some research, however, seeks to use subjects who are not legally competent, either temporarily or permanently. This broad category includes adults who are unable to make or communicate an informed decision because of lack of consciousness, medication, pain, developmental disabilities, or mental disorders, as well as children (97). Such research raises difficult legal, as well as ethical, questions about whether incompetent persons can participate in research at all and, if so, whether some form of surrogate authorization is necessary or sufficient (98).

There are many examples of research studies that, rightly or wrongly, used incompetent adults and children as subjects (4,29,39,97–99). The ethical justification for using such subjects is that it is impossible to discover the etiology or find a therapy for conditions that render people legally incompetent or that affect only those who lack competence — such as schizophrenia or childhood leukemia — unless the investigational modality is tested in that very group of people as research subjects (97,98). In such cases the need for research results directly conflicts with the general principle prohibiting the use of

individuals as research subjects without their voluntary, competent and informed consent (100).

The states have the sovereign authority (the *parens patriae* power) to protect the safety and welfare of incompetent adults and children. Parents of children and legal guardians of incompetent adults also have the obligation to protect the safety and welfare of their children and wards and to act in their best interests. Whether such legal representatives have the legal authority to consent to the use of their children and wards as research subjects has not been finally decided in the common law.

There is ample precedent for the principle that incompetent adults have the common law right to *refuse* to participate in research, and that their legal representatives can refuse on their behalf. This assumption is based on an analogy to the right to refuse medical treatment (60). Virtually all state courts that have considered the issue have decided that incompetent adults have the same rights as competent adults to refuse *medical treatment*, even if the treatment would save the person's life (101,102). The right to refuse medical treatment is part of the common law right to bodily integrity and self-determination — the same right the allows competent individuals to refuse to participate in research. The U.S. Supreme Court has also assumed, without deciding, that the Due Process Clause of the federal Constitution protects the right to refuse treatment (67,68). The U.S. Supreme Court has also noted that individuals who are mentally ill — including those who are involuntarily committed to a mental hospital — retain a constitutionally protected liberty interested in avoiding unwanted medication (103,104). There is general agreement that a person who refuses or objects to participate in research must not be used as a research subject (97,98).

There is very little legal precedent for the prevalent assumption that a parent or guardian can *consent* to an incompetent person's participation in research. Legal guardians are obligated to act in the best interests of the incompetent person (105). Participation in research is rarely in the best interest of a subject. Where the research is intended solely to gain generalized knowledge, no benefit can be expected to accrue to the subjects. In the case of so-called therapeutic research, where an experimental technique offers the possibility of curing a disease suffered by the person and no standard or accepted treatment has proved effective for that person, one might argue that the person might benefit from participating in the research. However, the potential risks of the research must be weighed against any benefit. It could be argued that a surrogate decision maker should be able to consent to research that offers possible benefit to an incompetent individual as long as it carries little or no risk (106). It is more difficult to argue that a legal representative could consent to potentially beneficial research that also entails significant risks. Therefore it is possible that researchers can be liable for battery or failure to obtain informed consent if they use incompetent subjects in research, whether or not they obtain surrogate consent.

There is little case law concerning research with subjects who are legally incompetent to consent to participation. Most cases in which the courts have considered the problem of incompetence involved isolated experiments related to medical treatment for an individual patient rather than a research study with a defined study population and detailed protocol. Some courts have permitted experiments on an incompetent person if a responsible family member or guardian gives informed consent and the subject himself or herself does not appear to object (105).

One of the few court decisions describing common law as well as constitutional limitations on using incompetent individuals as research subjects involved a group of patients in hospitals and facilities licensed and operated by the New York State Office of Mental Health (OMH) (107). In *T.D. v. New York State Office of Mental Health*, the patients challenged OMH regulations that allowed them to be used as research subjects even though they were considered incapable of giving informed consent. The court found that the state regulations, which provided for surrogate consent, were unlawful because they had not been approved by the Commissioner of Health. The court also found that the regulations violated the common law and constitutional rights of the patients, but the Court of Appeals, New York state's highest court, noted that this part of the decision was unnecessary and therefore an inappropriate advisory opinion. Still it suggests what courts might decide in the future. The lower court recognized that the goal of achieving important medical advances might not always be compatible with the goal of protecting human rights:

> It may very well be that for some categories of greater than minimal risk nontherapeutic experiments, devised to achieve a future benefit, there is at present no constitutionally acceptable protocol for obtaining the participation of incapable individuals who have not, when previously competent, either given specific consent or designated a suitable surrogate from whom such consent may be obtained. The alternative of allowing such experiments to continue, without proper consent and in violation of the rights of the incapable individuals who participate, is clearly unacceptable (107).

The New York regulations analyzed in *T.D.* did not apply to federally funded research that complied with federal regulations. However, federal regulations do not yet address the participation of incompetent adults. Federal regulations do permit certain types of research with children under narrowly defined conditions (25). In 1996 the Food and Drug Administration (FDA) adopted a regulation to permit research using emergency medical technologies without the consent of individuals who are incompetent to consent to participate because they are temporarily unconscious or in pain as a result of an emergency medical crisis (108). Several states prohibit research with residents of mental institutions by statute. Others permit surrogate consent in carefully defined situations. In one case California's law permitting experimentation with incompetent persons was held to be limited to research studies and could not authorize a surrogate to allow a physician to use an experimental bone graft in the treatment of an incompetent patient (109).

Thus there remain significant gaps in the law with respect to whether and, if so, how children and adults who

are incapable of giving informed consent may participate as subjects of research. Several government agencies have recognized these gaps and have recommended regulations permitting people who are not legally competent to participate as subjects of research in certain circumstances, often with the informed consent of legally authorized representatives (97,110,111). The recommendations offer slightly different answers to such questions as what types of research should be permitted, who should be eligible to give surrogate consent, and what procedures should be used to protect research subjects. Although this is an area where legislation or regulations may prove useful to clarify individual rights and duties, the proposals would apply to specific jurisdictions, such as federally funded research or research conducted within a particular state. The common law, which is less likely to permit incompetent individuals to participate in research, will still apply to research that is not governed by federal regulations or specific state laws.

Liability for the Design and Implementation of Research

Negligence. Researchers have common law duties to design and carry out research properly and can be liable for injuries suffered by research subjects that are caused by negligence in the design or implementation of the research study itself (112). The fact that a research subject has consented to participate in the study and has been informed of the foreseeable risks of participation does not mean that the subject assumes the risk of the researcher's own negligence (113).

The law of negligence requires individuals and organizations to conduct themselves as reasonably prudent persons so as to avoid causing harm to others (13). In ordinary circumstances reasonably prudent conduct can be judged by the ordinary citizen, as represented by the jury in a jury trial. The standard of conduct to be observed by professionals (e.g., physicians) in the conduct of their profession, however, is that of an expert in the field, as established by the profession itself, for ordinary citizens are not expected to be familiar with specialized knowledge and skills. Those who conduct research are legally responsible to subjects who suffer injury as a result of their negligence (39,112,113). Thus organizations that design and supervise research, as well as investigators who carry out studies with human subjects, can be liable for the injuries they cause when they fail to conform their conduct to professional standards of care. Researchers, like physicians, can also be liable for intentionally or negligently disclosing confidential personal information about research subjects or invading their privacy without their consent.

A research subject who claims injury as a result of negligence must prove the following four things:

1. The researcher had a duty to the research subject.
2. The researcher breached that duty.
3. The research subject suffered physical injury.
4. The subject's injury was caused by the researcher's breach of his duty to the subject.

Violation of a statute or regulation can sometimes be considered to be negligence per se if the injured person belongs to the class of people that the law intends to protect (114). Private individuals cannot sue researchers merely because the researchers have violated federal law governing federally funded research with human subjects (115). However, if the statute imposes specific requirements on the researcher's conduct, injured persons may sue for common law negligence and use the statute as evidence of the standard of care that the defendant should have followed (116).

Individual Investigators. The case of *Vodopest v. MacGregor* illustrates negligence by an individual investigator (117). Patricia Vodopest suffered permanent brain damage from cerebral edema in a high altitude climbing study in the Himalayas in Nepal. The study was intended to test breathing methods as a way to prevent altitude sickness. Vodopest experienced symptoms of altitude sickness at 8700 feet, but MacGregor, the project leader who was also a nurse, discounted the symptoms and told Vodopest to "breathe away" the symptoms and continue climbing higher. By 11,300 feet, Vodopest had developed cerebral edema and had to be evacuated. Vodopest was able to prove all four elements of negligence in this case. As study leader, MacGregor had a duty to act responsibly to protect the safety and welfare of the research subjects. Because the study was designed to prevent altitude sickness, the investigators should have been able to recognize its symptoms and respond with medically appropriate care. MacGregor failed to recognize (or refused to acknowledge) obvious symptoms of altitude sickness (nausea, headache, dizziness, and mental confusion). She also failed to follow the standard of care for treatment, which required having the person descend. Vodopest suffered physical injury (permanent brain damage), and the evidence indicated that her injury was proximately caused by MacGregor's negligence.

Organizations. Organizations can be held responsible for negligent injury to research subjects in two ways. First, and most simply, they can be vicariously liable for negligent acts committed by their employees and agents (118). (The employees and agents themselves also remain personally liable for their own negligence. However, employees may have insufficient assets to satisfy a large judgment or their employers may agree to pay their liability awards directly or through liability insurance.) The doctrine of *respondeat superior* (from the Latin "let the master answer" for the wrongs of his servant) holds an employer legally accountable for the unlawful acts of its employees committed during the course of employment. The purpose of the rule is to encourage employers to supervise their employees to ensure that they act responsibly. For example, the plaintiffs in *Mink*, described above, sued both the University of Chicago, where the research took place, and Eli Lilly & Company, the DES manufacturer, for vicarious liability for the acts of their employees in failing to tell plaintiffs that they were given DES as part of a medical experiment (44). In *Schwartz v. Boston Hospital for Women*, a federal district court found that a hospital could be liable for a physician's failure to obtain a subject's informed consent to a study procedure where the hospital paid the physician as assistant project

director of a study of diabetic pregnancies (119). Similarly organizations can be held liable for the negligence of those who act as agents on their behalf, where the organization controls and directs the actions of the agent.

The second basis for organizational liability is direct corporate liability for the organization's own negligent acts or violations of duty. Organizations have duties to use care in the selection, retention, and supervision of their research staff and can be responsible for injuries caused by individuals who are incompetent or should not have been hired or retained by the organization (119,120). Organizations also have a duty to maintain a safe environment, which requires them to keep their premises safe and equipment in good working order to avoid preventable injuries. Plaintiffs in the *Mink* case also claimed that the corporations breached their own duties to notify the women about the experiment when they learned that DES could cause an increased risk of cancer in the children of women who took DES (44). The court agreed that both corporations had a duty to notify the women as soon as they became aware or should have become aware of the relationship between DES and cancer. However, the women were unable to recover damages on the failure to notify claim because they could not show that they had suffered any physical injury as a result of the failure to notify.

Two subjects in a 1960s study were awarded $8 million, including $5 million in punitive damages, by a 1999 jury verdict against investigator Dr. William Sweet and Massachusetts General Hospital, where the study took place (121). The study was intended to determine whether radiation to treat brain tumors could be focused selectively on the tumor without destroying other brain tissue by using boron neutron capture to selectively attract radiation. Subjects who were terminally ill with brain cancer had a boron compound injected into their arteries to see if it collected in the tumor. The compound caused severe illness and premature death in many subjects. The investigator argued that the subjects could not be harmed because they were already terminally ill, but the jury found that the review process was negligent in allowing the study to proceed at all.

Media reports of studies inducing or allowing psychosis in patients with schizophrenia at almost a dozen medical schools in the 1980s and 1990s illustrate several possible grounds for negligence, although few lawsuits have been brought (122). A study of schizophrenic patients at the University of California at Los Angeles (UCLA) which began in the 1980s was intended to identify schizophrenic patients who could function without antipsychotic medication because long term use of certain drugs can cause tardive dyskinesia, a condition producing involuntary movements for which there is no known treatment. Antipsychotic medication was withdrawn from schizophrenic patients who had recovered from acute psychotic disorders. The patient-subjects were observed until they relapsed and exhibited severe psychiatric symptoms of "bizarre behavior, self-neglect, hostility, depressive mood [or] suicidability" (123).

One possible negligence claim might be that the researchers breached their duty of care by designing the study so as to induce subjects to experience severe psychiatric symptoms that had been controlled by their medication. Put more generally, the claim might be that the study design posed unreasonable risks and should not have been conducted in that manner or, possibly, at all. Another might be that researchers failed to adequately monitor or treat subjects who experienced symptoms during the study, allowing them to suffer unnecessarily and possibly risking irreversible deterioration. In addition it might be possible to claim failure to obtain informed consent if the subjects were not adequately informed about the nature of the study and the risks it posed. Although the subjects were asked to participate in a research study, the original consent form was ambiguous, allowing an inference that study was linked to their medical care, and did not make clear that, unlike regular patients, the subjects would not receive medication unless and until they had a severe relapse (123). By the late 1980s, 88 percent of the subjects had suffered a relapse.

After the parents of one subject complained, the federal Office for Protection from Research Risks (OPRR) in DHHS investigated and required UCLA to change some of its internal monitoring procedures and modify the consent form to point out the risks of participation and the fact that the study was not intended to meet the subject's own personal medical needs (123). OPRR also investigated psychiatric challenge and relapse studies at the University of Cincinnati, University of Maryland, Bronx Veterans Affairs Medical Center, New York State Psychiatric Institute, and the National Institutes of Mental Health, and found similar problems with researchers' informed consent practices (122,124).

The families of two subjects sued UCLA in 1992 for fraud, deceit, lack of informed consent, and civil rights violations. Gregory Aller dropped out of college, threatened his mother with a butcher knife to exorcise the devil he believed inhabited her, and tried to hitchhike to Washington, DC, to assassinate then President Bush. His parents asked the researchers to give Greg his medication, and claimed they did not do so for many months. Antonio LaMadrid jumped off a 12 story UCLA building and died about three months after participating in the study (125,126).

A study of schizophrenia subtypes took place at University Hospital in Cincinnati in the 1980s with similar results. According to newspaper reports, a patient with schizophrenia sought treatment to adjust her medication dosage to prevent a possible manic episode (126). Lacking the resources to pay for treatment, she was enrolled in a challenge study, withdrawn from her medication, given a different medication, and then placed in restraints when her manic and delusional behavior erupted. Like several other studies, this one used a medication to cause or exacerbate psychiatric symptoms. She sued the researchers, but the case was dismissed because it was not brought within the statutory time limit.

Relatively few lawsuits have focused on negligence in carrying out research, as compared with failure to obtain consent. It may be that problems are more likely to arise from expected risks or lack of consent than from poor implementation of a study. In some cases research

subjects may be partly responsible for their own injuries if they deliberately or carelessly fail to follow instructions designed for their protection and are harmed as a result. The doctrines of contributory and comparative negligence reduce, and in some cases may eliminate, the damages to which an injured person would otherwise be entitled (127).

Increased publicity about research in the 1990s may encourage closer scrutiny of research design — and perhaps legal claims of negligence — in the future, especially where researchers are pushing the scientific envelope or stand to gain financially from the success of products they study. The research design was questioned in a gene therapy trial to test — for the first time in human beings — the safety of delivering genetic material missing in people with ornithine transcarbamylase (OTC) deficiency, a sometimes fatal genetic mutation that prevents the liver from breaking down ammonia (128). Among the review committee's concerns were the use of adenovirus as the vector for transmitting the OTC gene to the liver of subjects because of adenovirus's potential to cause liver damage and toxic, sometimes fatal, inflammatory reactions. There were also reservations about the method of infusing directly into the liver instead of into a distant blood vessel. In addition the study used healthy, asymptomatic patients with OTC deficiency as the first human subjects to receive the gene, instead of patients for whom an accepted regimen of diet and medications was not effective. In 1999 Jesse Gelsinger, an 18-year-old subject, died from an inflammatory response apparently caused by the adenovirus vector. Although there is no indication that any lawsuit will be brought, the circumstances illustrate the opportunity for claims of negligence against the investigators for using excessively risky methods in designing and carrying out the research and for exposing healthy volunteers to unreasonable risks.

Conflicts of interest among researchers who have financial interests in the products they investigate may trigger or provide supporting evidence for future legal claims (129,130). James Wilson, a researcher in the OTC gene therapy study, founded a company to sell the rights to his discoveries, including the liver-directed gene vector approach studied in the OTC gene therapy trial (128). The question is whether a researcher who stands to gain financially from the success of a product might neglect evidence of the product's risks, rush to use it in human subjects prematurely, or select a research design that poses unnecessary risks to human subjects. It is not clear whether disclosure of the researcher's financial interest is sufficient to preclude a claim that the researcher acted negligently or even fraudulently.

Duties to Third Parties. Research subjects may not be the only persons injured as a result of research. However, liability for negligence is predicated on violating a legal duty to the injured person, and in the absence of any duty to third parties, there is generally no liability to third parties. Indeed, excluding some groups, especially pregnant women, from participation in research studies was often thought to ensure freedom from liability to women and their offspring (131). This raises the question whether researchers have any duty to avoid harm to future generations. If so, may subjects waive any potential claim that a future child might have for injury? Such questions are most likely to arise in connection with clinical research with fertile or pregnant women, and sometimes fertile men, using investigational drugs with teratogenic effects or genetic material that may affect a fetus or future child (132,133). Although tort law has gradually extended its application to cases of prenatal injury in which fetuses are injured by intentional or negligent conduct, the causes of action available to children remain limited and the cases rarely involve research (134). In one example, however, the University of Chicago settled a lawsuit brought by the daughters of women who had received DES as part of an experiment in the 1950s (135). The daughters claimed to have an increased risk of cancer as a result of exposure to DES in utero. Future research involving germ-line gene transfer may expose the next generation to the effects of research with today's subjects (136).

Recently it has been recognized that not using research subjects who are representative of the populations that will ultimately use or be affected by the research may also raise liability concerns if the investigational product or service is later marketed. For example, if research subjects do not include women, the research may fail to discover adverse effects unique to women. If women are injured from using the product when later marketed, they might argue that the manufacturer or researcher negligently caused their injuries by failing to use a reasonable research design to identify possible risks to women (131). Researchers have begun to respond to concerns about inappropriately underrepresenting specific groups in their study populations (132). Ethical principles for research design include the equitable selection of research subjects, and federal regulations may include more specific requirements. In the long run this may create a standard of care in research design that requires a reasonably representative study population. So far this concern has not given rise to any cause of action for individuals who are excluded from research. In the late 1990s, however, a family sued UCLA because their Asian-American child was not admitted to its experimental elementary school, claiming that the use of race as an admission criterion violated the equal protection clause of the U.S. Constitution. The students enrolled were research subjects protected by federal and state research guidelines, but the case was limited to the constitutional issue. The court found the use of race was justified in order to select a representative sample of local students to ensure the validity of the research, and that research itself served the state's compelling interest in improving urban public education (137). Just as there is no legal right to be a research subject, there is no duty on the part of researchers to include specific individuals in a particular research study. On the contrary, there may be a duty not to include vulnerable or incompetent persons in some circumstances (7,97,105,138).

Products Liability. Product manufacturers can also be liable for negligence in the manufacture, design, or distribution of their products. A cause of action for negligence parallels the four part format for negligence described above. A person who is injured by a product

must prove that the manufacturer failed to adhere to the appropriate standard of care, which failure caused the person's injury (139). Manufacturers of specialized products, such as biotechnology products, are held to the standards of an expert in the field.

In the 1960s state courts began to adopt the doctrine of strict liability which holds manufacturers liable for injuries caused by defective products without requiring proof that the manufacturer acted negligently. The justification was partly that negligence could be presumed when a product turned out to be defective and partly because it was difficult for plaintiffs to obtain evidence of a manufacturer's internal manufacturing processes to prove negligence. In theory, strict liability focuses on the condition of the product, while negligence focuses on the conduct of the manufacturer. When a person is injured by a defective product (as opposed to the behavior of individuals), they often assert both negligence and strict liability claims. Courts, however, often analyze both types of claims under products liability, a special subset of tort law that borrows negligence principles from tort law as well as warranty principles from contract law (140). Manufacturers may also be liable for breach of express warranty or implied warranty of a product's fitness for a particular purpose or merchantability (141). These causes of action, although based in contract law, impose similar duties and are usually subsumed by products liability today.

Products liability law holds all manufacturers, sellers and distributors of products legally responsible under state common law for personal injuries caused by a defect in their products (140). The earlier *Restatement* summarized the rule as follows: "One who sells any product in a defective condition unreasonably dangerous to the user or consumer or to his property [is liable] for physical harm thereby caused to the ultimate user or consumer. . . ." (142). Product defects include (*1*) manufacturing defects or flaws, in which the manufacturing process contains an error that produces something different from the product intended by the manufacturer; (*2*) defects in product design, in which the product specifications themselves pose foreseeable risks of harm that could have been avoided or reduced; and (*3*) errors or omissions in directions or warnings that accompany the product (142–144).

Manufacturing defects, such as contamination, adulteration, or production errors, are rare in pharmaceutical and biotechnology products but have occurred less rarely in medical devices (145,146). A product design is considered defective if the product could have been designed in a different way that would reduce its inherent risks without significantly decreasing its utility or effectiveness (142–144). Whether a product could have been differently designed, however, is a technical question that depends on the state of scientific knowledge when the product was sold, the nature and results of product tests, and the feasibility of alternative designs (147). An experimental product can rarely, if ever, be accused of having a defective design while it is being studied precisely to determine whether the design is safe and effective. The fact of the research is ordinarily a defense against a claim of defect. In addition it is often difficult to sort out the cause of an injury to a research subject, especially where the product has not been widely tested or the injury might result from other sources to which the research subject was exposed. The most plausible claim would be that the product was not ready for testing in human beings, and that additional laboratory or animal studies would have revealed dangers before human being were harmed, which is really a claim of negligence. Thus manufacturers are not liable for a defective design of an experimental product as long as they conduct reasonable studies and do not use a product design that is foreseeably and unnecessarily dangerous.

The recently issued *Restatement* of products liability law contains a somewhat narrower definition of design defect for prescription drugs and medical devices, which requires proof that the product produces no net benefits for *any* class of patients or that no reasonable manufacturer would produce it if it knew of the defect (148). This permits the marketing of drugs and devices with serious risks as long as they may benefit at least one class of patients. Outside the realm of ordinary consumer products and asbestos, claims of defective design have been directed primarily at lawfully marketed, nonexperimental medical devices, such as the Dalkon Shield, the Copper-7 IUD, the Bjork-Shiley heart valve, and silicon-gel breast implants (149).

Defects in product distribution include errors and omissions in product directions and warnings. Manufacturers are responsible for failing to provide to the user a warning of dangers inherent in the use of the product, or for providing an inadequate warning that failed to alert the user to the danger (150). The cause of action is similar to a cause of action against a physician or researcher for failure to obtain informed consent (151,152). The theory is that a consumer would not have used the product and suffered an injury had he or she been warned of the risk that caused the injury. Warning defects have been claimed in a substantial proportion of lawsuits against manufacturers of licensed pharmaceuticals and vaccines (153). Manufacturers of prescription drugs and vaccines that can be obtained only from a physician are not required to issue warnings directly to patients or subjects of research. The "learned intermediary" rule holds the manufacturer responsible only for providing adequate warnings to the physician, who in turn is responsible for judging the appropriateness of the medicine for an individual patient and informing the patient of its risks and benefits (154–158). In the future, however, courts may decline to apply the learned intermediary rule in the case of some prescription drugs that are marketed by direct-to-consumer advertising if the physician has only a minor role in determining whether the patient should use the drug (159,160).

Human subjects who participate in research involving investigational products should be made aware that they could be exposed to risks. If subjects are adequately informed, their consent to participate means that they assume responsibility for the risks that have been disclosed to them. The investigator's duty to obtain the voluntary informed consent of subjects also protects the product manufacturer.

Products liability law focuses on commercially marketed products — those lawfully for sale or distribution

in the market. The new *Restatement* of products liability defines products as "tangible personal property distributed commercially for use or consumption" (161), and may not even apply to investigational products. (Although human blood and tissue qualify as products when they are sold or distributed commercially, most states have enacted blood shield statutes that limit liability for contaminated human blood and human tissue to liability for negligence.) It would not ordinarily apply to investigational products that are being tested in research studies before any commercial marketing. Such investigational products, like investigational new drugs, are not held out as safe or effective commercial products ready for marketing to consumers. Rather, they are being studied to determine their effectiveness and to identify possible defects and risks. Of course, some research studies commercial products that are already on the market. For example, two marketed products, such as diagnostic tests, may be compared for relative efficacy, or a product marketed for one purpose may be studied to determine whether it has another use, such as an "off-label" use of a drug that has been approved by the FDA for a different specific use.

In *Proctor v. Davis*, the Upjohn Company was found liable for failing to warn physicians of the risks of an off-label use of its corticosteroid suspension, Depo-Medrol, by injection near the eye to treat eye conditions (162). FDA had approved the drug for intramuscular, intrajoint, and intralesional use only. The court found that "Upjohn fostered and encouraged this unapproved use as experimentation on human beings" (162). The company gave ophthalmologists financial and technical assistance to test periocular use and write favorable reports. Although the company received reports of adverse reactions, including blindness, it did not issue any warning. The company's conduct was found to justify punitive damages of just over $6 million, as well as compensatory damages of over $3 million, to the plaintiff who lost his eye after the drug was injected in an off-label use.

Independent researchers — who are not employed by a manufacturer — are not likely to be subject to products liability unless they also qualify as commercial sellers or distributors who are engaged in the business of selling or distributing the products they study. Investigators ordinarily provide services, and services are not considered products. Courts are unanimous in refusing to categorize commercially provided services as products for purposes of strict products liability in tort. Thus strict products liability does not extend to professionally provided services, such as medical or legal help (163,164).

Most states have held that public policy should exempt hospitals, physicians, and dentists from liability for injury from products used to treat their patients, such as defective pacemakers, forceps, and syringes, even though similar product/service combinations have been the subject of liability (165). The justification for this exemption is that the need for medical care outweighs any need to hold providers strictly liable for the products they use or implant in medical treatment. However, courts do not ordinarily consider research to be part of medical care, so the public policy rationale may not apply to research. A few recent cases have found that liability might be imposed where a hospital or physician selected or sold the product (166,167). A few cases have held a hospital liable for injuries resulting from defective products that were not directly related to medical care, but neither did they relate to research (168,169).

Although most legal principles apply equally in the case of experimental and marketed products, few investigational products being tested in research studies are likely to qualify as commercial products or otherwise meet the criteria for products liability claims. Most new drugs, biologics, and medical devices are developed in compliance with FDA regulations governing investigational use, so the federal rules are the primary source of law governing such research with human subjects. A violation of federal regulations may be evidence of common law negligence in most states, although states vary with respect to whether such a violation should be considered negligence per se, evidence of a deviation of the applicable standard of care, or excluded from consideration as inflammatory and irrevelant to the issue of causation.

Institutional Review Boards. Most IRBs are created by hospitals, universities, and research organizations to comply with federal law as a condition of receiving federal funding for research. The institutions agree by contract (the "general assurance") with a federal agency (typically HHS, FDA, or DOE) to be bound by federal IRB regulations specifying general membership qualifications and obligations of an IRB (170). Thus IRB duties are defined in the first instance not by common law principles but by federal statutes and regulations — specifically the Common Rule — issued by virtue of the federal spending power (10).

Nonetheless, the institution retains significant discretion over IRB operations, determining procedures, funding and personnel (11). An IRB created by a hospital or university is typically part of that organization, and not a separate legal entity, so that the organization is legally responsible for IRB actions. A growing number of IRBs, however, are independent legal entities, not part of any hospital, university, or research organization, and therefore responsible for their own legal obligations. Some have been created by community research groups and by private commercial enterprises like pharmaceutical companies to review their own research studies. Others are independently organized and offer their services to any group that wishes to conduct clinical research, usually small commercial companies.

Two decades ago Robertson considered possible liability issues that IRBs and their parent organizations might face in theory, such as defamation of an investigator or termination of employment for conduct involving research (11). The rarity of reported claims against IRBs suggests that such concerns remain largely theoretical. There have been several publicized scandals but little litigation (171,172).

In theory, possible claims by injured subjects of research against an IRB might include IRB negligence in approving a study, failure to attach conditions to the study to protect subjects, failure to require adequate information or informed consent, negligent assessment of the study's

risks and benefits, failure to review ongoing research that poses risks to subjects, failure to stop research when subjects are being harmed, and failure to notify subjects of a significant risk or harm during or after study. IRBs may have a common law duty to act with reasonable care. In addition a violation of federal regulations might be considered to be evidence of negligence in many states. To hold the IRB liable, however, a subject would have to prove that the injury would not have happened but for IRB misconduct. There are many steps between IRB approval and injuries to subjects that could negate IRB responsibility: An approved study might not be funded or carried out, important information might not have been available, investigators might disregard IRB requirements in conducting the study, investigators might have prevented the injury, or the subject might not act on warnings provided. As a Texas court noted, "Other than disapproving all or part of a study, the IRB does not and cannot control the direction, results, or use of the research" (173).

A 1996 U.S. General Accounting Office (GAO) report on IRBs serving NIH studies outlines some difficulties experienced by IRBs that could, in theory, prevent them from living up to relevant standards of conduct (17). IRB review is labor intensive and subject to considerable time pressure, with some IRBs devoting only a few minutes to reviewing a protocol. Members may rely on a single primary reviewer for their assessment of a proposed study and miss key issues in the protocol. Most IRBs studied by the GAO were composed of volunteer members and few had training in ethics or federal regulations governing the protection of human subjects. Lay volunteers may be reluctant to challenge the opinions of members with scientific backgrounds who may empathize more with other investigators than with potential subjects. Most institutions with IRBs derive substantial income from research grants and may put institutional pressure on the IRB to approve protocols, thereby creating a conflict of interest. Most IRBs were reported to spend much of their time reviewing the informed consent form rather than whether the research design posed unacceptable risks to potential subjects. In addition there was little time for continuing review of ongoing research, which might discover problems.

Despite these potential difficulties, the GAO report confirmed that there had been few complaints against institutions or IRBs that review research funded by the federal DHHS (17). FDA issued only 31 "Warning Letters" to institutions noting serious deficiencies in IRB oversight of drug research. Deficiencies included allowing researchers to participate in IRB review of their own research and false claims that research studies did not require IRB review. FDA had never disqualified an institution from submitting research studies. On the other hand, FDA issued 99 sanctions against 84 individual investigators between 1980 and 1995. Most violations were minor; serious violations included forging a subject's signature to a consent document, failing to obtain informed consent, fabricating data to make subjects appear eligible for a study, falsifying laboratory tests, and failing to report adverse reactions to investigational drugs. However, a 1999 report of 1000 FDA spot checks found that 213 researchers failed to obtain necessary informed consent, 364 researchers failed to follow their research protocol, and 140 did not report adverse reactions experienced by research subjects (174).

These reports, as well as reports with similar findings by the Office of the Inspector General, suggest that IRBs may find it difficult to perform their legal duties adequately (175,176). If improvements are not made, they could be liable in the future if their wrongful actions cause harm. At the same time other factors argue against a significant upsurge in claims. Human subjects who are injured by research are more likely to have a cause of action against, and to sue, the investigator or the institution for research conduct that is unrelated to IRB actions. An IRB's contribution to any injury may be difficult to discover. Few nonmembers appear at IRB meetings, see its records, or are in a position to identify IRB misconduct.

Federal agencies that oversee IRBs are in a good position to discover problems and act directly to halt research that may threaten human subjects. In the late 1990s OPRR began to impose sanctions on research institutions that do not comply with federal regulations. In May 1999 OPRR suspended Duke University's authorization to conduct federally funded research with human subjects (177). The action was taken only after the university had been given an opportunity to improve its system and failed to respond adequately. OPRR has imposed similar suspensions at other institutions, including the University of Rochester, University of Southern Florida, University of Minnesota, Mt. Sinai, Rush-Presbyterian-St.Lukes Medical Center in Chicago, University of Illinois in Chicago, and University of Colorado (178,179). Reasons for suspensions varied but were consistent with the problems identified in the GAO report — failures to submit studies for IRB approval, inadequate consent procedures and records, failure to keep track of research studies after they began, and failure to document why studies were approved. Enforcement action by OPRR and FDA may herald more consistent and searching scrutiny of federally funded research. It may also encourage a re-evaluation of the entire system for protecting human subjects of research, including the use of IRBs themselves.

There have been a few lawsuits by individuals attempting to discover IRB records, often to use as evidence in their lawsuits against other parties. IRBs typically treat research protocols and investigators' reports on ongoing research as confidential and do not voluntarily release them outside the IRB, except to federal agencies that are entitled by statute to review IRB compliance with a general assurance. Federal regulations prohibit the disclosure of IRB records that identify individual research subjects. The common law, however, does not specifically protect IRB records from disclosure to investigators, subjects, or the public, or discovery in a lawsuit, although some state statutes may grant protection (180). Most state statutes that grant a privilege against discovery to hospital peer review committee proceedings and records do not apply to IRBs because IRBs are not peer review committees that oversee patient care (181). Nonetheless, even courts that have permitted

disclosure of IRB records have kept confidential the names of subjects, as well as a company's proprietary information about its investigational products (181,182).

One might argue that there is a scientific or academic privilege protecting against discovery of IRB records (183,184). Research data, however, do not appear to be privileged. Data may be protected from discovery where they are not relevant or necessary to a lawsuit, where disclosure would be unduly burdensome on the researcher, or to protect a researcher's interest in completing a study or in publishing conclusions first in a peer-reviewed journal (185,186).

The fact that IRB records are not automatically protected from discovery does not necessarily mean that IRBs have any obligation to publicly disclose them. For example, an Iowa Supreme Court decision found that a public hospital (not an IRB) was not required to make public its summaries of nosocomial infection data that the hospital was required by statute to collect (187).

What counts as IRB records is not clear. They could be limited to minutes of the meetings, so that research protocols and monitoring reports would not be covered (188). A litigant might be able to obtain more information, including the research protocol, from the researcher who conducted the research.

Defenses Against Liability

Researchers charged with liability for injuries to research subjects can interpose several defenses, in addition to simply denying the facts alleged in the claim, to defeat a cause of action.

Release of Liability. Under the doctrine of informed consent, because researchers must disclose the risks of research to all subjects who agree to participate in a study, subjects assume the risk of being injured as a result of the risks that were disclosed to them. An analogous principle of contract law may allow human subjects to release researchers from liability for the foreseeable risks of participating in research, as long as those risks were disclosed and the research was properly carried out (189). However, the general rule is that human subjects cannot release researchers from tort liability for the researcher's own negligence (190). In addition federal regulations forbid the use of any document that waives or otherwise excuses, releases, or indemnifies researchers from liability for injuries to research subjects in federally funded research (191): "No informed consent, whether oral or written, may include any exculpatory language through which the subject or the representative is made to waive or appear to waive any of the subject's legal rights, or releases or appears to release the investigator, the sponsor, the institution, or its agents from liability for negligence." As a practical matter, most IRBs are thought to reject attempts to use any such releases (192).

In *Vodopest v. MacGregor*, described above, the research subject Vodopest had signed a form releasing the investigators from "all liability, claims and causes of action arising out of or in any way connected with my participation in this trek" (117). The Supreme Court of Washington found that the release was against public policy and therefore void and unenforceable because it would release researchers from their own negligence in the conduct of research. The Court noted that if Vodopest had fallen off a trail, the release might have been effective because falling is an expected risk of climbing mountains and was not necessarily part of the research. However, Vodopest's injury specifically related to the research — monitoring subjects for high altitude sickness — and therefore liability could not be released. The court concluded, "The public's interest in the safety of human subjects and the public's interest in the integrity of legitimate and necessary research militate against allowing researchers to negligently conduct research with impunity."

Sovereign Immunity. Neither the federal nor a state government can be sued without its consent. Sovereign immunity — based on the ancient notion that the king can do no wrong — can be a defense to tort claims brought against officials, employees, and agents of state and federal governmental institutions, such as state universities and city hospitals (13). The Federal Tort Claims Act permits certain tort claims to be brought against the federal government (57). All states have enacted statutes that permit certain claims to be made against the state, although some states limit the amount of damages for which a state can be held liable. For example, Massachusetts limits its own liability for personal injury to $100,000 per claim (193).

Researchers who are employees or agents of state entities may be protected by sovereign immunity if they are acting on behalf of the state. A Virginia court applied sovereign immunity to protect a researcher from liability for a research subject's death after an overdose of asthma medication (194). The researcher was employed as an allergy fellow by the University of Virginia Hospital, a state university protected by sovereign immunity. The court found that the hospital's employees and agents were also protected if they (*1*) are subject to the control of the state and have little or no control over the patients they see, and (*2*) participate in activities in which the state has a strong interest, especially where the activity is not readily available in the private sector. The court also noted that a third factor ordinarily required for sovereign immunity in Virginia — that the employees have duties that require them to exercise a substantial degree of judgment or discretion in the activity complained — is likely to exist in most medical research studies. In this case the court found that the research was important to the state. However, not all research is necessarily important to the state or anyone else, and much research can be performed as well or better by the private sector. Thus the importance of the research is not necessarily a reliable factor for predicting whether specific research will be protected by sovereign immunity elsewhere.

Physicians who are independent contractors — not employees or agents — are not protected by sovereign immunity. They may also be liable for the negligence of public hospital employees, including residents, if their supervision of those employees is negligent (195).

Sovereign immunity protected an assistant attorney general of the State of New York from liability for

concealing the fact that a man's death from an injection of synthetic mescaline was the result of a covert experiment by the U.S. Army Chemical Corps to test the drugs as a chemical warfare agent (196). The assistant attorney general received absolute immunity because he acted as the lawyer for the state psychiatric institution in a lawsuit brought by the decedent's family. However, several federal attorneys representing the U.S. Army, which was not party to the lawsuit, were not entitled to immunity for their actions in fraudulently concealing the experiment and thereby depriving family members of their cause of action for battery.

Charitable Immunity. Some private, nonprofit, charitable organizations are also protected from liability for personal injury to patients and research subjects by the doctrine of charitable immunity. Massachusetts was the first state to adopt the doctrine, in 1876, and still retains it (197,198). However, charitable immunity has been repealed in most states, and states that still apply the doctrine have crafted so many exceptions that immunity is more the exception than the rule (120). The trend appears to be in favor of holding all private organizations liable for their own negligence. Reasons for ending charitable immunity include the fact that most research organizations that qualify as "charitable" for tax purposes are no longer charities in the historical sense, and also often purchase liability insurance to protect against depleting their assets (199). In addition the nineteenth-century assumption that patients who accept charity care must also accept the risk of harm is inconsistent with the premise that all patients and research subjects are entitled to the same legal safeguards (200). Nonprofit hospitals are among the few organizations that still qualify for the remnants of charitable immunity. Most private companies that conduct research in biotechnology are not charities and would not be protected.

Federal Preemption. Federal law sometimes supersedes (preempts) state common law, so individuals cannot use the state law as the basis for a lawsuit. In the absence of a federal statute explicitly preempting state law, however, there is a general presumption that federal law does not supersede state law, especially with respect to health concerns and common law claims of negligence and product liability (201). Even where a federal statute regulates an industry, the presumption against preemption of state common law actions remains (202,203). Although some products are subject to federal labeling requirements, few such laws specifically preempt state common law claims of failure to warn, inadequate directions or warnings, or fraudulent misrepresentation (204). Thus, federal preemption of state common law claims for personal injury is rare.

The Medical Devices Amendments of 1976 contain a provision that preempts state law safety or effectiveness requirements that differ from requirements for medical devices imposed by FDA under federal law (205). The Act states that "No State or political subdivision of a State may establish or continue in effect with respect to any device intended for human use any requirement—(*1*) which is different from, or in addition to, any requirement applicable under this chapter to the device, and (*2*) which relates to the safety or effectiveness of the device or to any other matter included in a requirement applicable to the device under this chapter." (This specific provision in the statute does not apply to drugs.) In *Medtronic, Inc. v. Lohr*, the U.S. Supreme Court found that the federal law did not preempt common law claims for personal injury resulting from certain medical devices that are not subject to specific review and premarket approval by FDA (206). In that case, which did not involve research, Lora Lohr claimed that her pacemaker lead had a defect that caused a complete heart block requiring emergency surgery and also that the manufacturer had not warned of the device's tendency to fail. The device was sold pursuant to the federal law's §510(k) premarket notification procedure, which does not require FDA approval of the device or its specifications before marketing (207). The Medical Device Amendments may preempt state common law claims involving a Class III medical device, which is subject to premarket review and approval with specific FDA requirements, but the precise scope of preemption has not yet been fully addressed by the Supreme Court. Aside from its limited scope, this preemption provision may have little effect on research subjects because it is limited to claims about the medical device itself and does not necessarily preclude claims concerning the conduct of research.

The federal Biomaterials Access Assurance Act of 1998 limits the liability of suppliers of raw materials and component parts that are used in medical devices (208). Independent suppliers (who are not organizationally related to a medical device manufacturer) are liable to individuals who are injured by an implanted medical device only if the parts supplied to make the implant did not meet the supplier's or device manufacturer's specifications. Suppliers who sell raw materials and component parts that are used in many different products are not ordinarily in a position to conduct the research necessary to ensure that their materials are safe for implantation in human beings. Medical device manufacturers remain responsible for ensuring that the resulting product complies with FDA requirements and may be sued by injured users. However, the supplier may be sued for its own negligence or intentional conduct if the device manufacturer demonstrates that the supplier's actions caused an injury and the device manufacturer should not be held solely responsible or will not be able to pay the full amount of damages awarded to the injured person.

Statutes of Limitations. A defense that is available to all public and private defendants is that a claim has been brought after the time permitted by law to begin a lawsuit. All states have "statutes of limitation" or "repose" that bar the bringing of a legal action, typically two to six years after the claim arose. The purpose of such statutes is to limit a defendant's exposure to potential liability and to encourage claims to be brought while evidence is still fresh and available (13). In an unusual decision, a California trial court dismissed a suit brought against the University of California at Los Angeles for fraud, deceit,

lack of informed consent and civil rights violations in a schizophrenia study because the plaintiffs had exceeded the five-year statute of limitations for bringing the case to trial after it was filed (209).

Different legal claims or causes of action often have different maximum time periods for bringing suit. Many states have shortened the statute of limitations for medical malpractice to two years, while the limit for ordinary negligence and other personal injury claims is typically three to six years. Several courts have held that the longer personal injury statutes of limitations (and not the medical malpractice statute) apply to cases involving research injuries, because research does not constitute medical care and does not create a physician–patient relationship — even if the research is conducted by physicians (47,52,210,211).

Most states apply a "discovery rule" that extends or "tolls" the time period in which a claim may be brought (13). The time does not begin to run until the date on which the plaintiff actually discovered or should have discovered that he or she had a legal claim (212). The discovery rule might be applied in cases of fraudulent concealment, in which investigators hid the fact of research from a research subject who suffered latent injuries, as in *Mink v. University of Chicago* or *Craft v. Vanderbilt University*, or if new information indicates a previously unsuspected causal relationship between an injury and earlier research (44,47). But the rule does not excuse a plaintiff with obvious injuries from acting on reasonably available information (74).

Because children are not ordinarily empowered by law to bring suit on their own behalf, many states toll the applicable statute of limitations until a child who has been injured reaches the age of majority, typically 18 years of age (13). Research sponsors and investigators who use young children as research subjects may be exposed to potential liability for many years. Some states have limited this exception by requiring that suit must be brought within a specific number of years following the injury, usually a longer period than the otherwise applicable time periods for adults.

CONCLUSION

The common law arguably offers the most comprehensive statement of legal principles for the protection of human subjects of research and redress for their injuries. Although the Common Rule is probably more familiar and influential, the common law is more comprehensive, in theory, than either federal regulations, which apply only to federally funded research, or state statutes, which address specific issues like fetal research. Moreover the common law is the only source of redress and compensation for human subjects injured as a result of research. Although the federal government can penalize investigators and researchers for violating federal rules, federal enforcement actions do not compensate research subjects personally.

Because they are the product of court decisions in particular disputes, common law principles are not easily summarized in a textbook. Caution must be used in extracting lessons from individual cases with particular facts. Reasonable people can disagree about precisely what is and is not required of researchers in specific circumstances. Moreover several issues have yet to be fully addressed in court decisions, including whether, and if so, when and how, individuals who are not legally competent might participate as human subjects of research; what, if any, duties investigators owe to people who are not subjects of their research; and whether, and if so, what kinds of conflicts of interest preclude researchers from conducting certain research.

Although the research enterprise — and the number of research subjects — is large and growing, there have been relatively few published court decisions addressing common law claims arising out of research. A majority of such lawsuits have complained of failure to tell subjects that research was being conducted at all. Few cases have been based on negligence in the conduct of research or invasion of privacy. In the absence of comprehensive empirical data on violations of legal obligations and injuries arising from research, it is hazardous to draw general conclusions about the incidence of particular problems from the few published court decisions and media reports.

Greater public visibility for research may focus new attention on the common law rights of research subjects and increase the potential for legal claims in the future. Publicity about cold war era radiation experiments conducted without the subjects' knowledge served to focus public awareness on the rights of human subjects in the 1990s and may have inspired lawsuits based on common law causes of action to compensate injured subjects or their families. A series of studies identifying problems and gaps in the current federal regulatory system may have encouraged federal agencies to initiate bolder enforcement of existing federal regulations governing federally funded research. Although these federal initiatives do not alter common law, they may lead to a more comprehensive restructuring of statutory legal protections for human subjects, which in turn may affect the scope of common law protections. The rise of biotechnology, new forms of financing research, innovative financial relationships between academe and private industry, and public–private partnerships between government and commercial enterprises all create new opportunities for research and for research subjects. These may foster a new era of public concern for protecting human subjects in research. Whatever form that concern takes, it is likely that the common law will remain the legal backstop for the rights of human subjects.

BIBLIOGRAPHY

1. G. Hughes, in A.B. Morrison, ed., *Fundamentals of American Law*, Oxford University Press, New York, 1996, pp. 9–33.
2. G.J. Annas and M.A. Grodin, eds., *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, New York, 1992, pp. 94–104.
3. G.J. Annas, in G.J. Annas and M.A. Grodin, eds., *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, New York, 1992, pp. 201–222.

4. J. Katz, *Experimentation with Human Beings*, Russell Sage Foundation, New York, 1972.
5. Tuskegee Syphilis Study Ad hoc Advisory Panel, Final Report, U.S. Department of Health, Education and Welfare, Washington, DC, 1973.
6. National Research Act of 1974, 88 Stat. 342, codified, as amended, at 42 U.S.C. 201-300aaa-13.
7. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, DHEW Publ. No. (OS) 78-0012, U.S. Government Printing Office, Washington, DC, 1978.
8. W.J. Curran, *Daedalus: Ethical Aspects Exp. Hum. Subj.* **98**, 542 (1969).
9. L.H. Glantz, in G.J. Annas and M.A. Grodin, eds., *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, New York, 1992, pp. 183–200.
10. Office of Science and Technology Policy, Federal Policy for the Protection of Human Subjects, *Fed. Regist.* **56**, 28002–28032 (1991).
11. J.A. Robertson, *UCLA Law Rev.* **26**, 484–549 (1979).
12. O.W. Holmes, *The Common Law*, Little, Brown, Boston, MA, 1881 (1963 ed.).
13. W.P. Keeton and W.L. Prosser, eds., *Prosser and Keeton on the Law of Torts*, 5th ed., West Publication, St. Paul, MN, 1984.
14. G. Williams, *Curr. Leg. Probl.* **4**, 137–146 (1951).
15. H.K. Beecher, *N. Engl. J. Med.* **274**, 1354–1360 (1966).
16. D.J. Rothman, *N. Engl. J. Med.* **317**, 1195–1199 (1987).
17. U.S. General Accounting Office (USGAO), *Scientific Research: Continued Vigilance Critical to Protecting Human Research Subjects*, GAO/HEHS-96-72, USGAO, Washington, DC, 1996.
18. U.S. Department of Health, Education and Welfare (USDHEW), *Secretary's Task Force on the Compensation of Injured Research Subjects*, Report, USDHEW, Washington, DC, 1977.
19. P.V. Cardon, F.W. Dommel, Jr., and R.R. Trumble, *N. Engl. J. Med.* **295**, 650–654 (1976).
20. T.A. Brennan et al., Incidence of adverse events and negligence in hospitalized patients — Results of the Harvard Medical Practice Study I. *N. Engl. J. Med.* **324**, 370–376 (1991).
21. A.R. Localio et al., *N. Engl. J. Med.* **325**, 245–251 (1991).
22. B. Freedman, *N. Engl. J. Med.* **317**, 141–145 (1987).
23. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Compensating for Research Injuries: The Ethical and Legal Implications of Programs to Redress Injured Subjects*, vol. 1, U.S. Government Printing Office, Washington, DC, 1982.
24. Council for International Organizations of Medical Sciences, World Health Organization, *International Ethical Guidelines for Biomedical Research Involving Human Subjects*, CIOMS, Geneva, 1993.
25. U.S. Department of Health and Human Services, Protection of human subjects, *Code of Federal Regulations*, Title 45, Part 46.
26. R.N. Proctor, *Racial Hygiene: Medicine under the Nazis*, Harvard University Press, Cambridge, MA, 1988.
27. R.J. Lifton, *Nazi Doctors*, Basic Books, New York, 1986.
28. J.H. Jones, *Bad Blood: The Tuskegee Syphilis Experiment*, Free Press, New York, 1981.
29. Advisory Committee on Human Radiation Experiments, *Final Report*, U.S. Government Printing Office, Washington, DC, 1996.
30. U.S. Congress, General Accounting Office, *Human Experimentation: An Overview on Cold War Era Programs*, Testimony Before the Legislation and National Security Subcommittee, Committee on Government Operations, House of Representatives, U.S. Congress, GAPO/T-NSIAD-94-266, U.S. Government Printing Office, Washington, DC, 1994.
31. U.S. Congress, House of Representatives, Committee on Energy and Commerce, Subcommittee on Energy Conservation and Power, *American Nuclear Guinea Pigs: Three Decades of Radiation Experiments on U.S. Citizens, Report*, U.S. Government Printing Office, Washington, DC, 1986.
32. R.M. Veatch, *The Patient as Partner: A Theory of Human Experimentation Ethics*, Indiana University Press, Bloomington, 1987.
33. H.K. Beecher, *J. Am. Med. Assoc.* **195**, 34–35 (1966).
34. *Union Pacific Railroad Co. v. Botsford*, 141 U.S. 250, 251 (1891).
35. *Scholendorff v. Society of New York Hospital*, 211 N.Y. 125 (1914).
36. *Canterbury v. Spence*, 464 F.2d 772, 780 (D.C. Cir.), *cert. denied*, 409 U.S. 1064 (1972).
37. *Natanson v. Kline*, 186 Kan. 393, 350 P.2d 1093 (1960).
38. *Cobbs v. Grant*, 8 Cal. 3d 229, 502 P.2d 1, 104 Cal. Rptr. 505 (1972).
39. G.J. Annas, L.H. Glantz, and B.F. Katz, *Informed Consent to Human Experimentation: The Subject's Dilemma*, Ballinger, Cambridge, MA, 1977.
40. R.R. Faden and T.L. Beauchamp, *A History and Theory of Informed Consent*, Oxford University Press, New York, 1986.
41. American Law Institute, *Restatement of the Law, Second, Torts*, §§13–20, ALI, Philadelphia, PA, 1965.
42. J.C. Fletcher, in K. Berg and K.E. Tranoy, eds., *Research Ethics*, Liss, New York, 1983, pp. 187–228.
43. *Hyman v. Jewish Chronic Disease Hospital*, 15 N.Y.2d 317 (1965), rev'g 21 A.D.2d 495, 251 N.Y.S.2d 818 (2d Dep't 1964).
44. *Mink v. University of Chicago*, 460 F. Supp. 713 (N. Dist. Ill. 1978).
45. American Law Institute, *Restatement of the Law, Second, Torts*, §§652A, 652B, 652D, 652I, ALI, Philadelphia, PA, 1977.
46. American Law Institute, *Restatement of the Law, Second, Torts*, §§304, 310, 525-530, 550, 551, 557A, ALI, Philadelphia, PA, 1977.
47. *Craft v. Vanderbilt University*, 18 F. Supp. 2d 786 (M.D. Tenn. 1998).
48. *Craft v. Vanderbilt University*, 174 F.R.D. 396 (M.D. Tenn. 1996).
49. *Boston Globe*, July 28, p. A5 (1998).
50. *Massachusetts Lawyers Weekly*, May 3, B6 (1999).
51. *Stadt v. University of Rochester*, 921 F. Supp. 1023 (W.D.N.Y. 1996).
52. *In re Cincinnati Radiation Litigation*, 874 F. Supp. 796 (S.D. Ohio 1995).
53. Associated Press, Judge Approves $5 Million Settlement in Radiation Case, May 5, 1999.

54. *Jaffee v. United States*, 663 F.2d 1226 (3d Cir. 1981), *cert. denied*, 456 U.S. 972 (1982).
55. *Central Intelligence Agency v. Sims*, 471 U.S. 159 (1985).
56. *United States v. Stanley*, 483 U.S. 669 (1987).
57. Federal Tort Claims Act, 28 U.S.C. 1348.
58. *Feres v. United States*, 340 U.S. 135 (1950).
59. *Doe v. Sullivan*, 938 F.2d 1370 (D.C. 1991).
60. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Making Health Care Decisions*, vols. 1, 3, U.S. Government Printing Office, Washington, DC, 1982.
61. R.J. Levine, *Ethics and Regulation of Clinical Research*, 2nd ed., Urban & Schwarzenberg, Baltimore, MD, 1986.
62. A.M. Capron, *Univ. Penn. Law Rev.* **123**, 341–438 (1974).
63. J. Katz, *St. Louis Univ. Law J.* **38**, 7–54 (1980).
64. R.J. Levine, *Arch. Intern. Med.* **143**, 1229–1231 (1983).
65. G.J. Annas, *Standard of Care: The Law of American Bioethics*, Oxford University Press, New York, 1993.
66. P.S. Appelbaum et al., *Hastings Cent. Rep.* **17**, 20–24 (1987).
67. *Cruzan v. Director, Missouri Department of Health*, 497 U.S. 261 (1990).
68. *Vacco v. Quill*, 521 U.S. 793 (1997).
69. T.L. Beauchamp and J.F. Childress, *Principles of Biomedical Ethics*, 4th ed., Oxford University Press, New York, 1994.
70. *Beno v. Shalala*, 853 F. Supp. 1195 (E.D. Cal. 1993).
71. *Halushka v. University of Saskatchewan*, 1965 D.L.R2d. 436, 52 W.W.R. 608 (Sask. 1965).
72. *Estrada v. Jaques*, 70 N.C. App. 627, 321 S.E.2d 240 (N.C. App. 1984).
73. *Ahern v. Veterans Administration*, 537 F. 2d 1098 (10th Cir. 1978).
74. *Monroe v. Harper*, 164 Mont. 23, 518 P.2d 788 (1974).
75. *Clemens v. Regents of the University of California*, 87 Cal. Rptr. 108 (Cal. Ct. App. 1970).
76. *Wilson v. Scott*, 412 S.W. 2d 299 (Tex. 1967).
77. *Kus v. Sherman Hospital*, 268 Ill. App. 3d 771, 744 N.E.2d 1214 (Ill. App. 1995).
78. *Karp v. Cooley*, 493 F.2d 408 (5th Cir.), *cert. denied*, 419 U.S. 845 (1974).
79. G.J. Annas, Death and the magic machine: Consent to the artificial heart. In G.J. Annas, *Standard of Care: The Law of American Bioethics*, Oxford University Press, New York, 1993, pp. 198–210.
80. *Whitlock v. Duke University*, 637 F. Supp. 1463 (M.D.N.C. 1986), aff'd, 829 F.2d 1340 (4th Cir. 1987).
81. W.K. Mariner, *Eval. Rev.* **14**, 538–564 (1990).
82. M. Chapman, *Am. J. Hum. Gene.* **42**, 491–498 (1992).
83. K. Rothenberg, *J. Law, Med. Ethics* **23**, 312–319 (1995).
84. K. Rothenberg, *Science* **275**, 1755–1757 (1997).
85. *Begay v. United States*, 591 F. Supp. 991 (D. Ariz. 1984), *aff'd*, 768 F.2d 1059 (1985).
86. *Moore v. Regents of the University of California*, 51 Cal. 3d 120, 793 P.2d 479, 271 Cal. Rptr. 146 (1990).
87. *Carpenter v. Blake*, 60 Barb. 488 (N.Y. Sup. Ct. 1871).
88. D.J. Rothman, *Strangers at the Bedside: A History of How Law and Bioethics Transformed Medical Decision Making*, Basic Books, New York, 1991.
89. *Owens v. McCleary*, 281 S.W. 682 (Mo. 1926).
90. *Fortner v. Koch*, 261 N.W. 762 (Mich. 1935).
91. *Stammer v. Board of Regents*, 39 N.E.2d 913 (N.Y. 1942).
92. *In re Guess*, 393 S.E.2d 833 (N.C. 1990), *cert. denied*, 498 U.S. 1047 (1991).
93. *Looney v. AMI Medical Center of North Hollywood* (1994), *Jury Verdict Review Publications, Inc.*, Feb. 1995.
94. N.M.P. King and G. Henderson, *Mercer Law Rev.* **42**(3), 1007–1050 (1991).
95. N.M.P. King, *Hastings Cent. Rep.* **25**, 6–15 (1995).
96. T. Grisso and P.S. Appelbaum, *Assessing Competence to Consent to Treatment: A Guide for Physicians and Health Care Professionals*, Oxford University Press, New York, 1998.
97. National Bioethics Advisory Commission (NBAC), *Research Involving Persons with Mental Disorders That May Affect Decisionmaking Capacity, vol. I*, Available at: NBAC, Rockville, MD, 1998. *http://bioethics.gov/capacity/TOC.htm*
98. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *Research Involving Those Institutionalized as Mentally Infirm*, U.S. Department of Health, Education and Welfare, Washington, DC, 1978.
99. S.E. Lederer and M.A. Grodin, in M.A. Grodin and L.H. Glantz, eds., *Children as Research Subjects—Science, Ethics, and Law*, Oxford University Press, New York, 1994, pp. 3–25.
100. A.M. Capron, *N. Engl. J. Med.* **340**(18), 1430–1434 (1999).
101. *Matter of Quinlan*, 70 N.J. 10, 335 A.2d 647, *cert denied*, 429 U.S. 922 (1976).
102. *Superintendent of Belchertown State School v. Saikewicz*, 373 Mass. 728, 370 N.E.2d 417 (1977).
103. *Washington v. Harper*, 494 U.S. 210 (1990).
104. *Riggins v. Nevada*, 504 U.S. 127 (1992).
105. L.H. Glantz, in M.A. Grodin and L.H. Glantz, eds., *Children as Research Subjects—Science, Ethics, and Law*, Oxford University Press, New York, 1994, pp. 103–130.
106. R. Michels, *N. Engl. J. Med.* **340**(18), 1427–1430 (1999).
107. *T. D. v. New York State Office of Mental Health*, 228 A.D.2d 95, 650 N.Y.S.2d 173 (1st Dept 1996), *app. dism'd mem. opinion*, 89 N.Y.2d 1029 (1997).
108. 21 CFR §50.23.
109. *Trantafello v. Medical Center of Tarzana*, 227 Cal. Rptr. 84 (Cal. Ct. App. 1976).
110. Office of the Maryland Attorney General, *Final Report of the Attorney General's Research Working Group*, June 12, 1998.
111. State of New York Department of Health Advisory Work Group on Human Subject Research Involving Protected Classes, *Recommendations on the Oversight of Human Subject Research Involving Protected Classes*, State of New York, New York, 1998.
112. N. Hershey and R.D. Miller, *Human Experimentation and the Law*, Aspen Systems Corp., Germantown, MD, 1976.
113. I. Ladimer and R.W. Newman, eds., *Clinical Investigation in Medicine: Legal, Ethical, and Moral Aspects*, Law-Medicine Research Institute, Boston University, Boston, MA, 1963.
114. *Johnson v. Farmers & Merchants State Bank*, 320 N.W.2d 892 (Minn. 1992).
115. 21 U.S.C. §337(a).
116. *Femrite v. Abbott Northwestern Hospital*, 568 N.W.2d 535 (Minn. App. 1997).
117. *Vodopest v. MacGregor*, 128 Wash. 2d 840, 913 P.2d 779 (1996).

118. K.R. Wing, M.S. Jacobs, and P.C. Kuszler, *The Law and American Health Care*, Aspen Publishers, New York, 1998.
119. *Schwartz v. Boston Hospital for Women*, 422 F. Supp. 53 (S.D.N.Y. 1976).
120. *Darling v. Charlestown Memorial Hospital*, 33 Ill. 2d 326, 211 N.E.2d 253 (1965), *cert. denied*, 383 U.S. 946 (1966).
121. *Heinrich v. Sweet*, No. 97-12134-WGY (D. Mass. Oct. 15, 1999).
122. R. Whitaker and D. Kong, *Boston Globe*, November 15, p. 1 (1998).
123. J. Katz, *St. Louis Univ. Law J.* **38**(1), 7–54 (1993) (quoting the research protocol).
124. D. Kong, *Boston Globe*, February 9, p. A8 (1999).
125. J. Willwerth, *Time* **42**, 21–42 (1992).
126. D. Kong, *Boston Globe*, November 16, p. 1 (1998).
127. American Law Institute, *Restatement of the Law, Second, Torts*, §463, ALI, Philadelphia, PA, 1977.
128. D. Nelson and R. Weiss, *Washington Post*, November 21, p. A01 (1999).
129. K. Eichenwald and G. Kolata, *N.Y. Times*, November 30, pp. A1, C16 (1999).
130. K. Eichenwald and G. Kolata, *N.Y. Times*, May 16, p. A1 (1999).
131. E. Flannery and S.N. Greenberg, in A.C. Mastroianni, R. Faden, and D. Federman, eds., *Women and Health Research — Ethical and Legal Issues of Including Women in Clinical Studies*, vol. 2, National Academy Press, Washington, DC, 1994, pp. 91–102.
132. A.C. Mastroianni, R. Faden, and D. Federman, eds., *Women and Health Research — Ethical and Legal Issues of Including Women in Clinical Studies*, vol. 2, National Academy Press, Washington, DC, 1994.
133. L. Walters and J. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, 1997.
134. E.W. Clayton, in A.C. Mastroianni, R. Faden, and D. Federman, eds., *Women and Health Research — Ethical and Legal Issues of Including Women in Clinical Studies*, vol. 2, National Academy Press, Washington, DC, 1994, pp. 103–112.
135. *Wetherill v. University of Chicago*, 565 F. Supp. 1553 (N.D.Ill. 1983).
136. P. Kitcher, *The Lives to Come*, Simon & Schuster, New York, 1996.
137. *Hunter v. Regents of the University of California*, 971 F. Supp. 1316 (C.D.Cal. 1997).
138. *Bailey v. Lally*, 481 F. Supp. 203 (D.Md. 1979).
139. American Law Institute, *Restatement of the Law, Second, Torts*, §282, 388-390, 394-398, ALI, Philadelphia, PA, 1977.
140. American Law Institute, *Restatement of the Law, Third, Torts: Products Liability*, §1, Comment a, ALI, Philadelphia, PA, 1998.
141. *Mitchell v. Collagen Corporation*, 126 F2d 902 (7th Cir. 1997).
142. American Law Institute, *Restatement of the Law Second, Torts*, §402A, ALI, Philadelphia, PA, 1977.
143. American Law Institute, *Restatement of the Law, Third, Torts: Products Liability*, §2, ALI, Philadelphia, PA, 1998.
144. W.P. Keeton, *Mo. Law Rev.* **45**, 579–596 (1980).
145. Institute of Medicine, *Vaccine Supply and Innovation*, National Academy Press, Washington, DC, 1985.
146. *Martin v. Telectronics Pacing Systems, Inc.*, 105 F.3d 1090 (6th Cir. 1997).
147. J.A. Henderson, Jr. and A.D. Twerski, *Hofstra Law Rev.* **26**(11), 667–695, 672 (1998).
148. American Law Institute, *Restatement of the Law, Third, Torts: Products Liability*, §6, ALI, Philadelphia, PA, 1998. (No examples of experimental products are given.)
149. S. Garber, *Product Liability and the Economics of Pharmaceuticals and Medical Devices*, R-4285-ICJ, Rand Institute for Civil Justice, Santa Monica, CA, 1993.
150. P.D. Rheingold, *Rutgers Law Rev.* **18**, 947–1018 (1964).
151. J.A. Henderson, Jr. and A.D. Twerski, *Product Liability: Problems and Processes*, 3rd ed., Aspen Law & Business, New York, 1997, pp. 342–350.
152. M.A. Franklin and R.L. Rabin, *Cases and Materials on Tort Law and Alternatives*, 6th ed., Foundation Press, Westbury, NY, 1996, pp. 608–614.
153. W.K. Mariner, in Office of Technology Assessment, Congress of the United States, *Adverse Reactions to HIV Vaccines: Medical, Ethical, and Legal Issues*, OTA-BP-H-163, U.S. Government Printing Office, Washington, DC, 1995, pp. 79–159.
154. *Hill v. Searle Laboratories*, 884 F.2d 1064 (8th Cir. 1989).
155. *Kirk v. Michael Reese Hospital and Medical Center*, 117 Ill. 2d 507, 513 N.E.2d 387 (Ill. 1987), *cert. denied*, 485 U.S. 905 (1988).
156. *Plummer v. Lederle Laboratories*, 819 F.2d 349 (2d Cir.), *cert. denied*, 484 U.S. 898 (1987).
157. *Payne v. Soft Sheen Products, Inc.*, 486 A.2d 712 (D.C. 1985).
158. *In re Norplant Contraceptive Product Liability Litigation v. American Home Products Corp.*, 165 F.3d 374 (1999).
159. *Edwards v. Basel Pharmaceuticals*, 116 F.3d 1341 (10th Cir. 1997).
160. *Perez v. Wyeth Laboratories, Inc.*, 161 N.J. 1, 734 A.2d 1245 (1999).
161. American Law Institute, *Restatement of the Law, Third, Torts: Products Liability*, §19, ALI, Philadelphia, PA, 1998.
162. *Proctor v. Davis*, 291 Ill. App. 3d 265, 683 N.E.2d 1203, 1212 (1997).
163. American Law Institute, *Restatement of the Law, Third, Torts: Products Liability*, §19, Comment j, Philadelphia, PA, 1998.
164. *Hoven v. Kelble*, 256 N.W.2d 379 (Wis. 1977).
165. *Estate of Hannis v. Ashland State General Hospital*, 554 A.2d 574 (Pa. Cmwlth. Ct. 1989).
166. *Bell v. Poplar Bluff Physicians Group*, 879 S.W.2d 618 (Mo. Ct. App. 1994).
167. *Parker v. St. Vincent Hospital*, 919 P.2d 1104 (N.M. Ct. App. 1996).
168. *Thomas v. St. Joseph Hospital*, 618 S.W.2d 791 (Tex. Civ. App. 1981).
169. *Johnson v. Sears, Roebuck & Co.*, 355 F.Supp. 1065 (E.D. Wis. 1973).
170. National Institutes of Health (NIH), *Protecting Human Subjects: Institutional Review Board Guidebook*, NIH, Bethesda, MD, 1993.
171. *Science* **189**, 383–396 (1977).
172. R. Weiss, *Washington Post*, August 1, p. A1 (1998).
173. *Cason v. E.I. Dupont de Nemours & Co.*, 1997 Tex. App. LEXIS 811, *62 (Ct. App. Tex. 1997).
174. S. Kaplan and S. Brownlee, *U.S. News and World Report*, October 11, pp. 34–38 (1999).
175. Office of the Inspector General, Department of Health and Human Services, *Institutional Review Boards: A Time for*

Reform, Testimony before the Committee on Government Reform and Oversight, Subcommittee on Human Resources, U.S. House of Representatives, Washington, DC, 1998.

176. Office of Inspector General, Department of Health and Human Services, *Investigational Devices: Four Case Studies*, OEI-05-94-00100, USDHHS, Washington, DC, 1995.
177. R. Weiss, *Washington Post*, May 12, 1999 Available at: *http://www.washingtonpost.com/wp-srv/national/daily/may99/duke12.htm*
178. A. Scharader, *Denver Post*, November 1 (1999).
179. L. Song, *The Chicago Tribune*, October 19 (1999).
180. *Doe v. Illinois Masonic Medical Center*, 297 Ill. App. 3d 240, 696 N.E.2d 707 (Ill. App. Ct. 1998).
181. *Konrady v. Oesterling*, 149 F.R.D. 592 (D.Minn. 1993).
182. *Deitchman v. E.R. Squibb & Sons, Inc.*, 740 F.2d 556 (7th Cir. 1984).
183. *In re American Tobacco Company*, 880 F.2d 1520 (2d Cir. 1989).
184. *Plough Inc. v. National Academy of Sciences*, 530 F.2d 1152 (D.C. Cir. 1987).
185. *Burka v. U.S. Department of Health and Human Services*, 87 F.2d 508, 520 n.13, 318 U.S. App. D.C. 274 (D.C. Cir. 1996) (collecting cases).
186. *Dow Chemical Company v. Allen*, 672 F.2d 1262 (7th Cir. 1982).
187. *Burton v. University of Iowa Hospital & Clinics*, 566 N.W. 2d 183 (Iowa 1997).
188. *Washington Research Project, Inc. v. Department of Health, Education, and Welfare*, 504 F.2d 238 (1974), *cert. denied*, 421 U.D. 963 (1975).
189. *Colton v. New York Hospital*, 98 Misc. 2d 957, 414 N.Y.S.2d 866 (1979).
190. American Law Institute, *Restatement of the Law, Second, Contracts*, §195, ALI, Philadelphia, PA, 1997.
191. 45 CFR 46.116.
192. W.F. Bowker, Exculpatory language in consent forms (letter to editor). *IRB: Rev. Hum. Sub. Res.*, March, pp. 1, 9 (1982).
193. Massachusetts General Laws, c. 258, §2.
194. *Hicks v. Pollart*, 27 Va. Cir. 7 (Va. Cir. Ct. 1991).
195. *Dang v. St. Paul Ramsey Medical Center, Inc.*, 490 N.W.2d 653 (Minn. App. 1992).
196. *Barrett v. United States*, 798 F.2d 565 (2d Cir. 1986).
197. *McDonald v. Massachusetts General Hospital*, 120 Mass. 432 (1876).
198. Massachusetts General Laws, c. 231, §85K.
199. *Harv. Law Rev.* **100**, 1382–1399 (1987).
200. *Malloy v. Fong*, 37 Cal. 2d 356, 232 P.2d 241 (1951).
201. *Levesque v. Miles, Inc.*, 816 F. Supp. 61 (D.N.H. 1993) (collecting cases).
202. *Rice v. Santa Fe Elevator Corp.*, 331 U.S. 218, 230 (1947).
203. *New York Conference of Blue Cross & Blue Shield Plans v. Travelers Ins. Co.*, 514 U.S. 645 (1995).
204. *Cipollone v. Liggett Group, Inc.*, 505 U.S. 504 (1992).
205. 21 U.S.C. §360(k)(a).
206. *Medtronic, Inc. v. Lohr*, 518 U.S. 470 (1996).
207. D.A. Kessler, S.M. Pape, and D.N. Sundwall, The federal regulation of medical devices. *N. Engl. J. Med.* **317**, 357–366 (1987).
208. Biomaterials Access Assurance Act of 1998, 21 U.S.C. §§1601–1606.
209. D. Anderson, Judge dismisses lawsuit over schizophrenia patients. *Associated Press Report*, June 5, WL 4869407 (1997).
210. *Collins v. Thakkar*, 552 N.E.2d 507 (Ind. Ct. App. 1990).
211. *Payette v. Rockefeller University*, 220 A.D.2d 69, 643 N.Y.S.2d 79 (1996).
212. *Baird v. American Medical Optics*, 301 N.J. Super. 7, 693 A.2d 904 (1997).

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, LAW, FDA RULES

BARUCH BRODY
Baylor College of Medicine
Houston, Texas

OUTLINE

INTRODUCTION

In the United States, as in most other countries, a new drug cannot be marketed unless a national regulatory agency has determined that it is safe in, and effective for, its intended use. In order to establish that a new drug is safe and effective, research must be carried out on human subjects. In the United States, as in most other countries, this preliminary research on human subjects must be carried out with the concurrence of a national regulatory agency. The Food and Drug Administration (FDA) is the national regulatory agency that carries out both of these functions (concurring with the research and approving the drug as safe and effective) in the United States; it has other functions as well. FDA has issued many important rules governing the performance of both functions. These rules, and the ethical and policy issues raised by them, are the focus of this article.

Each FDA rule raises its own set of ethical and policy issues. But there are two themes that run through FDA's treatment of rules and it is helpful to identify them in advance. The first is the recognition that complex ethical and policy issues are best resolved by recognizing the legitimacy of many values, even if they are sometimes conflicting values, and by attempting to formulate rules that properly balance these different values. The second is the recognition that participation in research can be both a benefit and a burden to subjects, and the rules governing research need to reflect this dual nature of participation.

BRIEF HISTORY OF FDA

A series of national scandals led to the legislation that created FDA as we know it today (1,2,3). In 1906, in response to a national outcry related to false labeling and marketing of patent medicines, Congress passed legislation creating an agency to deal with adulteration or mislabeling of drugs. In 1938, in response to the sale of a liquid form of sulfanilamide which turned out to be poisonous and caused the death of more than 100 children, Congress passed a law prohibiting the sale of drugs in interstate commerce until the seller had submitted to that agency a New Drug Application (NDA) that demonstrated that the drug was safe in its intended use. In 1962, in response to the worldwide outbreak of phocomelia in children whose mothers had taken thalidomide during pregnancy, Congress passed a law requiring (*1*) that no drug be tested in human subjects until its sponsors submitted to that agency an Investigational New Drug (IND) application and (*2*) that no drug be approved for sale until the seller demonstrated in its NDA that the drug is effective as well as safe. FDA, as it exists today, is largely the product of these three legislative acts. Because these legislative acts were passed in response to scandals arising out of new drugs hurting those who used them, their emphasis was on the value of protecting research subjects and the general public.

FDA is also the product of its own regulatory responses to public criticisms. In response to criticisms raised through the 1970s and early 1980s that the FDA process took too long and resulted in useful drugs being available elsewhere but not in the United States (the "drug lag" claim), FDA issued in 1985 and 1987 new regulations, the NDA Rewrite (4) and the IND Rewrite (5), that were designed to speed up the approval process. In response to the demand for quicker access to drugs by desperate Acquired Immune Deficiency Syndrome (AIDS) and cancer patients, FDA developed in the late 1980s and the early 1990s the Treatment IND program and the accelerated approval programs. In response to the claim that it was insensitive to the needs of special populations, it developed in the 1990s policies relating to emergency room patients, geriatric patients, pediatric patients, and patients who were women of childbearing potential. Many of these regulatory changes were incorporated into the Food and Drug Administration Modernization Act of 1997 (6). Because these regulatory responses, and the resulting legislation, arose in response to concerns about overregulation and overprotection, their emphasis was on widening access of research subjects to the benefits of research and on speeding the availability of the results of research to the general public.

CURRENT PROCESS FOR NEW APPROVALS

The sponsor of a new drug begins the process by submitting an IND application to the FDA. The application must include information about the composition of the drug, information about preclinical testing of the drug (including animal studies), information about all proposed protocols for research on humans, information about the approval of those protocols by an independent Institutional Review Board (IRB), and information about the informed consent process proposed in such protocols. In certain special circumstances involving emergency research, that last requirement may be waived. The protocols may be for Phase I studies in a limited number of subjects designed primarily to study the effects of increasing dosages of the drug. For such studies, the FDA's focus in its review of the application will be on the safety of the proposed protocols. The protocols may be for Phase II or Phase III studies in larger numbers of patients designed to study effectiveness and the overall benefit–risk ratio of the drug. For such studies FDA's focus in its review of the application will also include the scientific quality of the studies to see whether they can generate the data that are sufficient to support an application for marketing approval.

Unless the FDA objects within 30 days (or a longer period if further information is requested of the sponsors), research on human subjects may commence. During that research period, the new drug is only to be provided to research subjects under the approved protocols. There has been a traditional exception for emergency use authorized by the FDA for an individual case. In recent years, in response to the AIDS crisis, the FDA has developed a Treatment IND Program under which whole classes of patients may receive a drug while large scale clinical trials are continuing; the details of that program are discussed below. In general, prior FDA approval is required to charge patients for drugs received in research protocols; in the case of drugs provided under the Treatment IND program, only notification of the FDA and nonobjection by the FDA is required. In either case, the price may not exceed the actual costs of manufacture, research and development, and handling of the investigational drug.

If the results of the research are satisfactory, the sponsors may file an NDA with FDA. This application must include information about the composition and production of the drug, about the proposed labeling for the drug, and about the research results which support the claim that the drug has a favorable benefit–risk ratio. FDA regulations provide an extensive description, to be analyzed below, of the research data that are required to support such a claim. Some of those requirements have been modified, in ways that will be described below, for the Accelerated Approval programs. FDA has 180 days to respond to a filing of an NDA, but that period is open to extension. In making a determination of the application's acceptability, FDA draws upon the expertise of advisory committees, but the final decision belongs to FDA. Congress has in

the last few years authorized FDA to charge user fees to those filing NDAs. The user fees have supported the hiring of additional staff, which has helped speed the NDA review process. Once the FDA responds with its approval, the new drug can be marketed in accordance with the approved labeling.

The set of regulations just described was developed for the approval of new drugs. It also applies to the approval of already approved drugs for new indications. The approval regulations are somewhat parallel to those governing the approval of new medical devices (7). Approval does not apply to new surgical procedures (as opposed to new devices that they may employ), to new testing services including genetic testing services (as opposed to new testing kits that they may employ), or to new physician uses of approved drugs for nonapproved purposes (as opposed to manufacturer promotion of such "off-label" new uses). The 1997 Food and Drug Administration Modernization Act partially addressed the "off-label" use issue by codifying the conditions under which manufacturers can distribute scientific information related to "off-label" uses (8). Much controversy exists both about the merits of these exceptions to the FDA's regulatory authority and about the new legislation related to "off-label" promotion.

FDA RULES FOR IND STAGE

Rules Governing Human Subjects Research

When a sponsor submits an IND application to the FDA, it must submit documentation that the proposed research protocols have been reviewed and approved by an appropriate IRB and that the protocols contain provisions for obtaining the informed consent of research subjects or their representatives. These requirements are consonant with the internationally accepted consensus about the conditions required for ethical research on human subjects.

All such research protocols must be reviewed and approved by an IRB that is independent of both the sponsors and the investigators. In approving the research, the IRB must determine that the following very standard requirements have been met (9):

- Risks to the subjects have been minimized and the remaining risks are reasonable in relation to the anticipated benefits.
- Selection of the subjects is equitable, and potentially vulnerable subjects are provided with additional safeguard.
- Informed consent has been obtained and documented.
- Privacy of subjects and confidentiality of data are protected.
- Adequate provisions are made for ongoing safety monitoring.

The requirements for the information that must be provided as part of the informed consent process are also standard. Subjects must be informed of at least the following (10):

- Purpose of the research, its duration, and the experimental procedures involved.
- Risks and benefits of participation and of the alternatives to participation.
- Extent to which research records are confidential.
- Any compensation and/or treatment for research related injuries.
- Right not to participate and right to discontinue participation.

Comparison to Federal Common Rule

In 1991 most federal agencies that conduct or sponsor research on human subjects adopted regulations governing human subjects research. Among these agencies was the Department of Health and Human Services (HHS) (11). The commonly adopted regulations are referred to as the Federal Common Rule. FDA did not adopt the Federal Common Rule, because of certain special features of FDA's regulatory process. The above-described FDA regulations closely resemble the Federal Common Rule. In 1996 FDA issued a description of the differences between its regulations and the Federal Common Rule (12). The three most important to be noted are the following:

- FDA regulations apply to all research submitted to it in IND applications and NDA submissions, regardless of the funding source for the research, while the Federal Common Rule only covers research funded and/or conducted by the relevant federal agency.
- FDA has developed special exemptions from the requirement of informed consent for emergency research; these will be described below.
- The Federal Common Rule has an elaborate system of reports by IRBs (called the assurance mechanism) to ensure that they meet the standards in the Federal Common Rule. No such system is imposed under the FDA regulations.

International Application of These Rules

As noted above, the FDA regulations apply to all research submitted to it, whether or not the research is conducted in the United States (13). Because of the significance of FDA approval in allowing access to the U.S. market, this requirement impacts greatly on the conduct of human subjects research in other countries. There are those who see this as a form of ethical imperialism, of the United States imposing its ethical standards on other countries. There are others who see this as a U.S. reaffirmation of fundamental moral truths, which remain true even in other countries that have not yet recognized their value. Both views assume that the standards are unique to the United States, and it is this assumption that is in error. It has recently been demonstrated that the essentials of the FDA standards have been broadly adopted throughout the world, even if the precise details vary from country to country, and that this broad adoption is strong evidence of the moral validity of FDA standards (14).

FDA regulations require that research conducted on human subjects in foreign countries meet the standards of

the Declaration of Helsinki, a declaration of the cross-national World Medical Association, rather than the standards of the FDA regulations. In fact the standards are very similar, reflecting the broadly accepted international consensus about human subjects research. Still, imposing the standards by reference to the Declaration of Helsinki clarifies that doing so is not a form of U.S. ethical imperialism. Moreover, if the country in which the research is conducted has its own more stringent standards on human subjects research, then FDA requires that those more stringent standards be met, showing respect for those stricter standards.

Exception for Emergency Research

The international consensus embodied in the FDA regulations requires that informed consent be obtained from all human subjects or from their representatives before the research is conducted. That requirement has often imposed great difficulties on those conducting vitally needed research on the management of medical emergencies. Such research often must be conducted as soon as possible after the patient-subject presents in the emergency room. The patient may be temporarily incompetent to give consent and no legally authorized individual may be present to give consent. Moreover the very short time frame for the effective use of the investigational intervention often precludes the possibility of accurately informing the competent subject and/or the available representative and giving them a meaningful choice about participation. So vitally needed research suffers from imposing that requirement of prospective informed consent. Even when there are enough subjects or representatives available and competent to give meaningful consent, so the research can go forward, that requirement has prevented many other noncompetent emergency patients from getting access to the most promising new interventions, and that seems inequitable. There are, then, important moral considerations favoring a waiver of the requirement of informed consent in these special circumstances (15).

There are, on the other hand, important moral considerations favoring retaining that requirement. To begin with, we should try to avoid violating the right of individuals not to be used as research subjects without their consent. Second, we need to protect these vulnerable subjects from being harmed by research to which they have not consented when (as often happens) the experimental interventions do not fulfill their promise.

Many proposals have been made as to how this value conflict should be resolved (16,17). One of the most promising is the regulations adopted by FDA in October 1996. FDA allows for an IRB to waive the requirement of prospective informed consent when (18):

- the subjects are in a life threatening condition and available treatments are unsatisfactory;
- animal and preclinical human studies support the likelihood of the intervention's being helpful;
- obtaining informed consent from the subjects or their surrogates is not feasible, so the clinical investigation cannot be carried out without the waiver; and
- additional procedural safeguards, as well as the usual IRB approval, are adopted, including attempting to find surrogates to obtain their consent and/or informing them and the subjects afterward, community consultation and notification, appointing an independent data safety monitoring board, and FDA approval.

If the values protected by the requirement of informed consent are absolute moral values, these regulations are inappropriate, since they allow for the waiver of the requirement of obtaining that consent. They are appropriate, however, if those values need to be balanced against the competing values of social need and potential subject benefit. As noted in the Introduction to this article, this type of balancing approach, which has recently won much favor among scholars, is characteristic of many of the FDA's newer rules.

Treatment INDs

In general, new drugs being tested under an IND may not be distributed for clinical use by patients not enrolled in research protocols. The one exception to this rule is the Treatment IND program which allows for the distribution outside of research protocols of some new drugs being tested for use in treating immediately life-threatening or serious illnesses (19).

The Treatment IND program, formally adopted by FDA in regulations in 1987, grew out of the experience with AZT for treating AIDS patients. In 1986, Phase II trials began for AZT. They were stopped after six months when it was clearly demonstrated that AZT was life-prolonging in the short run. An NDA was filed in December 1986 and approved in March 1987. From the time the drug was stopped until the application was finally approved, over 4000 patients received the drug outside any research protocol. This experience demonstrated the need for such a program in certain cases (20). From a moral perspective, the value of quicker access takes precedence in the relevant cases over the value of protecting desperate patients from not-yet-approved drugs. This is one more example of the need to balance values, rather than treating any one value as absolute, in research ethics.

Although the program was first introduced in response to the need for an AIDS drug not yet approved, it has been used to allow access to drugs for treating other conditions. In the first six years of the program, 28 Treatment INDs were issued. Nine were for AIDS-related drugs, 9 were for cancer-related drugs, and 10 were for drugs to treat other conditions (e.g., newborn respiratory distress syndrome and neurological conditions such as Alzheimer's dementia, MS, and Parkinson's disease) (21). Some Treatment INDs were issued after the clinical trials were completed but before the NDA was approved, while other INDs were issued while clinical trials were still being conducted.

Under the regulations the following are the requirements for issuing a Treatment IND (22):

- The drug is to be used for treating an immediately life-threatening condition (death will occur in a matter of months, or premature death will result from nontreatment) or a serious disease.

- There are no comparable or satisfactory treatments available.
- The drug is being investigated in controlled trials, or the trials have been completed and the sponsor is actively pursuing approval.
- For the drug to treat immediately life-threatening conditions, there must be a reasonable basis for believing that the drug may be effective and would not expose the patient to unreasonable additional risks. For the drug to treat a serious illness, there must be sufficient evidence of safety and effectiveness.

These provisions were adopted by Congress in the Food and Drug Administration Modernization Act of 1997 (23).

Special Rules for Genetic Research

Investigational biologic agents used for somatic cell and gene therapy must undergo the same process of FDA approval as other investigational agents (24). INDs must be secured for the research to begin, and NDAs must be approved before the agent can be distributed for general use. Until recently IRB and FDA approval of gene therapy protocols was not sufficient; a special NIH committee, the Recombinant DNA Advisory Committee (RAC), had to approve all federally funded research protocols. That requirement of additional approval has recently been eliminated, although RAC continues to review individual protocols, and it has been suggested that its authority to approve protocols be reinstituted.

Biologic agents pose special problems of product development which have received special attention by the FDA. Measures must be taken to control the biologic sources of the material, the production process, and the final product. These are necessary to deal with concerns both of safety and potency. The FDA issued in 1991 a *Points to Consider* document to clarify many of these technical issues (25).

In contrast to gene therapy, new forms of gene testing have not traditionally required FDA approval unless they use new testing kits; if they do, the kits, as opposed to the tests themselves, do require FDA approval. In recent years concern has been expressed about the proliferation of new genetic tests not employing new kits and not therefore being approved by any regulatory agency. Many feel that this allows for the marketing of tests that may not be analytically or clinically valid and useful, and that this can lead to a variety of problems for those being tested. Others have less concern, thinking that the decisions to use these tests should be left to the individuals in question, guided by their physicians or other health care providers. A recent report from a joint task force of the National Institutes of Health and the Department of Energy (26) supports the former approach, arguing that there is a need for review at a national level of new genetic tests before they enter into clinical practice and suggesting that the FDA does have the regulatory authority to conduct that review. It also points that if the FDA were to adopt this role, it would need to look at a broader set of issues, including the contribution of testing to producing better long-term outcomes, than the FDA usually does when reviewing lab tests. It remains to be seen whether this approach will be adopted.

FDA RULES FOR NDA STAGE

Traditional Rules for Adequate Evidence

The research conducted under an IND is designed to acquire data sufficient to support the approval of an NDA. In order to do that, the data must establish with a high enough degree of certainty that the drug has a sufficiently favorable risk–benefit ratio to justify its intended use. This section will review the FDA regulations related to the NDA approval process.

It is helpful in reviewing FDA regulations to understand that any drug approval process must answer two different, although related, questions, and that answers to these questions necessarily involve trade-offs among important values. The first of these questions is a *content* question: How should the values of effectiveness and safety be balanced in deciding whether the drug in question has a sufficiently favorable risk–benefit ratio? We want drugs that are effective. We also want drugs that are safe. But the most effective drugs often carry with them risks. The content question asks about how these two legitimate values should be balanced in deciding what is a sufficiently favorable risk–benefit ratio. The second of these questions is an *epistemic* question: How should the demand for adequate evidence and the demand for speedy approval be balanced in deciding that the evidence in question has established the drug's risk–benefit ratio? We want firm evidence that the risk–benefit ratio is sufficiently favorable. We want good drugs to be approved as soon as possible. But getting firm enough evidence often requires delays in the approval process. The epistemic question asks about how these two legitimate values should be balanced in deciding whether the evidence is firm enough so that the drug should be approved without further delay. The FDA regulations have much more to say about the epistemic question than the content question.

There is a special section of the NDA regulations that address the type of evidence required (27). The basic point of the section is that the FDA considers adequate and well controlled studies as the primary basis for supporting the claim that there is sufficient evidence to justify the approval of a NDA. The following characteristics of such studies are listed as essential:

- The study must involve a control group to compare with the treatment group in order for the drug's effect to be assessed quantitatively. Five types of control groups are identified. Four are concurrent control groups (placebo concurrent, dose-comparison concurrent, no treatment concurrent, and active treatment concurrent). For these types of control groups, subjects should be randomly assigned to the treatment group or the control group. The fifth type is a historical control group; in that type of study, randomization is not possible.
- Patients must be assigned to the treatment and control groups by a method that minimizes bias.

Randomization is the way to accomplish this in all concurrently controlled trials.

- Procedures such as blinding subjects and investigators must be adopted to minimize bias in the conduct or interpretation of the study.
- The protocol for the study must carefully define the population to be studied and the methods to ensure that the subjects are part of that population, the nature and duration of the treatment to be studied, the number of subjects to be studied, the methods to assess the response of the subjects, and the planned statistical analysis.

Based on its reading of the 1962 statute, FDA has traditionally insisted that two adequate and controlled studies provide sufficient evidence to justify the approval of an NDA. That is, of course, much more that just a point of statutory interpretation. It reflects the well-known scientific ideal that good scientific results are reproducible results. Nevertheless, such a requirement can often result in approval delays if the two trials are run sequentially, and additional costs even when the two are run concurrently. As part of its effort to develop a new balancing of values in response to the epistemic question, Congress, in the Food and Drug Administration Modernization Act of 1997, rejected that interpretation and left the need for a second trial to the discretion of the agency (28).

There are many important issues raised by this traditional set of rules. Probably the most controversial are those surrounding the choice of the control group, particularly when there already are available other drugs for treating the condition in question. From the perspective of easily getting well-established answers (the perspective of investigators, sponsors, and future patients), placebo or no treatment concurrently controlled trials are preferable because they allow for smaller and more easily interpretable studies. From the perspective of insuring access to at least some treatment (the perspective of the subjects and of their treating clinicians), dose-comparison or active treatment concurrently controlled trials are preferable. Both sets of values are legitimate, and the issue becomes one of how to balance the values. Several crucial official statements seem to support the use of only the latter types of control groups whenever an effective alternative treatment exists (29,30). This approach seems too absolute in stressing only the subject-centered values. FDA, by contrast, has stressed the scientific advantages of the first types of control groups, and has urged their use (with appropriate subject-protection mechanisms such as early rescue and minimization of study duration) except when existing treatments are life-prolonging (12). This approach seems to not sufficiently weigh the subject-centered values; it does not, for example, take into account the losses to subjects in the control group when existing therapies effectively limit morbidities and discomfort. What is needed in the case of each controlled trial is a careful look at all possible trial designs (different types of control groups and different types of protection mechanisms) and a determination of which best balances the research-centered values with the subject-centered values (31).

Accelerated Approval Rules

The accelerated approval rules, like the Treatment IND rules discussed above, arose in response to those suffering from AIDS demanding quicker access to promising drugs. They were joined in this demand by others, such as patients with cancer, who were dissatisfied with the available treatments and who wanted quicker access to promising drugs. Unlike the Treatment IND rules, the accelerated approval rules relate to final approval of the use of the drug, and not just to interim access while approval is being considered.

The accelerated approval rules apply to drugs being tested for the treatment of serious or life-threatening illnesses (32). The drugs in question must have the potential of providing meaningful therapeutic benefits to patients over existing treatments, either because patient response is improved with the new drugs or because patients are unresponsive to, or intolerant of, the existing treatments. The crucial provision of the rules is that such drugs can be approved on the basis of well-controlled trials that establish a favorable effect on surrogate endpoints (e.g., tumor shrinkage for cancer patients or reduced viral loads for AIDS patients), endpoints that are thought to be predictive of true clinical benefits (e.g., improved length of survival or decreased morbidity). This provision accelerates the approval because it is often possible to get data on surrogate endpoints quicker than data on true clinical benefits.

Critics have pointed to many examples where the use of surrogate endpoints has led to mistaken conclusions about the effectiveness of the new drugs as measured in terms of the true clinical benefits (33). They are certainly right to raise these concerns. Even after accelerated approval is given on the basis of data concerning surrogate endpoints, postmarketing studies of the effectiveness of the drug on true clinical endpoints need to be conducted when there is doubt about the effect on the true clinical endpoints. But this practice does not undercut the moral validity of the accelerated approval rules. There is rather a need to balance the demand for adequate evidence with the demand for speedier approval. The accelerated approval rules do so for drugs whose promise is supported by surrogate endpoint data but not fully established by true clinical endpoint data in cases where the patients are very sick and existing treatments are not helpful. The actual choice to use the new drugs will be made by patients with the advice of their doctors, but society will not stand in their way once promising surrogate endpoint evidence becomes available. When these provisions are coupled with the requirement of further studies as appropriate, they seem like a reasonable balancing of the competing values. The Food and Drug Administration Modernization Act of 1997 accepted this conclusion and incorporated the accelerated approval rules into the statute governing the FDA (34).

Rules Governing Special Populations

There are a number of populations that traditionally have been perceived as requiring special protection in the research setting. One such group is the elderly, who often

live in conditions of dependency, a dependency that makes them vulnerable to being exploited, and who sometimes suffer from impaired intellectual functioning that may impair their capacity to protect themselves. A second group is children, whose immaturity often impairs their capacity to protect themselves and whose dependency on their parents makes it difficult for them to make their own independent decisions about participation in research projects. A third group is women of childbearing potential, whose potential fetuses may require special protection. One way to provide special protection for these subjects is to exclude them from research, and this ethical concern is one of several factors that have often led in the past to the exclusion of these groups from research projects.

While this value of protection of the vulnerable is deserving of respect, there are other values that need to be considered as well, especially the value of justice. Individual members of excluded groups may unfairly be denied the personal benefit of obtaining access to promising new treatments. The exclusion of entire groups may result in the nongeneralizability of the results of the research to the members of the group, so they and their treating physicians are unfairly denied the basis for making decisions as to whether to use new treatments whose use in the general population is supported by the research in question (35).

In the earlier discussions of research ethics, greater emphasis seems to have been placed on the value of protecting the vulnerable from the risks of research. In recent years, as the benefits from participating in research have become better understood, more emphasis is being placed on balancing that concern with the justice-based concern of not excluding these individuals and groups from obtaining the benefits of research. As part of that more recent understanding, FDA has made several major initiatives in its rules dealing with these populations.

In 1989, in response to concerns about the exclusion of the elderly from research, FDA issued guidelines that, while explicitly not introducing further requirements for drug approval, provided recommendations about the inclusion of the elderly in research on drugs that are likely to have significant use in the elderly (36). The crucial recommendation was that the elderly should be included in such trials in reasonable numbers so that the patients enrolled in the trial reflect the patients who will use the drug if it is approved on the basis of the research. Following this recommendation would result only in the detection of fairly large age-related differences, but it was thought that those are the only differences important enough to be of concern.

These 1989 guidelines concerning the elderly were part of a larger effort in 1988 and 1989 by FDA to ensure that the subjects in research protocols were more representative of the population that would use the drug if it were approved (37). This more general effort also attempted to address the issue of the inclusion of women of childbearing potential in clinical trials. There is some controversy as to whether this effort was successful. The Government Accounting Office (GAO) claimed that there was a continuing problem, especially in connection with cardiovascular trials, while the FDA argued that the remaining differences resulted from the need to include more younger men because they were more vulnerable than younger women to cardiovascular disease. Nevertheless, FDA issued further clarifying guidelines in 1993 (38):

- Enough members of both genders should be included in order to detect "clinically significant gender-related differences," and the total study population should "reflect the population that will receive the drug when it is marketed."
- The integrated analysis of data from the trials should contain an analysis of gender differences in terms of both safety and effectiveness; additional studies might be required if these differences are significant enough.
- The definitive trials to support approval should employ gender-differing dosages if pilot pharmacokinetic studies reveal gender differences.
- Research protocols, unless specifically designed to study the effects of a drug during pregnancy, should include measures to minimize the exposure of fetuses to the drug being tested; the most important of these measures are pregnancy testing before administering the drug being tested and proper counseling about the selection and use of reliable methods of contraception.

In 1997 FDA proposed still further regulations to ensure that women of childbearing potential were included in clinical trials (39). These proposed regulations would only apply to trials of drugs intended for treating life-threatening illnesses. Unlike the above-discussed regulations, which have a direct effect only on the full set of data needed to secure approval and do not affect specific clinical trials, these proposed regulations will put constraints on certain trials. Under the proposed regulations, FDA can put a hold on trials, preventing further enrollment, if women of childbearing potential are being excluded because of fears of reproductive or developmental toxicities. FDA's argument for these proposed regulations is that such exclusion unfairly denies these women the choice to participate because of risks that can be minimized by the measures to minimize fetal exposure described above. These proposed regulations represent one further step in FDA's balancing of values in this complex area.

A third population for which FDA has developed special rules is the pediatric population. A very high percentage of the drugs used in the treatment of children have not been adequately tested for use in that population. In 1992, for example, 79 percent of the new drugs approved by FDA that could be used in children were approved without labeling for use in children because they had not been sufficiently studied in pediatric subjects (40). Based on 1994 data, the FDA identified the 10 most commonly used drugs in the pediatric population for which there was no pediatric labeling. The list included: albuterol inhalation for asthma, ampicillin injections for infection, and Prozac for the treatment of depression (41). These are serious drugs for serious medical problems, and their

pediatric use should be guided by better data. The lack of research data results in a fundamental ethical dilemma for pediatricians: Should they take the risk of using these new drugs without adequate information about their safety and efficacy at various dosages in children, or should they not use these drugs and risk denying to their patients valuable treatments? Clearly, policies that promote more testing of new drugs in children would help resolve this dilemma, and would constitute one more balancing of protecting subjects from research risks with insuring the access of patients to the benefits of research.

In 1994 FDA attempted to address this problem by adopting what can be called "the extrapolation" approach (42). According to that approach, FDA would allow labeling for pediatric use on the basis of clinical trials conducted with adults if there was sufficient pharmacokinetic and adverse reaction data from pediatric subjects to justify extrapolating the adult results of a favorable risk–benefit ratio to the pediatric population. That rule did not, however, require sponsors to conduct the studies required to justify the extrapolation. By 1997 FDA concluded that this approach had been inadequate to resolve the problem. It proposed supplementing that rule with an "including children" approach, similar to (although not exactly the same as) the approach it had adopted toward the elderly and toward women of childbearing potential. According to this proposed approach, the NDA application for drugs that represent a meaningful therapeutic advance and that are likely to be used in a substantial number of pediatric patients must include (unless a waiver is issued) data about safety, effectiveness, dosages, and mode of administration in the pediatric population. This proposal drew much criticism from those who were concerned with the cost of new drug development and from those who were concerned with the pace at which new drugs are approved for use in the adult population, but a version of it was adopted in 1998 (43).

CONCLUSION

As one reviews FDA's rules, two major themes clearly emerge. The first is the need to recognize that participation in research can be a benefit as well as a burden. It is this general recognition that has been central to the FDA's efforts to expand the types of subjects enrolled in clinical trials to include the elderly, children, and women of childbearing potential. The second is the need to balance values rather than to treat some as absolute. It is this general recognition that has led to FDA's rules on emergency research (which balances the need for informed consent with the individual and social need for emergency research), on Treatment INDs (which balances the need to complete clinical trials with the need for speedier access), and on accelerated approval (which balances the value of optimal evidence with the need for speedier access). These recognitions of the legitimacy of multiple values and of the need to balance them to arrive at optimal rules are a strength of FDA's rules governing human subjects research.

BIBLIOGRAPHY

1. H.G. Grabowski and J.M. Vernon, *The Regulation of Pharmaceuticals*, American Enterprise Institute, Washington, DC, 1983.
2. P. Temin, *Taking Your Medicine*, Harvard University Press, Cambridge, MA, 1980.
3. B. Brody, *Ethical Issues in Drug Testing Approval and Pricing*, Oxford University Press, New York, 1995.
4. 50 Fed. Reg., 7452–7519 (February 22, 1985).
5. 52 Fed. Reg., 8798–8847 (March 19, 1987).
6. Report 105-399 (November 9, 1997).
7. R. Merrill, "Regulation of Drugs and Devices: An Evolution" *Health Affairs* 47-69 (Summer, 1984).
8. PL105-115 Section 401.
9. 21 *Code Fed. Regulations*, 56.111.
10. 21 *Code Fed. Regulations*, 50.25.
11. 45 *Code Fed. Regulations*, 46.
12. FDA, *IRB Operations and Clinical Investigation Requirements*, Food and Drug Administration, Washington, DC, 1996, App. E.
13. 21 *Code Fed. Regulations*, 312.120.
14. B. Brody, *The Ethics of Biomedical Research*, Oxford University Press, New York, 1998.
15. B. Brody, *Hastings Center Rep.* **27**(1), 7–9 (1997).
16. N. Abramson and P. Safar, *An. Emergency Med.* **19**, 781–784 (1990).
17. P. Grim et al., *J. Am. Med. Assoc.* **262**, 252–255 (1989).
18. 21 *Code Fed. Regulations*, 50.24.
19. 21 *Code Fed. Regulations*, 312.34.
20. F. Young et al., *J. Am. Med. Assoc.* **259**, 2267–2270 (1988).
21. FDA, *Treatment Investigational New Drugs Allowed to Proceed*, Food and Drug Administration, Washington, DC, 1993.
22. 21 *Code Fed. Regulations*, 312.34.
23. 21 U.S.C., §360bbb.
24. M. Couts, *Kennedy Inst. Ethics J.* **4**, 63–83 (1994).
25. FDA, *Hum. Gene Ther.* **2**, 251–256 (1991).
26. Task Force on Genetic Testing, *Promoting Safe and Effective Genetic Testing in the United States*, NIH ELSI Program, Washington, DC, 1997.
27. 21 *Code Fed. Regulations*, 314.126.
28. PL105-115 Section 115.
29. World Medical Association, *Declaration of Helsinki*, Principle II.3.
30. *NCBHR Communique*, No. 7, pp. 20–21 (1996).
31. AMA, *CEJA Rep.* 2-A-96 (1996).
32. 21 *Code Fed. Regulations*, 314.150.
33. T. Fleming and D. DeMets, *An. Internal Medi.* **125**, 605–613 (1996).
34. 21 U.S.C., §356.
35. B. Brody, in J. Kahn, A. Mastroianni, and J. Sugarman, eds., *Beyond Consent*, Oxford University Press, New York, 1998, pp. 32–46.
36. FDA, *Guideline for the Study of Drugs Likely to Be Used in the Elderly*, Food and Drug Administration, Washington, DC, 1989.
37. FDA, *Guideline for the Format and Content of the Clinical and Statistical Sections of New Drug Applications*, Food and Drug Administration, Washington, DC, 1988.

38. FDA, *Fed. Reg.* (July 22, 1993), 39406–39416.
39. FDA, *Fed. Reg.* (September 24, 1997), 49946–49954.
40. Committee on Drugs, *Pediatrics* **95**, 286–294 (1995).
41. FDA, *Fed. Reg.* (August 15, 1997), 43900.
42. FDA, *Fed. Reg.* (December 13, 1994), 64240.
43. 21 *Code Fed. Regulations*, 201.23.

See other entries FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE; GENE THERAPY, LAW AND FDA ROLE IN REGULATION; see also HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, LAW, HHS RULES

CHARLES R. MCCARTHY
Kennedy Institute of Ethics
Georgetown University
Washington, District of Columbia

OUTLINE

INTRODUCTION

The U.S. Department of Health and Human Services (HHS) is the single largest supporter of biomedical research in the world. HHS conducts or supports more biomedical and behavioral research involving human subjects than all other federal agencies combined (1). It is not surprising, therefore, that HHS is charged by law to issue regulations for the protection of human research subjects. HHS has played and continues to play the leading role among U.S. departments and agencies in promulgating and implementing policies and protections for the rights and the welfare of human research subjects.

This article will present a brief history of the origins, development and astonishing growth of the HHS agencies, particularly that of the National Institutes of Health (NIH), and its laws, policies, programs, and regulations for the protection of human subjects involved in research.

ORIGINS OF FEDERAL SUPPORT FOR BIOMEDICAL RESEARCH

Marine Hospital Service

In 1798 the U.S. Congress authorized, and President John Adams signed into law, a bill establishing the federal Marine Hospital Service to care for sick and disabled seamen. No one living at that time could have foreseen that the Marine Hospital Service, a Division of the Department of the Treasury, would evolve into a complex network of federal departments and agencies that collectively support the world's largest health research enterprise.

Federal support for biomedical research did not begin for nearly a century after the health care service program for merchant seamen was initiated. The creation, in 1887, of the one-room Bacteriological Laboratory for investigation of cholera and other infectious diseases at the Marine Hospital on Staten Island, New York, initiated federal support of biomedical research dedicated to public health (2).

U.S. Public Health Service

The U.S. Congress gradually widened the responsibilities of the Marine Hospital Service, and assigned additional research responsibilities to its laboratory, designated as the Marine Hospital Service Hygienic Laboratory in 1891.

In 1902 the Marine Hospital Service was renamed the Public Health and Marine Hospital Service (MHS), and in 1912 the mission of the MHS was further broadened; the Service was reorganized and it was renamed the U.S. Public Health Service (PHS). Over the next decade, a small military unit called the Public Health Service Commissioned Corps was added to the PHS. The Corps was composed of trained medical and research personnel who dedicated themselves to the protection and promotion of the health of the U.S. population. PHS Commissioned Corps officers were placed under the command of the Surgeon General of the United States (SG). Members of the Corps were expected to accept assignment to areas of the country that were medically underserved, or assignment to areas of the country where an outbreak of disease created special demands for medical personnel. The officers of the Commissioned Corps frequently labored alongside their counterparts, civilian employees of the PHS agencies. From the earliest days of the Corps, some of its officers have been assigned to conduct or administer biomedical research programs.

The Chamberlain-Kahn Act of 1918 (40 Stat.L. 309) directed the SG to conduct a new research initiative into the causes, prevention, and cure of venereal diseases, and empowered the SG to initiate grants-in-aid to further such research. Accordingly in 1918, the SG authorized grants-in-aid-of-research to twenty five extramural (outside the

PHS) institutions. These grants were the first biomedical research awards involving human subjects made by the federal government to research institutions in the private sector. They mark the beginning of an unprecedented partnership involving academe, industry, and the general public (represented by the federal government).

Creation of Categorical Research Institutes

The Ransdell Act of 1930 (P.L. 71-251) reorganized, expanded and renamed the Hygienic Laboratory as the National Institute of Health. Over time the Ransdell Act served as a template for what eventually would become a cascade of legislation that created additional research institutes aimed at studying disease categories, or studying human organs that are subject to specific categories of disease. The National Institute of Health gradually became a federation of categorical research agencies called the National Institutes of Health (NIH).

The first of the new of categorical institutes (later called the National Institute of Mental Health) was authorized by Congress to carry out research into the causes, treatment, and prevention of mental and nervous disorders, and to conduct research into narcotics abuse (P.L. 71-357).

The Social Security Act of 1935 (P.L. 74-271) is often considered to be the high water mark of the pre–World War II Roosevelt administration because it entitles persons who qualify — for reasons of age or disability — to receive income from the federal government. However, few people recall that the Social Security Act also contained landmark provisions that authorized the PHS to advise and assist state and local health officials to offer services and conduct research to prevent the spread of disease within and across state lines, and to improve regional and local health programs. This responsibility has, for the most part, been assigned to the National Centers for Disease Control (CDC) in Atlanta, Georgia.

The Congress created the National Cancer Institute (NCI) in 1937 (P.L. 75-244). The new institute not only provided for grants-in-aid to study the causes, treatment, and prevention of the many varieties of cancer, but it provided for fellowships, personnel training, and cancer prevention and control programs within and across the several states. The legislation that established NCI served as a model for the subsequent establishment by statute of 24 additional institutes and centers that currently constitute NIH.

In 1939 PHS was transferred from the Treasury Department to the Federal Security Agency (FSA) (P.L. 76-19). In 1953 FSA was folded into the newly created Department of Health Education and Welfare (HEW) headed by a Secretary who is a senior cabinet officer. When the Department of Education was separated from HEW and established as a new Department in 1979, HEW was transformed into the Department of Health and Human Services (HHS) (3).

World War II: A Time of Change

The Second World War brought profound changes to health research in the United States. In 1941 President Franklin D. Roosevelt established a federal Committee for Medical Research (CMR) to address war-related disease and injury. The CMR carried out its responsibilities by issuing multiple contracts for targeted research. Spectacular results including, development of sulfanilimides, gamma globulin, adrenal steroids, cortisone, and a wide variety of new surgical techniques and treatment regimens, were credited to the CMR (4).

Enactment of the PHS Act with Section 301 Authority

The enthusiasm generated by successful wartime research efforts led by the CMR was channeled, following World War II, into unprecedented congressional support for biomedical research. Before World War II had come to a close, Congress responded to the success of wartime research by enacting the landmark PHS Act of 1944 (P.L. 78-410) that revised and consolidated into a single law, all of the authorities that governed the various PHS agencies. The PHS Act required major reorganization of PHS agencies. Most important, it gave the PHS unprecedented generic powers to support research "into [all of] the diseases and disabilities of man" (commonly known as "Section 301 authority"); and authorized the involvement of subjects in research conducted at PHS medical facilities. The significance of "301 authority" lay in the fact that the PHS agencies, especially NIH, were authorized to exploit promising research opportunities in basic research and in a wide variety of medical specialty fields without waiting for Congress to authorize such research. Thus, although the institutes of NIH were created with categorical missions, they were able to devote a large portion of their growing research budgets to basic research. The result was a spectacular expansion of the biological knowledge base. Similarly the institutes of the NIH, using "301 authority," were able to follow promising research leads without waiting for Congress to authorize and fund their efforts.

The NIH Clinical Center was erected in 1953 under Section 301 authority. Today Section 301 authority is seldom cited because Congress has created so many research programs targeted at specific diseases and disabilities that there is less need to invoke generic research authority. Furthermore Congress now recognizes that giving direction to specific research programs into diseases such as AIDS, heart disease, cancer, and diabetes is politically more attractive to the tax-paying public than giving nonspecific authority to support expansion of the biomedical research knowledge base. Much basic biological research is still supported by NIH, but it is called "cancer research" or "heart research" or "diabetes research," even though it focuses on basic biological structures and functions that may find application in many disease categories. Since the enactment of the PHS Act, virtually all federal legislated research initiatives have taken the form of amendments to the PHS Act that authorize or underwrite both intramural and extramural research programs.

Absence of Research Ethics Policies in Pre-war Years

The research agencies and programs of the U.S. government were created and repeatedly reshaped by the Congress during the first half of the twentieth century. It is

interesting to note, however, that during that same period of time, the Congress wrote no laws, held no hearings, and created no policies for the protection of human subjects involved in research.

By failing to provide either substantive or procedural ethical rules for research funded by PHS, the Congress tacitly implied that responsibility for the ethical conduct of research should be left almost entirely to awardee institutions and to the conscience of each investigator who conducted research supported by the federal government (5).

The absence of ethics policy governing research involving human subjects conducted or supported by the federal government can be explained by several factors. In the first place, the Hippocratic tradition of medicine was the centerpiece of the prevailing ethics of medicine and research conveyed to physicians in training, including those who would function as clinical research investigators. That tradition was imparted by mentors, role models, and institutional traditions. Bioethics was not recognized as a distinct discipline that lent itself to a methodological approach and systematic teaching. Biomedical ethics, including research ethics, was not taught, nor was it even available in the library collections of most of the medical schools of that period.

Many medical ethical problems were addressed by moral theologians teaching in educational institutions established and operated by religious denominations. These theologians systematically addressed fundamental questions of medical ethics, but because theological deliberations were grounded in religious faith, traditions and practice — as well as human reason — neither the Congress nor the Executive Branch demonstrated a strong interest in incorporating the opinions and conclusions of moral theologians into public policy. Furthermore, although a few theologians enjoyed broad knowledge of medical practice, few had credentials in research. Most theologians lacked credibility with the research community (6).

Second, the distinction between medical care and biomedical research was seldom made — by the Congress, by the public, by subjects of research, or even by research investigators themselves. Consequently the ethics of medical research (dedicated to systematic development of generalizable medical knowledge) was not clearly distinguished from the ethics of medical practice (dedicated to the best interests of each individual patient). Research subjects often assumed that research investigators were serving their best interests, whereas in fact the primary concern of research investigators was to gain generalizable knowledge for the sake of the public health. (This error is sometimes referred to as the "therapeutic mistake.") In some cases the best interests of subjects coincided with the development of new knowledge in research, but in most cases the interests of the subjects were subjugated to the development of general knowledge. Today the distinction between medical practice and medical research is widely recognized. Prior to World War II and for many years thereafter, it was seldom cited or acknowledged, even by the strongest proponents of research.

Third, prior to World War II, research subjects had often been patients of physicians whose medical practice evolved beyond standard care of patients into innovative therapy, and finally into carefully constructed research projects. Neither medical practitioners who evolved into research investigators, nor their patients who evolved into research subjects, seemed to recognize the fundamental change in their relationship. Rothman notes that research subjects of this period had a high level of trust in research investigators because they usually thought of the investigators as their private physicians. Public confidence in the ethical integrity of physicians who provided care to patients was at an all time high. That confidence was easily extended to physicians who crossed the threshold from practice to research (5).

As the funding of research by the federal government expanded, research became a full-time career for many investigators. Because their research cohorts were made up mostly of subjects referred to them by physicians, it gradually became less common for research investigators to have prior trust relationships with their research subjects.

Postwar Development

After the close of World War II in 1945, a series of laws created the National Institute of Mental Health (P.L. 79-487), the National Heart Institute (P.L. 80-655), and the National Dental Institute (P.L. 80-755). The Hill-Burton Act of 1946 (P.L. 79-725) authorized grants to states for the construction of hospitals and public health centers, each of which included a research component. The Omnibus Medical Research Act of 1950 (P.L. 81-692) created the National Institute of Neurological Diseases and Blindness and the National Institute of Allergy and Infectious Diseases.

As research funding components of NIH multiplied following World War II, the biomedical research budgets of PHS agencies, particularly that of NIH, experienced meteoric rises that reflected unprecedented enthusiasm for the support of research by the tax-paying public. Between 1946 and 1949 NIH budgets leaped from $180 thousand to more than $800 million (7).

In 1953 two events of great importance to medical research in general and to research involving human subjects in particular occurred. First, PHS, including SG and PHS Commissioned Corps, was transferred to the new HEW. Second, the NIH Clinical Center, a state-of-the-art research hospital, opened its doors. The Korean conflict was in progress at the time, and young American men (including physicians and scientists) were subject to military draft. Physicians who qualified as research fellows at the CC at that time were given credit for military service. Consequently keen competition for NIH Fellowships developed. Fellows from the NIH CC program subsequently assumed leadership roles in the research programs of medical schools and in private industry. As a consequence of NIH's "doctor draft," biomedical research in America not only expanded, but the quality of such research improved dramatically.

It is not surprising, therefore, that HHS (and its predecessor, HEW) played and continue to play a leading

role in the United States in providing protections for the rights and the welfare of human research subjects. That leadership was scarcely discernable after World War II, but since 1966 it has gradually expanded and has become increasingly prominent.

The HHS includes many agencies. The largest — measured in terms of disbursement of funds — are the Social Security Administration (SSA) and the Health Care Financing Administration (HCFA) that administers the Medicare program. Biomedical research supported by HHS is, with a few exceptions, conducted or supported by eight agencies that comprise PHS (9). (The eight agencies today include the Agency for Health Care Policy and Research, the Health Resources and Services Administration, NIH, the Indian Health Service, CDC, FDA the Substance Abuse and Mental Health Services Administration, and the Agency for Toxic Substances and Disease Registry.) The PHS agencies were, in the past, coordinated by the PHS administration, operating under the direction of the Assistant Secretary for Health (ASH) and SG (sometimes these positions have been held by a single person who was at both ASH and SG, at other times they were held by different persons). SG and ASH reported to the Secretary of HHS.

Beginning in 1998, the administrative Offices of ASH and SG have been absorbed into the offices of the Secretary of HHS. PHS agencies are expected to continue to report to the Secretary of HHS through ASH and SG, but an intermediate level of administration will largely disappear.

NIH is by far the largest biomedical research agency within HHS. NIH is currently composed of 25 institutes and centers. The fiscal year (FY) 1999 budget of NIH is $15.56 billion dollars (8).

In FY 1998 the CRISP data collection program showed that 27,782 research projects involving human research subjects were conducted or supported by HHS. The vast majority of these projects are supported by NIH. Because a study involving human subjects may involve as few as one or two subjects, and as many as 18,000 or more (e.g., the ongoing study on prevention of prostate cancer), it is nearly impossible to estimate how many subjects are involved in research conducted or supported by HHS.

First U.S. Policy for the Protection of Human Subjects

The most visible effect of the newfound public endorsement of research can be seen in completion of the NIH state-of-the-art Clinical Center (CC) in 1953. The CC is a 500-bed hospital dedicated exclusively to the conduct of research. Only subjects actively participating in research studies are eligible for admission to the Center. Even in the setting of an institution totally dedicated to research, it was tacitly assumed that research procedures involving Clinical Center "patients" should be conducted in accord with the traditional doctor-patient relationship — a relationship of trust in which the patient had reason to believe that the doctor was uniformly acting in the best interests of the patient. To require a formal process of informing subjects of the risks and benefits involved in research and eliciting informed consent was judged to be intrusive in the doctor-patient relationship. It was left to each research investigator to decide what information, if any, would be conveyed to "patients," whether a consent process was to be employed, and how the process would be recorded.

However, the decision not to intrude on the doctor-patient relationship that was mistakenly thought to be the same as the physician-patient relationship could not be applied to "normal volunteers." Normal volunteers, recruited from colleges in nearby states, served to provide "control data" that was compared with data derived from "patients" in clinical studies. Since normal volunteers were expected to take medical risks and to receive no direct benefit from their participation in research, they could not be identified as "patients." A special policy was developed by the CC for normal volunteers (9). The policy stated: (*1*) that informed consent would be obtained from normal volunteers prior to their participation in any research project, and (*2*) that a research project involving normal volunteers could go forward only if it had been approved by a committee of scientists. Once a committee of scientists had approved a research protocol involving normal volunteers, and the normal volunteers had consented to participate, responsibility for the normal volunteers' safety and health was assigned to the principal investigator conducting the research. Dr. Philippe V. Cardon, M.D. was assigned to supervise the program for normal volunteers. The program would, in subsequent years, provide the first template for the PHS policy that was to come.

A Policy Vacuum

During the 1950s the United States was simultaneously recovering from the human and materiel expenditures of World War II and focusing its attention on the increasingly complex nature of the cold war. The United States was also enmeshed in the Korean conflict. American citizens were dissatisfied with the ambiguous stalemate that ended military conflict in Korea but produced no lasting peace. The U.S. economy was expanding, and America turned its attention away from military conflict to fighting battles against disease. Little attention was given to the ethics of research. Within NIH attention focused more on the development of a fair and just peer review process (employed to decide which research projects were worthy of public funding) than on the rights and welfare of human research subjects.

In 1959 Senator Estes Kefauver (D.TN) began a series of hearings focused on the impact of the rising cost of prescription drugs on the health care costs of patients. In the course of the Kefauver hearings, the media revealed that hundreds of infants born to mothers who had taken the drug thalidomide were grotesquely deformed. By this time many American homes had television sets that brought graphic pictures of the thalidomide tragedy into their living rooms. Most of the babies affected by thalidomide were born in the United Kingdom, in Europe, or in Canada. Tragedy had been averted in the United States only because one FDA employee, Dr. Frances Kelsey, had demanded more animal tests before she would agree to its being prescribed for use by pregnant women in

the United States. Nevertheless, fear of a similar tragedy in this country preoccupied much of the nation.

Senator Kefauver quickly changed the focus of his hearings from the economics of drug testing to the safety of drug testing. He unveiled for the public the fact that drugs were commonly tested on patients without their knowledge. Physicians simply administered experimental pharmaceuticals without informing their patients that the drugs had not been tested for safety or efficacy. The Food and Drug Administration (FDA) had little authority to prevent the widespread practice, though it encouraged the presentation of animal data prior to licensing drugs for market use.

In 1962 Senator Kefauver, despite opposition from the pharmaceutical industry, was successful in persuading the Congress to enact the Kefauver-Harris amendments to the Food, Drug and Cosmetic Act (P.L. 87-781). The Act, as amended, required rigorous testing of drugs for safety and efficacy before they could be marketed. When the bill came to the floor of the Senate, Senators Jacob Javits and John Carrol introduced an amendment that required that subjects provide informed consent before being involved in the testing of investigational drugs.

The thalidomide episode seemed to prompt disclosure of other research problems. *The New York Times* carried a story of research involving the injection of live cancer cells into elderly, indigent, and possibly incompetent patients without their consent. These "charity" patients already were suffering from various forms of cancer, and the investigators wished to know whether their immune response systems would reject additional cancer cells. Drs. Southam and Mandel were found by the New York State Board of Regents to be "guilty of fraud or deceit and unprofessional conduct..." (10). Dr. Denton Cooley attempted transplantation of a sheep's heart into a dying human patient (11). He sought permission from no one. His act was soundly condemned as unethical by many of his colleagues in the medical community, but he did not violate any law or regulation. He defended himself by contending that he was acting in the best interests of his patient who died several days later.

Dr. James Shannon, the dynamic Director of NIH at the time, was alarmed by these incidents. He not only regarded them as violations of the rights of the research subjects, but he believed that they jeopardized the future of publicly supported medical research. Intense discussions began to take place in his office and across the NIH campus in Bethesda, MD, concerning appropriate professional ethical standards for the conduct of research. Questions of informed consent, minimizing research risks, excellence of research design, protection of subjects' privacy, and submission of research to review by other scientists were discussed. It was initially thought that NIH could not, without additional legal authority, impose standards of conduct on investigators who received support for their research from the agencies of PHS. Ethical behavior toward research subjects was considered to be a major responsibility of the research investigator.

Gradually the view emerged that federal agencies could place conditions, including ethical conditions, on the awards that they made to research institutions and their investigators. Dr. Shannon created a committee headed by Dr. Robert Livingstone to recommend a suitable set of protections for subjects involved in Public Health Service-supported research. Livingstone's report recommended making a careful assessment of "ethically responsible relationships" and an examination of the range and tenor of present professional ethical practices. The report concluded that "the NIH is not in a position to shape the educational foundations of medical ethics..." Shannon found this conclusion "wholly unsatisfactory" (12).

FORMULATION OF THE FIRST EXTRAMURAL FEDERAL POLICY

Concerns for the ethics of publicly supported research involving humans had been relatively dormant during the decade of the 1950s, at least until the thalidomide hearings raised serious questions about the conduct of research. The years from 1960 to 1972 provided a sharp contrast to the previous decade. A generation of young adults had grown to adulthood under the constant threat of nuclear destruction. Trust of science-based technology had gradually eroded. For a new generation of young Americans, suspicion of technology and fear of its misuse replaced the confidence and trust that had characterized the period following World War II. Virtually every kind of science-based technology was now viewed as a potential threat to health, well-being, or survival of the planet unless it was carefully controlled.

New federal regulatory agencies — the Environmental Protection Agency (EPA), the Occupational Safety and Health Administration (OSHA) — were created to make the environment and the workplace safer for all living things and especially for humans. Both old and new federal departments and agencies produced volumes of regulations to implement laws restricting perceived threats to the environment, public health, and public safety.

The new-found public suspicion of science and technology provided a dramatic and sharp contrast to the attitude of many in the research community who believed that "the scientific method rests [on] the integrity and independence of the research worker and his freedom from control, direction, regimentation, and outside interference" (13).

The director of NIH had few illusions about the difficulty of formulating an ethical framework that could serve as a guide for research involving human subjects. "To win general acceptance within, not only the medical research community, but also our society at large," Dr. Shannon wrote, "the final statement of principles should probably emerge from ... representatives of the whole ethical, moral, and legal interests of our society" (12, p. 152). Shannon knew that creating acceptable ethical guidance for research, and building a consensus for a statement of ethical principles, would be the work of many years. He also believed that the Congress would hold the PHS agencies, particularly NIH, responsible for failures to protect human subjects. He had the vision to see that failure to protect human research subjects was tantamount to a failure to protect the public support of biomedical research.

At Shannon's request Surgeon General William Stewart issued the first PHS extramural Policy for the Protection of Human Subjects. The policy was elegant, simple, and easy to understand. To this day, although it has been revised many times, the core of the policy remains essentially unchanged. Each awardee institution was required to file an Assurance of Compliance document with PHS stating that the institution would adhere to the following procedures.

> The awardee institution is to provide prior review of the judgment of the principal investigator or program director by a committee of his institutional associates. This review should assure an independent determination: (*1*) of the rights and welfare of the individual or the individuals involved [as subjects], (*2*) of the appropriateness of the methods used to secure informed consent, and (*3*) of the risks and potential medical benefits of the investigation (14).

It almost seemed that Henry K. Beecher, a prominent Harvard research investigator, wished to underline the importance of the policy when he published a shocking article in the *New England Journal of Medicine* in June 1966 (15). Beecher identified 22 research projects published in refereed research journals that, in the author's judgment, violated the rights of the research subjects involved. Because he enjoyed a prestigious position as a researcher in a leading academic institution, Beecher's article sent shock waves through much of the American research community. His article lent urgency and needed credibility to the new PHS policy.

Responsibility for implementing the PHS policy was assigned to the tiny Institutional Relations Branch of the Division of Research Grants within NIH. The process by which the policy was to be implemented required considerable clarification. It underwent minor revisions in the summer of 1966, and further revision in 1967. In 1969 the Surgeon General revised the policy to make it clear that the policy extended to behavioral and social science research as well as to biomedical research. The Institutional Relations Branch implemented the policy by negotiating Assurance of Compliance documents with each awardee institution. Few universities, clinics, and laboratories welcomed the policy, but after Beecher's article was published, most of them accepted the policy as inevitable.

Because the policy was directed toward extramural research, it did not apply directly to NIH's own intramural CC. It would be many years and many policy revisions later before the NIH intramural program came into full compliance with the policy that governed extramural institutions that received awards for research involving human subjects. Nevertheless, in 1966 Dr. Jack Masur, director of CC, appointed a committee headed by Dr. Nathaniel Berlin to update the CC policy of 1953. Masur was responding, in part, to the recently published PHS policy. Clinical Research Committees (CRCs) were created within the intramural programs of the categorical institutes (16). Consent of subjects (still referred to as "patients") was required only to the extent that the investigator was expected to make a note in each "patient's" chart that verbal consent had been obtained.

From the outset, the Institutional Relations Branch used education and negotiation as the primary tool of promoting compliance with the policy. Although the Institutional Relations Branch had authority to withhold awarded funds, for many years no sanctions were imposed on any institution for failure to comply.

In 1971 the PHS policy was revised by Dr. Donald S. Chalkley, director of the Institutional Relations Branch. The scope of the policy was expanded to cover all research supported by any agency, office, or unit within HEW (17). Consistent with the educational approach described above, the new HEW Policy for the Protection of Human Subjects (commonly described as the "Yellow Book" because of the color of the cover of the pamphlet in which it was published) not only set forth requirements that institutions were to meet (review by committee, informed consent, and evaluation of risks and benefits), but it provided a running commentary presenting reasons why these requirements were necessary.

In July 1972 the public press uncovered details of the infamous Tuskegee Syphilis Study involving 600 black males from Macon County, Alabama Started in 1932, the study systematically denied treatment for syphilis to approximately 400 men who were afflicted with the disease (the other 200 men were used as control subjects). Subjects were not informed of their diagnosis. They did not know they were involved in a research study, and even after penicillin became the drug of choice to treat syphilis, they were denied treatment.

The Tuskegee Study received national publicity because of hearings held in the Senate Health Subcommittee chaired by newly elected Senator Edward Kennedy (D. MA). The study was conducted by a series of PHS investigators over a period of more than 35 years. Assistant Secretary for Health, Dr. Monty DuVal convened a committee chaired by Dr. Jay Katz, M.D., J.D., of Yale University to review the study. At the recommendation of Dr. Katz, the study was closed. HEW subsequently paid millions of dollars to survivors and to families of those who did not survive to compensate them for harms caused by the study. In FY 1995 HHS paid $2.8 million and in FY 1996 HHS paid $1.88 million in compensation to survivors and heirs of participants in the Tuskegee Syphilis Study (18). On May 26, 1997, President Clinton, on behalf of the U.S. nation, apologized to survivors of the study and to their families. (Despite the fact that HEW/HHS has paid compensation to subjects in the Tuskegee Study, it has never adopted a policy of providing compensation to injured subjects in other situations.)

The Tuskegee Study was given wide publicity in the media, and it triggered public disclosure of a number of other alleged research abuses relating to psychosurgery, fetal experimentation, and illicit experimentation involving contraceptives. All of these matters were aired in the Senate Health Subcommittee hearings held periodically between 1972 and 1974.

In the meantime Dr. Robert Q. Marston, who succeeded Dr. Shannon as director of NIH, made a decision to promote protections for the rights and welfare of human research subjects. In an address to the College of Nursing at the University of Virginia, he declared an obligation

on the part of society to carry out research aimed at improving the health of vulnerable populations and to protect the rights of vulnerable research subjects in the process (22).

Dr. Marston then created a PHS-wide drafting committee chaired by Dr. Ronald Lamont-Havers to revise and upgrade the HEW policy of 1971. The mandate to the PHS committee was based on the Marston talk at the University of Virginia. In 1972 Dr. Marston also upgraded the Institutional Relations Branch to a position within the Office of the Director, NIH. Marston changed the name of the office to the Office for Protection from Research Risks (OPRR). Dr. Chalkley, who had directed the Institutional Relations Branch, was named the first director of OPRR.

Largely as a result of the Kennedy hearings and the Tuskegee scandal, a series of legislative initiatives pertaining to research involving human subjects were introduced in Congress in 1973–74. These included a bill sponsored by Senator Kennedy to create a regulatory commission (similar to the Securities and Exchange Commission) to oversee research. Senator Walter Mondale (D. MN) introduced legislation to create a National Advisory Commission to study the impact of advances in science and technology on American society. Congressman Paul Rogers (D.FL) introduced a bill that called for a "National Advisory Commission for the Protection of Human Subjects of Biomedical and Behavioral Research" (20). Congressman Angelo D. Roncallo (D. NY) introduced legislation to ban federal support for human fetal research. Roncallo claimed, erroneously, that NIH had been conducting gruesome research involving live aborted fetuses ex utero (12). Senator Jacob Javits (R. NY) urged passage of legislation extending and strengthening the requirements for informed consent from research subjects. Congressman Rogers and Senator Kennedy cobbled together a bill that included some features of all of these legislative initiatives. Senator Kennedy indicated that he would support the compromise bill if HEW would publish regulations for the protection of human subjects. HEW then hurriedly transformed the HEW policy contained in the "Yellow Book" into regulatory form and published it as Regulations for the Protection of Human Subjects (21). The Congress then passed the pending compromise bill as Title II of the National Research Act (P.L. 93-348) signed July 12, 1974. The Act required: (*1*) that HEW promulgate regulations requiring institutions to ensure compliance with the regulations in a manner acceptable to the Secretary HEW (this requirement was completed prior to enactment of the law), (*2*) creation of a National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research to study a wide range of health research issues, (*3*) a temporary moratorium on fetal research to remain in place until such time as the National Commission reviewed the matter of fetal research and made recommendations concerning continuation of the moratorium, and (*4*) a special study of the impact of science and technology. To prevent the recommendations of the new Commission from being ignored, the Act contained a "forcing clause," that is, the Secretary of HEW was required by law to accept recommendations of the National Commission, or publish in the Federal Register reasons for not accepting them. The law also recommended that the commission be succeeded by a HEW Secretary's Ethics Advisory Board to address the question of federal funding of human fetal research and other matters that the Secretary might assign to it.

Between 1974 and 1978 hearings and deliberations of the National Commission, chaired by Dr. Kenneth Ryan of Harvard University, captured the attention of the research community and the media. The National Commission endorsed the approach taken by HEW in 45 CFR 46, and it incorporated much of the work of the PHS committee headed by Dr. Lamont-Havers. The commission recommended adding additional protections for pregnant women and human fetuses (1975), prisoner subjects of research (1977), children who are subjects of research (1978), and persons who are institutionalized because they are suffering mental infirmity. The National Commission deliberated for four years, and it published 9 reports. Perhaps its best known report is the *Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research* (22).

After the commission completed its work, the Secretary directed the Public Health Service to accept all of its recommendations and to implement them.

HEW Secretary's Ethics Advisory Board

In 1978 HEW Secretary Joseph Califano established the Secretary's Ethics Advisory Board (EAB) to review controversial biomedical ethical issues and to issue reports and recommendations concerning them. James Gaither, an attorney from San Francisco who had worked with Califano in the Johnson administration, was designated as chairperson. The first issue that the EAB addressed was whether HEW should fund human in vitro fertilization (IVF) research. Shortly after the EAB began its work, Louise Brown, the first "test tube" baby, was born in England as a result of the experimental work of Drs. Steptoe and Edwards. Enormous publicity followed the birth of Louise Brown, and considerable ethics debate both in England and the United States followed the publicity. In the United States several research investigators submitted applications for federal support for human IVF research. Suddenly the deliberations of EAB were followed by reporters from all the major media and from the tabloid press as well. In May 1974 EAB issued a Report and Conclusions: HEW Support of Research Involving Human In vitro Fertilization and Embryo Transfer. While the board stopped short of recommending to the Secretary that such research should be funded, it unanimously declared that such research, if it meets certain conditions, is acceptable from an ethical standpoint. On the basis of the report, Secretary Califano was preparing to authorize investigators seeking support for human IVF research to compete for awards when he became engaged in a controversy with Hamilton Jordan, assistant to President Carter. The President fired Califano and replaced him with Patricia Harris. Secretary Harris had little interest in the ethical questions pertaining to research and never took action on the recommendations of EAB. EAB issued reports on (*1*) fetoscopy (recommending

that the department make funds available for the procedure), (*2*) nosocomial (hospital-induced) infections (recommending that the public be allowed to see data maintained by CDC showing the rate of hospital-induced infections), and (*3*) clinical trial Data (recommending that investigators be allowed to withhold release of preliminary clinical trial data under the provisions of the Freedom of Information Act).

In 1980 Secretary Harris failed to re-charter EAB, and its remaining budget was transferred to the newly created President's Commission for the Study of Medicine and Biomedical and Behavioral Research.

REVISION OF HHS RULES FOR THE PROTECTION OF HUMAN SUBJECTS

During the final two years of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research and for several years thereafter (1976–1980), a PHS committee chaired by the director of OPRR scrutinized hundreds of public comments on the reports of the National Commission and, in the light of those reports and the public comment on them, revised and expanded the regulations for the Protection of Human Subjects (45 CFR 46) last issued in 1974.

The work was completed in the final months of 1980. The proposed regulations included (*1*) Subpart A: Generic Regulations Providing Protections for all Human Subjects participating in Research; (*2*) Subpart B: Additional Protections for Research Involving Pregnant Women and Human Fetuses; (*3*) Subpart C: Protections for Prisoners involved in Research; and (*4*) Subpart D. Protections for Children involved in Research. HHS Secretary Harris signed the regulations on January 19, 1981, and they were published in the Federal Register six days later.

In addition to the recommendations made by the National Commission, the PHS drafting committee created some categories of research that were exempt from the regulations, and other categories that would allow for expedited review by the chairperson of the institutional review board or a person designated by the chair. The exempt and expedited categories were created precisely to reduce the workload of IRBs that were already heavily burdened. They were welcomed by the research community that felt that a reasonable trade-off had occurred — greater protections for high-risk research and fewer safeguards for research involving negligible or minimal risks.

In 1978 Congress enacted legislation that created a new ethics Commission called the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (23). Although the bulk of the work of the President's Commission was dedicated to ethical concerns in the delivery of health care, the commission issued two reports that dealt directly with the system of protections for human research subjects. It also issued a report on research involving genetic engineering (24).

From the point of view of federal policies for the protection of human subjects, the most important contribution of the President's Commission was the following set of recommendations: (*1*) All federal departments and agencies should adopt regulations of HHS, (*2*) the Secretary of HHS should establish an office to coordinate and monitor governmentwide implementation of the regulations, and (*3*) each federal agency should apply one set of rules for the protection of human subjects consistently to all research conducted or supported by the federal government (25).

The Secretary of HHS, through ASH, designated OPRR as the "lead" office to develop a common set of regulations across the government. However, OPRR was dealing with reduced budgets and severe restrictions on its outreach due to downsizing. Requests for personnel and budgets to carry out the Secretary's orders were denied. OPRR approached each agency in the federal system with a request for compliance with the recommendations of the President's Commission. Most of the agencies replied that they too were facing downsizing and could undertake no new initiatives. Nevertheless, OPRR was able to obtain some backing from the Office of Management and Budget on the ground that what was proposed was a simplification of a complex regulatory structure. The initial response from the agencies was disheartening. Each was willing to adopt the HHS rules — but with conditions attached. Each agency had different conditions so that, if all had been accepted, there would have been little left of the HHS Regulations for the Protection of Human Subjects that were used as a model for the proposed new governmentwide regulations. The tortuous process of winning agreement from all of the departments and agencies continued for more than seven years. Finally, the OPRR, after rebutting objections from the President's legal advisor, was able to gain support from the Office of the President's Science Advisor and from OMB. On June 18, 1991, final clearance was obtained. Sixteen departments and agencies, in addition to the HHS, simultaneously published the "Common Rule."

APPLICATION OF THE REGULATIONS TO WOMEN AND MINORITIES

As noted above, the Common Rule was based, in many ways, on the recommendations of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. That commission tended to regard biomedical research as risky and dangerous. It feared that persons who were sociologically disadvantaged would be exploited as research subjects. It was concerned that the burdens of research would fall on the poor, the uneducated, minority groups, and women. All of these were regarded as "vulnerable" populations. However, in the decade of the 1980s, a long period passed when there were few reports of research injuries or deaths. Furthermore, with the coming of the AIDS epidemic, most persons infected with HIV found that the best care, and the most advanced "treatment" for AIDS, could be obtained by participation in clinical trials. Consequently the protections that had been employed to prevent exploitation of the disadvantaged were now seen as discriminatory because they offered protections that served to discourage overinvolvement of these populations in research.

Furthermore, in the late 1970s and early 1980s the feminist movement identified many forms of discrimination against women. Women's rights became a major national political issue. Along with other concerns for women's

rights came the awareness that many drugs administered to women had not been tested in women. Physicians, fearing malpractice suits, began to warn women that the drugs they were prescribing had not been tested in women and that women must be willing to take them at their own risk. Dr. Bernadine Healy, who became director of NIH in 1987, made repeated demands on the Congress for increased funding of research involving women. Dr. Healy created the Office of Women's Health Research at NIH and dedicated discretionary research funds to projects related to women's health (26).

Although AIDs research was gradually modified to include women, complaints that women, especially pregnant women whose offspring might be infected with HIV, were excluded from highly desirable research were given wide publicity. Clinical studies that a decade before had been considered to be a heavy burden for pregnant women were now considered to be a prized benefit (20).

CURRENT PROBLEMS

The history of federal protections for human subjects continues to develop. In 1994 President Clinton created the Advisory Committee on Human Radiation Experiments (ACHRE) that dealt with some 4000 studies sponsored by the federal government that exposed human subjects to radiation. Most of these studies were conducted prior to the existence of any federal policies or regulations. Although much of the information concerning these studies is incomplete and fragmentary, ACHRE identified some cases of clear abuse and recommended compensation for those who had suffered injury as a result of such studies. ACHRE further identified serious deficiencies in the current system for protecting the rights and the well-being of human subjects (27).

Partly in response to ACHRE's report, President Clinton issued an Executive Order establishing the National Bioethics Advisory Commission (NBAC) to address questions dealing with the adequacy of federal regulations, particularly regulations for the protection of the cognitively impaired. This commission has attempted to address the implementation of the Common Rule and has raised the question whether the location of OPRR within the NIH constitutes an apparent conflict of interest. Since NIH funds research, OPRR can easily be caught in a position of opposing its own agency if it decides that NIH-funded research is not being conducted in compliance with federal regulations. Partly as a consequence of the possible conflict of interest, a decision has been made to transfer OPRR to the Office of the Secretary of HHS.

NBAC also has issued reports on the ethics of research involving human cloning, and research involving persons who are cognitively impaired. The list of issues pertaining to human subjects that could be explored or revisited by NBAC is almost endless (28).

Research involving human subjects funded by the U.S. government is expanding each year. The United States is already the largest single source of such funding, and such research appears to be a growth industry. Slowly but surely such research has come under increasing public scrutiny and regulation. Some believe that much more needs to be done. Others believe that overregulation may stifle the biological revolution that is extending human life, conquering disease, and providing better health not only to Americans but to humans in every part of the globe. The new millennium offers new and ever more complex challenges to the twin goals of advancing biomedical and behavioral science and protecting the rights and welfare of human research subjects.

BIBLIOGRAPHY

1. *NIH Almanac*, 1997, Available at: *http:www.nih/gov/welcome/alamanac97.htm*
2. Available at: *http//www.nih/gov/welcome/alamanac97/chapt1/legis.htm*
3. Department of Health and Human Services, October 1979, P.L. 96-88.
4. *Records of the Office of Scientific Research and Development*, Committee on Medical Research, Contractor Records, National Archives, Washington, DC Record Group 127.
5. D.J. Rothman, *Strangers at the Bedside*, Basic Books, New York, 1991.
6. A.R. Jonsen, *The Birth of Bioethics*, Oxford University Press, New York, 1988.
7. C.R. McCarthy, in *Encyclopedia of Bioethics*, vol. 4, Macmillan, New York, 1995, pp. 2285–2290.
8. Available at: *http://www.nih.govod/ofmbudget/totalbymechanism.htm*
9. *Group Consideration of Clinical Research Procedures Deviating from Accepted Medical Practice and Involving Unusual Hazard*, a policy statement issued by the NIH Clinical Center, 1953.
10. E. Langer, *Science* **143**, 551–553 (1964).
11. D.A. Cooley et al., *Am. J. Cardiol.* **22**, 804–810 (1968).
12. M. Frankel, *Public Policy Making for Biomedical Research: The Case of Human Experimentation, Dissertation*, George Washington University, Kansas City, MO, 1976.
13. C.J. Van Slyke, *Science* **104**, 569 (1946).
14. *PHS Policy and Procedure Order 129*, Feb. 8, 1966.
15. H.K. Beecher, *N. Engl. J. Med.* **274**, 1360–1364 (1966).
16. *NIH Clinical Center Policy*, issued in July 1966, Updated 1976, 1977.
17. *Institutional Guide to the Department of Health Education and Welfare Policy on Protection of Human Subjects*, Publication No. 72–102 HEW, Washington, DC, 1971.
18. Personal communication from the office for Protection from Research Risks to the author.
19. R.Q. Marston, *Medical Science, the Clinical Trial, and Society*, Presented at dedication of McLeod Nursing Building and Jordan Medical Education Building, University of Virginia, November 10, 1972.
20. C.R. McCarthy, in J. Kahn, A.C. Mastroianni, and J. Sugarman, eds., *Beyond Consent: Justice in Federal Research Policy*, Oxford University Press, New York, 1998, p. 20.
21. Title 45 Part 46, *Code of Federal Regulations*, May 30, 1974, *Regulations for the Protection of Human Subjects*.
22. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, U.S. Government Printing Office, Washington, DC, 1979.
23. The President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, P.L. 95-622 Title III 1978, P.L. 93-377, 1982.

24. President's Comission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life*, U.S. Government Printing Office, Washington, DC, 1983.
25. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. U.S. Government Printing Office, *Implementing Human Research Regulations*, Washington, DC, 1981.
26. C.R. McCarthy, *Acad. Med.* **69**(9), 695–698 (1994).
27. Advisory Committee on Human Radiation Experimiments, *Final Report*, Oxford University Press, New York, 1996.
28. A. Capron, *Kennedy Inst. Ethics J.* **7**(1), 63–80 (1997).

See other HUMAN SUBJECTS RESEARCH entries.

HUMAN SUBJECTS RESEARCH, LAW, LAW OF INFORMED CONSENT

MARSHALL B. KAPP
Wright State University School of Medicine
Dayton, OH

OUTLINE

INTRODUCTION

This article discusses the legal implications of research involving somatic cell gene transfer for the treatment of human disease. The genetic process is based on the transfer of normal genetic material (deoxyribonucleic acid, or recombinant DNA) from another organism into a diseased human being. Gene delivery can be achieved by either directly administrating the gene-containing viruses or DNA to the blood or tissues (i.e., in vivo) or indirectly introducting cells manipulated in the laboratory to harbor foreign DNA (i.e., in vitro). Gene therapy seeks to treat disease in an individual person by the administration of normal DNA rather than a drug (1). This therapy would be available at all stages of human development.

Somatic cell gene transfer is distinct from medical interventions that target the manipulation or engineering of germ-line cells (i.e., eggs and sperm) (2), although the ethical distinction has been questioned by some writers (3). Additionally therapeutic gene transfer is distinct from genetic manipulations done for personal enhancement. Although both forms of cell gene transfer attempt to "improve" the human body, therapeutic manipulations aim at normalizing a person with a genetic disease or abnormality, while genetic enhancement tries to optimize some designated area of a normal person's body or performance (e.g., athletic prowess) (2).

Recently there has developed much enthusiasm in the scientific and public communities about the therapeutic potential of genetic manipulation at the individual patient level (4). It has followed the successful efforts of biologists who in 1998 cultivated human embryonic stem cells. These embryonic stem cells are the primordial human cells from which an entire person may be created, via fertilized human eggs (ova) that are implanted in the uterus (5). As philosopher Daniel Callahan has stated it:

> [N]o excitement [in the biomedical arena] has quite matched that which genetic research has engendered. The claim in its behalf is sweeping and radical: genetic research and its clinical application promise to finally bring medicine to the root causes of disease.
>
> Once these casual, molecular mechanisms are understood, clinical medicine and medical technology will be in a superb position to eliminate many, if not most, of the deadliest diseases (6, p. 70).

According to a 1995 federal report:

> Somatic gene therapy is a logical and natural progression in the application of fundamental biomedical science to medicine and offers extraordinary potential, in the long-term, for the management and correction of human disease, including inherited and acquired disorders, cancer, and AIDS. The concept that gene transfer might be used to treat disease is founded on the remarkable advances of the past two decades in recombinant DNA technology (1, p. 1).

The positive potential of somatic cell gene transfer was recognized in a 1982 report of the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (7) and in a 1984 report of the Office of Technology Assessment (OTA) (8).

LEGAL REGULATION OF HUMAN SUBJECTS RESEARCH

Research involving somatic cell gene transfer in human beings is governed by the same federal laws that apply to human subjects research generally, namely regulations codified at 45 Code of Federal Regulations Part 46. These 1981 regulations define research as "a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge," 45 Code of Federal Regulations §46.102 (d). The requirements established by these regulations apply explicitly to all human subjects research that is federally funded, although most research-sponsoring institutions subject all of their human subjects protocols to these requirements. Compliance with federal regulations is overseen by the Office of Protection from Research Risks (OPRR), which Congress in 1999 placed in the Office of the Secretary, U.S. Department of Health and Human Services.

Under federal law, research protocols utilizing humans as research subjects must be reviewed and approved by a local, interdisciplinary Institutional Review Board (IRB) whose composition includes at least one public

member (9). The IRB is supposed to evaluate the ethical implications of several aspects of a research protocol, including the following items, in terms of human subjects protection:

- Fairness or equity of subject selection
- Minimization of risks to subjects (including nonphysical risks, e.g., inconvenience and financial or loss of privacy costs), and whether any foreseeable risks are justified by reasonably expected direct, indirect, or social benefits
- Sufficiency of the informed consent process and its documentation
- Sufficiency of confidentiality safeguards

Questions have arisen regarding the propriety and degree of IRB involvement in evaluating the scientific merit of research proposals. The most cogent position is that usually the scientific merits of a research proposal cannot be disentangled from ethical considerations of subject protection. As a 1998 report (10, p. 32) asserted, "Any research involving humans entails some risk. Even for research in which the risk to subjects is minimal, the risk should not be taken unless the research has scientific merit."

In addition to federal requirements, a number of individual states have enacted their own statutes or promulgated regulations that apply to human subjects research generally. State requirements may be more, but not less, stringent in terms of human subjects protection than those created by federal law, and they apply with full force to research on somatic cell transfer.

One Canadian working group has suggested a detailed schema for IRB review of gene therapy/gene transfer protocols. This schema proposes that the IRB delve in depth into the following aspects of research protocols in this sphere: background and justification; research design; procedures; confidentiality; subject selection; risks, discomforts, and benefits; and information to subjects (11).

Another potential source of regulation is the civil justice system under which an individual human subject may sue investigators and/or the institutions and sponsors for injuries resulting from a negligently conducted experiment (12). The extent of possible liability exposure under traditional or novel tort theories remains uncharted at this time, but certainly investigators and their institutions will need to develop and implement appropriate risk management strategies to minimize that exposure (13).

SPECIFIC LEGAL ISSUES IN HUMAN GENE TRANSFER RESEARCH

Conflation of Research and Therapy

Background on the Confusion. The tremendous potential of somatic gene transfer technology to correct or manage many serious human diseases and abnormalities has been vastly oversold thus far, to both the scientific and general public communities. This excessively exuberant publicity has created, or at least fostered, a widespread confusion among members of the public who comprise the pool of potential human subjects for gene transfer research protocols regarding the important distinction between medical interventions that are properly categorized as therapy, on one hand, versus medical interventions more correctly considered research studies, on the other (14,15).

The terminology frequently used by many of those who participate in or report on this activity illustrates, and contributes to, the conflation of therapy and research in the public's mind. When terms such as "gene therapy research," "patient," "immunotherapy," "vaccine," "drug," and "cure" are repeatedly employed in this context, the public impression is established and strengthened that this form of medical intervention must be undertaken primarily for the benefit of the individual person receiving the recombinant DNA from another organism. As one set of authors has observed, "The call of activists, patients, researchers, and regulators for greater access to research protocols emerges as a symptom of the overselling of medical research as therapy." (14, p. 43) Although the commonly used terminology represents the long-range aspirations of researchers, it neglects the reality that interventions currently being studied in clinical trials are a significant way off from achieving the status of accepted therapies.

Properly understood, the practice of accepted therapy consists of professionally agreed-upon interventions that are expected to provide direct, personal benefit to an individual patient. By contrast, research has new, generalizable knowledge as its primary aim; any direct or indirect benefit enjoyed by specific human subjects, while certainly welcome, is quite secondary and incidental (16). Recognizing that overselling the current capabilities of gene cell transfer misleads potential human subjects into expecting personal benefit as a direct consequence of their participation, an NIH panel has advised more restraint by investigators and their sponsors in presenting information about this activity to the public (1).

"Informed" Element of Informed Consent. Federal regulations and common law doctrine (17) mandate that individuals may participate as human subjects in biomedical and behavioral research protocols only if and when they have given consent to such participation that is voluntary, competent, and informed. Widespread public failure to accurately distinguish between research and therapy in the gene cell transfer context is manifested by a tendency on the part of many people to already see this intervention as a benefit to which there is a right rather than as a hazard from which vulnerable people might need external protection (14). This popular attitude presents investigators and IRBs with a major challenge for ensuring that the process of informed consent operates in this sphere.

The 45 Code of Federal Regulations §46.116 (a) (1) requires that in seeking informed consent for research participation, each potential subject must be provided with "a statement that the study involves research, [and] an explanation of the purposes of the research ...." Fulfilling this requirement obligates the investigator to make clear to each potential subject that the protocol

is intended to gather generalizable information that may help others in the future who have the same disease or abnormality as the subject, but that any immediate, personal benefit to that particular person would be unexpected and merely incidental to the research endeavor. In the written consent document that the possible participant is asked to sign, the educational effort about the research versus therapy distinction ought to be promoted by replacing misleading words like "treatment," "patient," and "therapy" with more accurate terms such as "potential subjects," "potential participants," "research," and "experiment." (15, p. 51).

Many human subjects in gene therapy research might benefit from long-term follow-up by the investigators. However, there often will be practical difficulties in the capacity of investigators and their institutions to conduct such follow-up over a sustained period of time. These limitations should be explicitly described to potential subjects as part of the informed consent process (18).

"Desperate Use" Exception. The widespread confusion regarding the research versus therapy status of gene cell transfer today is illustrated by the "desperate use" exception that has been officially recognized as a circumstance where gene cell transfer investigators are legally authorized to conduct transfer interventions on a person without first complying with the ordinary informed consent requirements. The "desperate use" exception for gene transfer research is modeled after earlier-created "emergency use" exceptions authorizing the conduct of limited types of research in the absence of timely informed consent from the participant (19,20). Although the "desperate use" exception has been defended on grounds of compassion for afflicted individuals who have no other available alternatives (an argument that assumes the likelihood of some direct benefit for that specific patient/subject), some commentators question the wisdom of this exception in light of the present dearth of scientific knowledge about the actual benefits of gene transfer (15, pp. 49–50). There also is concern that, despite the lack of any scientific basis, the "desperate use" exception was recognized nonetheless as a result of pressure brought on Congress by patients, and over the strong scientific objections of the NIH's Recombinant DNA Advisory Committee (RAC) (15, pp. 49–50).

Voluntariness When the Subject Is Also a Patient. Voluntariness is an essential element of informed consent for research participation. The 42 Code of Federal Regulations §46.116 provides, "An investigator shall seek such consent only under circumstances that provide the prospective subject or the representative [of a decisionally incapacitated subject] sufficient opportunity to consider whether or not to participate and that minimize the possibility of coercion or undue influence."

Voluntariness is jeopardized in the gene transfer research context when, as is ordinarily and logically the case, potential human subjects are recruited from the ranks of patients who are currently under a physician's care for the precise disease or abnormality that is the focus of the research protocol. Many lay persons (as well as clinicians) have difficulty making the distinction between the patient seeking direct, immediate therapeutic effect from a medical intervention and the research subject altruistically contributing to generalizable knowledge of practical application only to others in the future. This confusion also impairs the "informed" element of informed consent. As a result patients may imagine, or accurately perceive, subtle or more blunt pressure to participate in research protocols as a way to continue access to whatever therapeutic treatment they are currently receiving, even though federal law expressly forbids investigators to conditionally link together research participation and the continued availability of standard treatment in that fashion.

As one commentator has noted, "Most of the persons enrolled in gene transfer research to date are a special class of research subjects — namely, sick persons who present special vulnerabilities and require special protections not relevant to healthy, 'normal' subjects." (15, p. 52) Clinical investigators and IRBs are legally and ethically challenged to establish an environment in which patients' perceived compulsion to "volunteer" for gene transfer research protocols solely in order to preserve their present patient status is eliminated or reduced as much as possible.

Conflicts of Interest. Investigators in gene transfer research involving human subjects may have actual or apparent conflicts of interest of two different sorts (21). First, since many potential human subjects for these research protocols are current patients receiving care for the exact disease or abnormality that the protocol intends to study, the clinical investigator may also occupy the professional role of attending physician currently caring therapeutically for the person (patient) that he or she is recruiting to now also become a research subject. This dual investigator/clinician position can easily contribute to the conflation of research and therapy in the minds of many patients/potential human subjects as just discussed, raising implications for both the voluntary and informed elements of informed consent to research participation. At the very least, the informed consent process — as enforced by IRBs — ought to require that potential human subjects be told in understandable lay terms about the dual investigator/clinician role, the distinction between the goals and expectations of research versus the goals and expectations of therapy, and the various kinds of professional and financial incentives that the treating clinician has to enroll his or her patients into research protocols for which that clinician also serves as investigator.

A second source of possible conflict of interest arises when the investigator in a gene transfer research protocol is also an investor in, or otherwise holds some financial stake (e.g., as a paid consultant), in a corporation that stands to profit financially from positive results of the particular study. A number of ethical commentators have argued that, as a matter of informed consent, financial interests of the investigator in the conduct or outcome of a study should be fully disclosed to potential human subjects at the time that their participation in that study is solicited (22–24).

Proxy Consent When the Subject Is a Child. As noted by one geneticist:

> Often the individual for whom DNA treatment [sic] is proposed is a very young child and thus unable to give consent except via proxy. In-depth counseling of the proxy — often the parents — should be provided to ensure full review of the state of knowledge concerning the risks of changing the DNA: what the known risks are, as well as what is unknown. The relative risks and benefits of alternative treatments and options should be fully discussed (25, p. 570).

The challenge of using minors as research subjects is exacerbated by the fact that many of them will achieve adulthood during the follow-up phase of the research.

> Thus, it may be necessary to obtain assent at the time of enrollment as well as consent at adulthood. This must be done in a way that both protects the rights of the patient [sic] to withdraw from research at any time and yet encourages follow-up, which may provide benefits to both the individual and society (26, p. 661).

National Institutes of Health (NIH) Oversight

In addition to the generic federal and state regulations applicable to human subjects research and common law principles of informed consent, gene transfer research involving human subjects that is funded by the federal government through the National Institutes of Health (NIH) also may be influenced by specific policies of that executive agency. Specifically, Section 402(b) (6) of the Public Health Service Act, as amended, codified at 42 United States Code §282(b) (6), established the NIH Recombinant DNA Advisory Committee (RAC). According to this Technical Committee's May 27, 1997, Charter, its function is to "advise the Director, NIH, concerning the current state of knowledge and technology regarding DNA recombinants, and recommend guidelines to be followed by investigators working with recombinant DNA." Review of specific protocols is conducted initially by the Human Gene Therapy Subcommittee (HGTS). The RAC is governed by the provisions of the Federal Advisory Committee Act, as amended, 5 United States Code Appendix 2.

ACKNOWLEDGMENT

Some of the writing of this entry took place while the author was the 1998–99 Dr. Arthur Grayson Memorial Distinguished Visiting Professor of Law and Medicine, Southern Illinois University School of Law, Carbondale, IL.

BIBLIOGRAPHY

1. S.H. Orkin and A.G. Motulsky, Co-chairs, *Panel to Assess the NIH Investment in Research on Gene Therapy, Report and Recommendation*, National Institutes of Health, Washington, DC, 1995.
2. T.H. Murray, *Int. J. Technol. Assess. Health Care* **10**, 573–582 (1994).
3. R. Moseley, *J. Med. Philos.* **16**, 641–647 (1991).
4. M.C. Coutts, *Kennedy Inst. Ethics J.* **4**, 63–83 (1994).
5. N. Wade, *N.Y. Times*, November 6, p. A1 (1998).
6. D. Callahan, *False Hopes*, Simon & Schuster, New York, 1998.
7. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life: A Report on the Social and Ethical Issues of Genetic Engineering With Human Beings*, President's Commission, Washington, DC, 1982.
8. U.S. Congress, Office of Technology Assessment (OTA), *Human Gene Therapy: Background Paper*, OTA, Washington, DC, 1984.
9. R. Levine, *Ethics and Regulation of Clinical Research*, 2nd ed., Urban & Schwartzenberg, Baltimore, MD, 1986.
10. N.A. Holtzman and M.S. Watson, eds., *Promoting Safe and Effective Genetic Testing in the United States: Final Report of the Task Force on Genetic Testing*, Johns Hopkins University Press, Baltimore, MD, 1998.
11. K.C. Glass et al., *IRB: Rev. Hum. Subj. Res.* **21**, 1–9 (1999).
12. J.G. Palmer, *Hum. Gene Ther.* **2**, 235–242 (1991).
13. J.G. Palmer, *Ann. N.Y. Acad. Sci.* **716**, 294–305 (1994).
14. L.R. Churchill et al., *J. Law, Med. Ethics* **26**, 38–47 (1998).
15. M.T. Lysaught, *J. Law, Med. Ethics* **26**, 48–54 (1998).
16. Office for Protection from Research Risks, *The Belmont Report*, OPRR Rep., ACHRE No. HHS-011795-A-2, U.S. Department of Health, Education and Welfare, Washington, DC, 1979.
17. Karp v. Cooley, 493 F2d 408 (5th Cir 1974).
18. F.D. Ledley, *Adv. Genet.* **32**, 1–16 (1995).
19. C. Marwick, *J. Am. Med. Assoc.* **269**, 83 (1993).
20. B.J. Crigger, *Hastings Cent. Rep.* **23**(3), 3 (1993).
21. R.G. Spece, Jr., D.S. Shimm, and A.E. Buchanan, eds., *Conflicts of Interest in Clinical Practice and Research*, Oxford University Press, New York, 1996.
22. J. LaPuma, C.B. Stocking, W.D. Rhoades, and C.M. Darling, *B. Med. J.* **310**, 1660–1661 (1995).
23. J. Moreno, A.L. Caplan, P.R. Wolpe, and Members of the Project on Informed Consent, Human Research Ethics Group, *J. Am. Med. Assoc.* **280**, 1951–1958 (1998).
24. E.J. Emanuel and D. Steiner, *N. Engl. J. Med.* **332**, 262–267 (1995).
25. P.A. Baird, *Perspect. Biol. Med.* **37**, 566–575 (1994).
26. F.D. Ledley, B. Brody, C.A. Kozinetz, and S.G. Mize, *Hum. Gene Ther.* **3**, 657–663 (1992).

See other HUMAN SUBJECTS RESEARCH entries.

INFORMING FEDERAL POLICY ON BIOTECHNOLOGY: EXECUTIVE BRANCH, DEPARTMENT OF ENERGY

DANIEL W. DRELL
U.S. Department of Energy
Washington, District of Columbia

OUTLINE

INTRODUCTION

The intention to establish an activity, linked to a major U.S. government science research program, directed at Ethical, Legal, and Social Issues (ELSI) was announced to an unsuspecting world on October 1, 1988, by Dr. James D. Watson, the co-discoverer with Francis Crick of the structure of the deoxyribonucleic acid (DNA) molecule. Watson was accepting the directorship of the newly created Office of Human Genome Research at the National Institutes of Health (NIH). Together with the Office of Health and Environmental Research (now the Office of Biological and Environmental Research) at the U.S. Department of Energy (DOE, which had been supporting genome research on a modest scale for the prior two years), these programs were preparing to map and sequence the human genome in its entirety, an estimated 3.2 billion nucleotides. In response to a question about dealing with the societal implications, Watson said that he thought that a small percentage of the research budget, initially 3 percent, should be used to study the implications of mapping and sequencing the human genome. Thus, even before the formal beginning of the Human Genome Project (HGP) in 1990, project managers, researchers, and lawmakers recognized that increasing knowledge about human biology and personal genetic information would raise a number of complex issues for individuals and society and that the agencies ultimately paying for the research needed to engage in anticipating its impacts. In response to Congressional mandates for identifying and defining such issues and developing effective policies to address them, DOE and NIH have devoted 3 to 5 percent of their annual HGP budgets, since the outset of HGP, to studies of the project's ELSI. [Much of this early history is well described in Cook-Deegan]. These expenditures for ELSI research by DOE and NIH now total nearly $85 million dollars.

DOE, which initiated HGP from roots in its radiation biology programs dating from the establishment of the Atomic Energy Commission (DOE's predecessor) in 1947, has for most of its 53-year history been an agency whose top priority was producing the materials necessary for the U.S. nuclear arsenal. This in turn made it an appropriate place to support studies of the biological effects of the radioactivity that characterizes the materials needed for nuclear weapons. Over the years, the greatest fear about radiation and radioactivity derived from its well-known capacity, first established by Herman Muller in the 1920s, to cause mutations and various forms of cancer. Mutations are alterations in the arrangement or sequence of the specific informational units in the genome, all made up of DNA. Just as a refrigerator magnet passed over a prerecorded cassette tape can alter the music on the tape, so too can various radiation exposures alter the information content of the DNA in the genome. To better devise tools and methods for assessing radiation effects, it is helpful to begin with knowledge of the original information prior to any exposures (e.g., the prerecorded music on the cassette *before* the magnet passes near it). This was DOE's rationale for initiating HGP. Additional reasons for DOE's activity included the availability of advanced technologies from the DOE National Laboratory system (e.g., for DNA fragment isolation, DNA sequencing, and computational analysis of sequences), better understandings of genetic contributions to workplace susceptibilities (given the particular nature of DOE facilities, characterized by mixed radioactive and toxic material waste), and the extensive experience of DOE with managing large interdisciplinary projects. For NIH, knowing the sequence of the human genome is the substrate for exploring what all the gene products do, how aberrations in them can lead to diseases, and how variations in individual genomes can lead to variation in individual humans. Thus both agencies had important and compelling mission-related reasons to participate in genome research. For the constituencies of both agencies, and for the larger general public, the genome project promised (and to a considerable degree has begun to deliver) many benefits in the way of new technologies, new industrial entities in the biotechnology sector of the economy, and new approaches to a better understanding of human biology in health and disease.

Still, there were worries about uses of genetic information, often of a very personal nature, that would not be welcomed by society. Early workshops enumerated several major ones. Among these were the implications of being able to predict future illnesses well before any symptoms or medical therapies existed; potential abuses of

the privacy of genetic information by employers, insurers, direct marketers, banks, credit raters, law enforcement agencies, and many others; the availability of large amounts of genetic information in largely unprotected data banks; and the possible discriminatory misuse of genetic information. One hypothetical outcome, albeit perhaps an extreme one, of the wide use of genetic screening would be the creation of a new genetic underclass, leading to a host of new societal conflicts and exacerbating others of long standing. Additional concerns included the difficulties that physicians (many of whom had little genetics in their medical school curricula) would have in incorporating rapidly advancing genomic knowledge and technologies into practice, especially in an Internet enabled world where a patient often might know more about a particular condition than his or her physician. As biotechnology entered the business world, issues of whether gene sequences could (or should) be patented arose. In addition, since newly discovered genes often lead first to a test for the presence or absence of alleles associated with a disease, and the demand for a new test could be high, tension could be forseen between the natural economic imperative to sell a new test and the uncertainty of the significance of test results when the strength and nature of the putative disease association was not clear (e.g., BRCA1 and susceptibility to breast and ovarian cancer). As genetic advances rushed forward, there was concern that a variety of communities, from specific groups such as judges in the court system, to high school biology classes, to legislators, the press, MDs of various specialties, nurses, ultimately to the wider public, could not keep up. One group of concern were genetic counselors, who were appreciated as vital to helping patients and their families understand the implications (to the degree they were understood at all) of genetic test results; as a community, genetic counselors are recognized to be in drastically short supply.

The ELSI component of HGP began by emphasizing the privacy of genetic information, its safe and effective introduction into the clinical setting, fairness in its use, and professional and public education. In 1991, recognizing the breadth of this scope, DOE narrowed its ELSI focus to concentrate on genetic education, privacy and fair use of personal genetic information, the implications of intellectual property protection (e.g., patenting) of genes, and genetics and the workplace. As its portion of the total ELSI program, DOE has supported peer-reviewed studies on the uses, impacts, and implications of personal genetic information in various settings the ownership, access, and protection of genetic information in computerized databases, tissue and sample archives, and the commercialization of products of genome research. DOE also supports studies of ways in which society and its institutions deal with ELSI issues surrounding the complex (and common) multigenic conditions and disease susceptibilities most of us fear. One of DOE's goals is to ensure the wisest possible distribution of the knowledge from HGP to the general public as well as to the scientific, academic, minority, judicial, medical, educational, sociological, and political communities.

To avoid unnecessary duplication of effort by their two independent ELSI programs, DOE and NIH have collaborated on a number of activities and maintained close communications over the years. In addition to the NIH–DOE Joint ELSI Working Group, which periodically consulted with program staff and assisted in coordination of the ELSI programs, collaborations also involved several research projects, joint conferences and workshops, and programs supported with other agencies, organizations, and commercial companies. DOE and NIH collaborated in supporting the ELSI Research Planning and Evaluation Group (called ERPEG), a successor to the ELSI Working Group.

EDUCATION ACTIVITIES

In keeping with the long-standing commitment of DOE to education, the DOE ELSI program has emphasized the promotion of knowledge about the HGP and its ELSI implications to such groups as institutional review boards, medical professionals, genetic researchers, judges, policy makers, and the public. These efforts have included conferences, seminars, publications, videos, Web sites, and radio and television programs. A few examples are given below (a more complete listing is available at *http://www.ornl.gov/hgmis*) maintained by the Human Genome Management Information System (HGMIS) at the DOE Oak Ridge National Laboratory. Although not an ELSI program, HGMIS helps the DOE Human Genome Program fulfill its educational commitment by making information accessible to scientists, policy makers, and the public about the program's goals, funded research, implications, and applications. HGMIS carries out this mission by publishing and distributing documents and information both in print and via its Web site, which serves a wide and varied audience that includes the general public, scientific investigators, and medical professionals. One HGMIS product is the *Human Genome News*, the newsletter of the U.S. Human Genome Project. This Newsletter, originally jointly sponsored by both the DOE and the NIH Genome Programs, facilitates communication among genome researchers and informs the broader public about the project.

The DOE ELSI program has sponsored, in whole or in part, several widely distributed booklets on the genome project. These include *To Know Ourselves*, by Douglas Vaughan, an overview of the underlying science of the Human Genome Project; the *DOE Primer on Molecular Genetics*, by Denise Casey, an introduction to the basic science of genetics; *Your World, Our World*, a magazine dealing with the science of genomics and its ELSI implications designed for grades 7 through 10 by the Pennsylvania Biotechnology Association, in cooperation with the Alliance for Science Education; *Your Genes, Your Choices*, a book and video designed for low-literacy adults, by Catharine Baker and Maria Sosa of the American Association for the Advancement of Science (AAAS); and a comprehensive bibliographic database of publications related to ethical, legal, and social issues surrounding the genome project and including thousands of books and articles from 1990 to 1995, compiled by Michael Yesley of the DOE Los Alamos National Laboratory.

The DOE ELSI program has also assisted with video documentary projects including, "Medicine at the Crossroads," a four-part documentary jointly sponsored with NIH, produced by George Page and Stefan Moore at WNET/Thirteen in New York and shown around the country on public television in the spring of 1993. A book, *Medicine at the Crossroads: The Crisis in Health Care* by Melvin Konner, also resulted from this series. Another documentary that DOE ELSI supported was a WGBH series, "The Secret of Life," produced by Paula Apsell and Graham Chedd and shown in the fall of 1993. "A Question of Genes," a two-hour television documentary produced by Noel Schwerin, looked at a series of families and individuals challenged by the outcomes of genetic testing for inherited diseases. This was broadcast on the Public Broadcasting System in September 1997, and it received an Emmy nomination. "Seeking Truth, Finding Justice," a three-hour television documentary series, is now in production by Noel Schwerin now of Backbone Media. This series will explore the impact of cutting-edge science (e.g., genetic technology) on the courts and its profound effect on democratic institutions, people's relationships, and notions of truth, justice, and individual rights.

DOE ELSI has also supported two radio programs. For the Spanish-speaking public, DOE ELSI has supported the broadcast of 50 Spanish-language radio episodes within a nationally syndicated science and technology series, produced by The Self Reliance Foundation in Santa Fe, NM. "The DNA Files," a nationally syndicated series of radio programs on the social implications of human genome research, by SoundVision Productions in Berkeley, CA was broadcast over National Public Radio stations in November 1998.

Other DOE education activities have included a web site launched by the Shriver Center in Waltham, MA called "The Gene Letter," (now called GeneSage) a freely available, online quarterly ELSI newsletter for interested professionals and consumers (*http://www.genesage.com/professionals.html*). The San Francisco Exploratorium and the Smithsonian's Museum of American History each received DOE/ELSI support in 1995 for exhibits on genetics and the HGP. "Diving Into the Gene Pool" was on display at the Exploratorium from April to September 1995. Most recently, at Stanford University, a project is producing an interactive multimedia CD-ROM medical education course for physicians, most of whom have received little or no training in clinical applications of molecular genetics. It is expected that many other groups will benefit as well from such a resource.

Over the years, DOE ELSI has supported many workshops and conferences to explore and educate audiences about the genome project and ELSI. Among the more notable of these have been a May 1997 meeting at the University of Maryland in Baltimore to inform minority communities about HGP and to make the aspirations and interests of these communities better known to genome project scientists and policy makers. In September 1996 DOE and NIH jointly sponsored a major conference at Tuskegee University on "Plain Talk about the Human Genome Project," which addressed some of the HGP's implications for African-Americans. An updated compilation of all the conference talks was published in 1997. The Cold Spring Harbor DNA Learning Center on Long Island conducted several workshops early in the Genome Program to inform opinion leaders and public policy makers on genomics and the genome project's implications for society. In a three-year DOE and NIH project that began in 1990, the Baylor College of Medicine and the Texas Medical Center Institute of Religion organized two national conferences on genetics, religion, and ethics. This project led to a book, *On the New Frontiers of Genetics and Religion* (1995). In April 1998, a symposium attended by about 850 people on "The Human Genome Project: Science, Law, and Social Change in the twenty-first Century" was held at the Whitehead Institute for Biomedical Research. Following the symposium, which was supported in part by DOE, a CD-ROM containing the meeting syllabus, transcripts of all plenary talks, and links to relevant Web sites, were distributed to over 4000 people. A second conference in this series was held in Spring 2000.

Education for Professional Groups

The Einstein Institute for Science, Health, and the Courts has convened workshops around the country for more than 1500 federal and state judges. These workshops, conducted since 1994 and continuing through 2002 (with co-funding from the National Instititue for Environmental Health Sciences of NIH), prepare judges for the expected onslaught of cases that will involve some aspect of genetics. In addition, the Summer 1997 issue of *The Judges' Journal of the American Bar Association*, with support from DOE, was devoted to "Genetics in the Courtroom." This project has a particular priority in the DOE ELSI program, since in our litigious society it can be anticipated that the courts will be the place of last resort for many issues. Many judges are touchingly honest about their very limited science education and their feeling of being ill-prepared since several recent decisions of the U.S. Supreme Court have assigned responsibility for deciding what scientific evidence is allowed into court to the trial judges and, for federal courts at least, altered the criteria by which scientific evidence must be judged before it is admitted. There are many potential issues for the courts, including numerous ones that are not rooted in medical practice. Issues from forensic uses, the penalty phase after a felony conviction, adoption cases, cases of the custody of a minor or of an incompetent parent, inheritance cases, liability, and many others may be influenced by assertions of genetic causation or involvement. Courts will have little alternative but to deal with cases arising from assertions that "my genes were responsible" cloaked in various guises.

Institutional review boards (IRBs) are responsible for overseeing clinical research procedures at such institutions as hospitals and research facilities, particularly protocols that might affect the rights and welfare of human research subjects. Increasingly complex ethical, regulatory, and scientific issues are proving challenging to these boards, which often are composed largely of nongenetic professionals. To assist IRBs in reviewing genomic protocols, expecially those involving human

subjects and tissues, a project with the Public Responsibility in Medicine and Research is focusing on educating IRBs in the special language, technologies, and issues that typify these protocols.

Curricula

The DOE ELSI program has supported the production and dissemination of a number of curricula for high schools on the genome project and its implications. These have included a nationwide series of workshops created by the Cold Spring Harbor Laboratory in New York which introduce high school biology teachers to a laboratory-based unit on human DNA polymorphisms (genetic differences) and the ELSI aspects of the genome project. An innovative program initiated at the University of Washington, Seattle, allows high school students to perform DNA synthesis and sequencing in the classroom. Local teachers, as well as those from other states, attend a weeklong summer workshop in Seattle and receive continuing assistance with the experiments through equipment and technical advice after they return home. This program has allowed high school students to actually contribute DNA sequence to the human genome project database making them participants in the project.

The Biological Sciences Curriculum Study (BSCS) group in Colorado Springs has produced 4 high school modules for the DOE ELSI program, roughly every two years since 1993. The first was, "Mapping and Sequencing the Human Genome Program: Science, Ethics, and Public Policy," followed by "The Human Genome Project: Biology, Computers, and Privacy." The third module was "Nontraditional Inheritance and the Nature of Science." Most recently BSCS has released "Genes, Environment, and Human Behavior" which deals with complex traits and behavioral genetics. At California State University in Los Angeles, an ELSI-funded project has translated into Spanish the first BSCS module for use by students and their parents in selected high schools where the majority of the student body is Hispanic. In conjunction with other local and national organizations, the University of Kansas Medical Center trained over 175 high school science teachers annually as state "resources" in molecular genetics and the latest biotechnology methods. In a "trainer of trainers" model, they then prepared thousands of additional teachers, who are now reaching millions of students. At the Fred Hutchinson Cancer Research Center in Seattle, a researcher has established an electronic educational resource on the Web for republishing classic genetics literature (both papers and monographs). This project helps promote a foundation for understanding the new genetics and genome technology (*http://www.esp.org*). A more complete list of educational curricula developed with support from the DOE ELSI program is available at *http://www.ornl.gov/hgmis*.

Exploring Public Policy Issues

There are many public policy issues raised by advances in genome research and no ELSI program could hope to explore all of them. At the outset of the HGP, when the scope of potential ELSI issues was uncertain, DOE and NIH jointly supported a study by the Institute of Medicine to address a variety of issues raised by the rapid proliferation of predictive genetic tests in otherwise healthy individuals. This study led to the 1994 publication of *Assessing Genetic Risks: Implications for Health and Social Policy*, a report with recommendations for the use of genetic information in healthcare. Concentrating on Florida and Georgia, researchers at Morehouse School of Medicine and the University of Florida explored differences among state-supported programs for genetic testing, screening, and counseling. Jointly sponsored with NIH, this 1992 to 1994 project addressed the issue of confidentiality in a mobile society of broad ethnic diversity. A series of published papers resulted from these studies.

At the University at Albany in New York a project is examining confidentiality concerns raised by the possible misuse of DNA-based test results in the managed-care (MCO) setting. This setting presents unique ethical dilemmas because the MCO is both payer and provider and because physicians, and quite often, testing laboratory personnel are MCO employees. Additionally, a significant fraction of working Americans get their health care through MCOs. The uses they make of genetic information are a valid concern of ELSI.

Many disorders associated with mental retardation have genetic contributions, with Down syndrome and fragile X syndrome the most common, and new genetic findings from HGP pose very difficult ethical questions and legal and social concerns to those with disabilities and their family members. To address these concerns, The Arc of the United States developed and distributed a series of reports, fact sheets, and a workshop training package to all 1100 of the organization's chapters.

Threats to privacy based on genetic testing in various circumstances have been cited often as reasons for ELSI research and societal concern. The fair use of genetic information raises particularly difficult practical and philosophical problems related to access and disclosure. Third parties such as insurers, employers, adoption agencies, and educational institutions may feel they need access to genetic data having predictive or diagnostic value, while others feel that such access will lead to discrimination and decisions based too heavily on genetics. The DOE ELSI program has focused on a broad range of genetic privacy issues from the perspectives of several disciplines, including philosophy, social science, and law. Through grants and commissioned papers, the program has supported studies to consider such factors as attitudes toward genetic privacy in different populations, the need for appropriate measures to protect genetic information in various contexts, and evolving policies of private institutions and state, federal, and foreign governments in this area. These studies were designed to help increase the growing body of knowledge and to promote informed discourse leading to policy development.

A study by a leading social scientist at Columbia University related existing social science work on privacy to anticipated genetic-privacy issues. This study also examined current privacy-protection measures, debates over the need to update privacy protection, and implications for social and legal policies to deal with expected

future genetic testing and applications of genetic data. The results will appear in book form. To strengthen the dialogue between the professional genetics community and federal policy makers, a congressional fellowship program was initiated in 1995. This program allows genetic professionals to spend a year as special legislative assistants on the staff of members of Congress or on congressional committees. The fourth fellow (of a planned five) began her fellowship at the beginning of 2000. Under a grant to the Library of Congress, Philip Kitcher, a philosopher at the University of California, San Diego, researched genome ELSI issues and wrote *The Lives to Come: The Genetic Revolution and Human Possibilities* (2). The book explores both the science and the ethical and moral dilemmas arising from the genome project and concludes by arguing that society should make active use of genetic testing to avoid the worst genetic conditions that presently are not amenable to medical therapies, but involve horrific suffering, for example conditions such as Tay-Sachs disease.

In a project jointly sponsored by NIH and DOE, researchers at the University of California, Berkeley, conducted a major study between 1992 and 1997 to illuminate the processes by which genetic screening and genetic concepts of health and illness were integrated into the health concerns of high-risk families. This study focused on issues of privacy, stigmatization, and discrimination and how these issues were managed within family and institutional networks in two contrasting communities: one in which the disorder was generally recognized as race-associated (sickle cell disease) and the other in which this association was not part of the public consciousness (cystic fibrosis). Interestingly, cases are being identified of individuals of the "wrong" community with these conditions, namely Cacausians with sickle cell disease and African-Americans with cystic fibrosis. A series of scholarly papers and book chapters have resulted.

DOE sponsored a workshop on "Medical Information and the Right to Privacy" in June 1994 in Washington, DC, which led to the book *Genetic Secrets*, edited by Mark Rothstein of the University of Houston (3). A compilation of chapters written by leading experts, the book explores the full range of issues related to genetic privacy, particularly focusing on issues arising from the possible use of genetics in the workplace, a particular focus of the DOE ELSI program.

Privacy Legislation

Focusing on privacy concerns, some proposed legislation has attempted to establish a legal framework of fair practices for health information and to regulate its access, disclosure, and use. A draft bill (the Genetic Privacy Act), was drawn up in 1995 by George Annas of the Boston University School of Public Health to assist legislators. The bill proposed that access to information in genetic data banks should be regulated during sample collection and when it is stored, disclosed, and used. Several state lawmakers used language and concepts from this draft bill in drawing up proposals for legislation in their own states. The Genetic Privacy Act and Commentary is on the Web (http://www.ornl.gov/hgmis). In 1994 a Shriver Center study surveyed existing bills and laws with a view to drafting model legislation for protecting the privacy of personal genetic information. They found that state legislative efforts to regulate the use of such data have increased, particularly in employment and insurance, but major gaps and deficiencies in statutory coverage persist. From 1997 to 1998, Mark Rothstein determined the effects of a unique Minnesota law limiting employee medical records to job-related matters, with a view to using the law as a model for protecting genetic privacy in the workplace. He found that the Minnesota law had little impact because people were not aware of it. A journal article describing this study was published in 1998. At this writing, numerous state measures affecting genetic privacy have been passed, but each is somewhat different from the others, and no federal legislation focused on this issue has been passed by both the House of Representatives and the Senate of the United States.

Data Banks

Interest in forensic DNA data banks is growing, with all states except Massachusetts having laws that authorize the collection of samples from convicted felons. Two separate DOE studies carried out at the Shriver Center in Waltham, Massachusetts, focused on the growing practice of banking individuals' DNA or genetic data in forensic, academic, military, and commercial settings. These studies involved research on privacy in these settings and on developing and refining proposed policies and guidelines. One project, which reported widespread uncertainty about the types of sample releases that are legally or ethically prohibited, led to the production of a 28-minute video called *Banking Our Genes*. [Video: Fanlight Productions (800/937-4113)] Another study carried out a survey of life insurance companies and showed that they were more interested in obtaining existing genetic test information than in performing new tests on applicants. Company ratings based on genetic conditions reflect a considerable degree of subjectivity rather than actuarial data. How this will evolve as HGP nears completion and technologies make the acquisition of personal genetic profiles easier and less expensive is an ongoing ELSI challenge.

Patents

Since 1996 Rebecca Eisenberg at the University of Michigan Law School has explored the role of patents in transferring technology generated by the genome project to society at large. This issue, which was not expected to be a concern at the outset of HGP, exploded suddenly when, in 1991, NIH filed patent applications on ESTs (expressed sequence tags, which are short sequences of DNA from a gene that is known to be expressed in a cell). While this specific application was rejected by the Patent and Trademark Office in 1993 and NIH elected not to appeal the rejection, many other biotechnology companies have aggressively sought patent protection on genetic information. One company, Incyte Pharmaceuticals in California, has been awarded a patent on ESTs. Eisenberg now is conducting a study with Michael Heller on the trend toward privatization and patenting of the early stages of

biomedical and microbial genome research. A series of influential publications has resulted from this work.

CONCLUSION

The continuing importance of ELSI studies is rooted in the basic fact that each person has a unique genome that both identifies her or him as an individual and has predictive implications for her or his future health. An "ideal" or "perfect" genome does not exist, even if such a concept could be defined. All genomes contain many small differences that could severely and adversely affect health under different circumstances or if not influenced or masked by other genes. This information has potential value to other people and groups who may have their own agendas. Thus, while HGP's potential benefits are enormous for medicine, bioremediation, agriculture, and many other socially and economically important areas, we must remain alert to the more problematic implications as well.

HGP is rapidly moving to its goal of obtaining the complete human reference DNA sequence by 2003 and a useful "working draft" by Summer of 2000, well ahead of the original schedule. All of us will therefore face, much sooner than anticipated, many questions surrounding differences in our individual sequences, uncertainties regarding their significance for health and longevity, and the implications of knowing about these subtle distinctions before the biological effects are understood. Although many ELSI issues are not novel to the genome project, they nonetheless remain challenging and need to be addressed. Only by dealing directly and openly with such issues through the collective best efforts of bioethicists, scientists, policy makers, and the public can the benefits of genome research be realized and the difficulties minimized.

The ELSI programs have had significant impacts, both in terms of the research and other activities they have supported, but also in other tangible ways. Within DOE, ELSI has led to the creation of parallel programs, particularly the BASIC (Bioremediation and its Societal Implications and Concerns) program element of the larger Natural and Accelerated Bioremediation Research (NABIR) program. NABIR is focused on exploring biological approaches to the legacies of radionuclide wastes created by the 50-year history of the nuclear weapons programs of DOE and its predecessor agencies. DOE also has an obligation to address the challenges of human subjects experimentation associated with this legacy. Outside of the DOE, ELSI has contributed to the creation of the National Bioethics Advisory Commission (NBAC) attached to the White House Office of Science and Technology Policy. NBAC has a charter to explore both human subjects issues and genetic information issues. The creation of NBAC also takes cognizance of the challenges that will result from the novel technologies arising from HGP, among them genechips and microarrays. These technologies, and others, will make it easier to acquire accurate and precise information about specific genetic variations present (or absent) from large numbers of people and to correlate them with clinical conditions.

Just as genome scientists need to become active participants in the discussion of these difficult issues, so too must the bioethicists, lawyers, social scientists, and other "ELSI scholars" learn the relevant genetic science so that these very important dialogues remain firmly grounded in scientific reality. This is an area where much remains to be done and where dialogue and crosstalk are less in evidence (and sometimes less enthusiastic) than could be wished. There is also a major challenge for the private sector that has aggressively entered genomics to a degree exceeding even the public sector, spurred by the promise of commercial yields from patents on human and other genes. What one company does can influence the way other companies are regarded and the way their products are received. The genomics revolution (it is nothing less) will have major impacts on our lives in the twenty-first century, and we must use all the wisdom and insights of our intellectual and political ancestors to inform us and to build on as we try to make wisely the many difficult and momentous decisions that lie ahead. These decisions will have great impacts on the lives of our children and our society years into the future.

INFORMATION SOURCES

Abstracts for current and past ELSI projects can be found in the DOE Human Genome Program reports and in Contractor-Grantee workshop reports all of which are available on the HGMIS Web site (*http://www.ornl.gov/hgmis*)

ACKNOWLEDGMENT

I want to acknowledge the valuable and helpful suggestions of my colleagues in the DOE Office of Biological and Environmental Research (OBER). I particularly want to acknowledge contributions from Ari Patrinos, Marvin Frazier, Mike Riches, and Betty Mansfield (Oak Ridge National Laboratory).

BIBLIOGRAPHY

1. R. Cook-Deegan, *The Gene Wars: Science Politics and the Human Genome*, Norton, New York, 1994.
2. P.Kitcher, *The Lives to Come: The Genetic Revolution and Human Possibilities*, Touchstone Books, 1997.
3. M. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997.

ADDITIONAL READINGS

The ELSI literature has grown extensively since the onset of the ELSI programs at the DOE and the NIH. The selection below is not representative and other contributors will make their own suggestions which can differ with these. I don't necessarily endorse or agree with everything that is in these volumes, but as a place to start, they offer good introductions to the history of the Human Genome Project and many of the issues that have arisen associated with it.

L.B. Andrews, *Medical Genetics: A Legal Frontier*, American Bar Foundation Press, Chicago, IL, 1987.

L.B. Andrews, J.E. Fullarton, N.A. Holtzman, and A.G. Motulsky, *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994.

G. Annas and S. Elias, *Gene Mapping: Using Law and Ethics as Guides*, Oxford University Press, New York, 1992.

R. Cook-Deegan, *The Gene Wars: Science Politics and the Human Genome*, Norton, New York, 1994.

C. Cranor, ed., *Are Genes Us? The Social Consequences of the New Genetics*, Rutgers University Press, New Brunswick, NJ, 1994.

M.S. Frankel and A. Teich, eds., *The Genetic Frontier: Ethics, Law and Policy*, AAAS Press, Washington, DC, 1994.

S.J. Gould, *The Mismeasure of Man* (revised and expanded), Norton, New York, 1996.

D. Hamer and P. Copeland, *Living With Our Genes*, Random House/Doubleday, New York, 1998.

D.J. Kevles, *In the Name of Eugenics: Genetics and the Uses of Human Heredity*, University of California Press, Berkeley, 1985.

B.M. Knoppers, C.M. Laberge, and M. Hirtle, eds., *Human DNA: Law and Privacy, International and Comparative Perspectives*, Kluwer Law International, The Hague, The Netherlands, 1997.

C. Long, ed., *Genetic Testing and the Use of Genetic Information*, AEI Press, Washington, DC, 1999.

R. Plomin, *Nature and Nurture*, Brooks Cole, Pacific Grove, CA, 1990.

M.A. Rothstein, ed., *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, Yale University Press, New Haven, CT, 1997.

Y. Segal, *The Human Genome Project: Legal, Social, and Ethical Implications. Proceedings of an International Workshop, Jerusalem, 1995*, Israel Academy of Sciences, Jerusalem, 1997.

E. Smith and W. Sapp, eds., *Plain Talk about the Human Genome Project: A Tuskegee University Conference on its Promise and Perils ... and Matters of Race*, Tuskegee University Press, Tuskegee, AL, 1997.

L. Weir, S.C. Lawrence, and E. Fales, *Genes and Human Self Knowledge: Historical and Philosophical Reflections on Modern Genetics*, University of Iowa Press, Iowa City, 1994.

See other FEDERAL POLICY MAKING FOR BIOTECHNOLOGY entries.

INTERNATIONAL ASPECTS: NATIONAL PROFILES, FRANCE

FRANÇOISE TOURAINE-MOULIN
Immunology Laboratory
Hôpital Neurologique Pierre Wertheimer
Lyon, France

MARISSA VICARI
La Tour de Salvagny
Lyon, France

MARIA RUANO
CATHERINE CHABERT
Alain Bensoussan — Avocats
Lyon, France

OUTLINE

INTRODUCTION

The development of biotechnology is in France, as in other European countries, moving ahead strongly. Today the ethical debate, which has urged the reevaluation of new biotechnologies and commercial dynamics of the biotechnology industry, is also proceeding in full swing at the level of the general public, special interest group, government structures, and research facilities.

Although France is considered as being behind in comparison with the United States in industrial development and innovative entrepreneurship support, it is, without a doubt, competitive in life sciences research as well as for legal and ethical rules on biotechnology.

There has always been a high regard for life-sciences research in France, where pioneering advances were made in certain fields; the classic example is Louis Pasteur's contributions in microbiology, and later vaccinology. There were also major contributions by André Lwolff, François Jacob, and Jacques Monod (his achievement recognized internationally by the 1965 Nobel Prize in Medicine) to molecular biology in the discovery of the role of messenger ribonucleic acid (mRNA), and by Pr. Jean Dausset to knowledge in organ and tissue transplantation regarding the major histocompatibility complex (MHC). More recently in the field of virology a French team at the Pasteur Institute in Paris, directed by Pr Luc Montagnier, won acclaim with their discovery of HIV1 which is responsible for AIDS.

ACADEMIC INVOLVEMENT

In the academic sphere the research in biotechnology is directed by sizable institutions for research and

technology. The Centre National de Recherche Scientifique (CNRS), National Center of Scientific Research comprises a large networks of university laboratories in which fundamental research in science and technology, including biotechnology, is carried out. In 1999, 25,400 people were employed in about 1300 CNRS research and service units throughout France and the French territories. The Institut National de la Santé et de la Recherche Médicale (INSERM), National Institute of Health and Medical Research is another such large organization dedicated to the "research, understanding and bettering of the human condition, with the main aim of promoting health for all." Created in 1964, INSERM includes 275 research laboratories and a community of over 10,000 scientific and medical professionals. Of special interest for bioethics INSERM houses a collection of 3000 books and 100 specialized journals on ethics in the Documentation Center on Ethics of Life Sciences and Health (CDEI). There are also the National Institute for Agricultural Research (INRA), the Centre d'Energie Atomique (Center for Atomic Energy) (CEA), and public universities which, like CNRS and INSERM, are under the direction of the Minister of Education, Research, and Technology.

INNOVATION AND ENTERPRISING IN BIOTECHNOLOGY

Before the law on innovation and research was passed on July 12, 1999, it was difficult for scientists working for public institutions to create their own companies to commercialize their innovative technologies. As a matter of fact, those scientists who are government employees cannot easily take part or become a partner in a private firm. The new law will allow French academic research to evolve into commercial enterprises.

Besides, the Ministry for National Education, Research and Technology (MENRT) a national database on biotechnologies, be set up, called "Biotechnologies/France," (1). The MENRT objective, among other things, is to show the wide range of French biotechnological activity and to promote the work abroad, as well as to inform and educate the public.

Biotechnology Companies

Until recently there were only around one hundred biotechnology companies in France. Their number has started to increase with the organization of an association called France Biotech and also the Syndicat National de l'Industrie Pharmaceutique (SNIP), the National Union of the Pharmaceutical Industry. The situation is helped by administrative efforts at the local level in the creation of regional "biopoles," or concentrations of biotechnology-related resources, and in facilitating the founding of biotechnological enterprises.

French biotech industry therefore may be said to be experiencing a full expansion, which is in keeping with a long tradition of quality scientific research. Health-related biotechnology companies in France employed in 1999 around 2000 people with a market of two billion French francs. Industrial and agricultural biotechnology has been strong in France since the 1970s within large pharmaceutical companies. Associated start-up companies could be seen by the 1980s, although without much development in technology transfer. Since 1990 the biotechnology industry in food and agriculture has seen another surge in the development of new companies.

Public Interest Groups

In French biotechnology research, associations and foundations, such the Centre d'Etudes de Polymorphisme Humain (CEPH), or (Center of Studies for Human Polymorphism), play an important role in upholding public interest. Their combined effort with the French government has led to expanded research and technological innovation in networking projects on the human genome (e.g., the first mapping of the human genome and in the robotics field, the newly designed "GenHomme" project) and on plant genomics (e.g., Genoplante, dedicated to transgenesis and genetically modified organisms).

The recent governmental projects aimed at fostering collaborations between various public and private laboratories have led to a concentration of start-up biotechnology companies in the area of genomics, and most of them are located in the Genopole (or "Genomic Valley") in Evry, near Paris.

THE "BIOETHICS LAWS"

History

Turning to the legislation on biotechnology, we find in France a historical consciousness of the ethical debate that has been cultivated and influenced by a strong collective memory of Nazi ideology and the disclosures of the Nuremberg proceedings (1947), by the Universal Declaration of Human Rights, and by the works of the World Medical Association, namely the Declaration of Helsinki. The translation of the human into legal terms in France has thus been difficult and attempted in consideration of issues in bioethics, specifically those that have arisen with the explosion of information coming from discoveries in molecular biology.

To rise to the challenge, France formulated the three Lois de Bioéthique, or "bioethics laws," of July 1994 (2), which specifically address medically assisted procreation, the protection of the embryo, and diagnostic medicine. These laws, which were among the first pieces of legislation adopted on the subject, were proceeded by numerous preparatory studies, beginning with the 1988 report by Guy Braibant, *Sciences de la vie, de l'éthique au droit* (Life sciences, from ethics to law) (3). Also intrinsic to the creation of the bioethics laws is the large work completed by Nöelle Lenoir in 1991, *Aux frontières de la vie* (4).

Une éthique biomédicale à la française (At the frontiers of life: French biomedical ethics) as well as *Les science de la vie et les droits de l'homme: bouleversement sans contrôle ou législation à la française? Questions clefs et réponses contradictoires*, (Life sciences and human rights: tumbling out of control or French legislation? Key questions and contradictory answers) by Senator F. Serusclat (1992) were also influential (5). The work continued in 1994 with a report by M.P.J.F. Mattéi, *La vie en questions:*

pour une éthique biomédicale (Life in questions: toward a biomedical ethic) (6). Punctuating these is a series of advisory reports issued by the Comité Consultatif National d'Ethique pour les sciences de la vie (CCNE, the National Consultative Ethics Committee for life sciences), created in 1983, the first committee of its kind (7). In addition, existing legislation such as the Caillavet law of 1976, which regulated organ donation, and regulations relative to the protection of people participating in biomedical research (1988), thus defining their legal status, were taken into consideration, namely the Law of Huriet Sérusclat, December 20, 1988 (modified July 1994).

Already present in the Code Civil were the principles of the inviolability and nonobjectification of the human body, which have been reaffirmed by the bioethics laws. The idea of the protection of the human body has been expanded as the necessary consideration of a persons' totality, including the genetic identity has become apparent.

It is to the credit of the bioethics laws that introduced into the Code Civil is a chapter on the respect of the human body. In this regard the law ensures the primacy of the person, prohibiting all diminution of their dignity and guaranteeing the respect for the human being from the beginning of his or her life. It is also interesting to note that the Conseil Constitutionnel (which is charged with controlling the conformity of the laws with the French Constitution) asked, with the formation of the bioethics laws, that the laws enunciate the set of principles designed to ensure the constitutional preservation of human dignity.

General Principles

The principles at the origin of the bioethics laws include (8):

- previous consent for all procedures that relate to the integrity of the body;
- the prohibition of all modification of heritable genetic material with eugenic ends and prohibition of all intervention aiming at affecting the descendants of the person concerned;
- noncommercialisation of the human body, therefore prohibiting the selling of organs and tissues;
- anonymity with regard to the donation of organs, elements, or products of the human body;
- the impossibility of contesting the familial relations or descent of a child for reasons related to medically assisted procreation;
- the regulation of the utilization of genetic tests and techniques of medically assisted reproduction;
- the regulation of information related to the donation of organs.

The bioethics laws thus offer very specific regulations on essential procedures that protect human dignity. The laws strive to consider the ethical issues brought up by each relevant technological step. They have arisen from the general apprehension of the ethical implications of biotechnology for the human and living organisms. The largely governmental initiative has materialized via the promulgation of numerous legislative reports and regulations, which leave little room for individual initiative. The ultimate aim of the heavy control of research in general, and biotechnology in particular, is to avoid at all costs, any legal lacunae. Nevertheless, the body of extremely complicated and rich legislation that has resulted is highly technical and, because of continual developments in this technology, eternally incomplete.

Provisions

The first of the bioethics laws is in the Code Civil, law 94-653 of July 29, 1994, which pertains showing to respect for human life. This law covers the study of the genetic characteristics of a person and the identification of a person via genetic fingerprinting, protection of the embryo, and filiation (legal status relations between one person and his or her descendants) in cases of medically assisted procreation.

On the respect for the human body, the law clearly prohibits any contravention from the beginning of life (Article L. 16), although the point at which life begins is not fully defined by French law. The human body is considered as inviolable (Article L. 16-1), and the integrity of the human body may only be challenged in cases where the intervention is therapeutic (Article L. 16-3). Any action challenging the integrity of the human species is prohibited, including all eugenic practices designed to select the sex of the embryo, and any modification of the genetic characteristics, with the exception of research directed at the prevention of genetic diseases (Article L. 16-4). No remuneration may be given to a person who agrees to submit to medical experimentation or to donate parts or products of his or her body (Article L. 16-6). Of interest is Article L. 16-7 which officially renders null and void any contract relevant to procreation or gestation regarding a third party. This is relevant to the donation of gametes and also to the issue of contracting surrogate mothers, which is prohibited. Any information on the identification of an organ donor, and the person who received the organ is strictly confidential (Article L. 16-8). Neither the donor's family nor the recipient may learn the identity of the other, and only in the case of medical necessity may the doctor of the patient receiving the donation have access to such information.

The second law (94-654) of July 29, 1994, relates to the donation and utilization of human fetal tissues, cells, and other parts, medically assisted reproduction and prenatal diagnosis. It is known as the Code de la Santé Publique (Public Health Code). This law also addresses organ donation from persons living and deceased, organ transplantation, the donation, use and conservation of fetal tissues, cells and products, medical assistance to procreation, and specific regulations regarding the donation and utilization of gametes. Notably this law prohibits in vitro fertilization (IVF) for the purpose of research or experimentation and has enacted a ban on all embryo experimentation, with certain exceptions for medical research. In line with the principle of noncommercialization of the human body, the law prohibits the creation and use of embryos for commercial or industrial purposes. In addition no remuneration may

be received for an embryo. The same principle is applied to parts of the fetus, including tissues, cells and organs. Because of the large number of surplus embryos resulting from in IVF procedures, and the therapeutic possibilities associated with breakthroughs in cloning, the mandatory five-year legislative review of this law has been the subject of much public debate.

There is also an information protection law that has been modified by a third bioethic law; law 94-548 of January 1, 1994, on the confidentiality of personal data, amending law 78-17 of January 6, 1978, on information, medical records, and access, and the confidentiality of personal data for research purposes in the field of health.

Medically Assisted Reproduction

The extent to which the laws control certain technical processes, and their features related to the protection of human dignity within the letter of the law, are topics of notable interest. A look at the procedures of medically assisted reproduction can provide a good idea of the legislations' effect.

By the reproduction legislation and the decrees that have followed, France has instituted a highly organized structure to oversee medically assisted reproduction cases. These laws regulate public hospitals and private clinics alike. In particular, the Commission Nationale de Médicine et de Biologie de la Reproduction et du Diagnostic Prénatal (CNMBRDP, National Commission for Biological and Reproductive Medicine and Prenatal Diagnosis) was created by decree in April 1988. The members convene regularly, give advice, and authorize IVF practices in individual cases (which must also be further ratified by the Minister of Health). This committee has been responsible for the elaboration of several decrees on reproductive medicine and prenatal diagnosis. As a result the techniques pertaining to medically assisted reproduction are strictly codified to the point where practically every necessary act is the object of a specific decree.

Medically assisted reproduction procedures are in fact covered by La Sécurité Sociale, the French national health insurance. Thus economically the procedure is available to the general population. The whole process is therefore mitigated by the state via Sécurité Sociale, the organization that actually reimburses the expenses related to these procedures.

Under the law 94-654, medically assisted reproduction is restricted to couples whose infertility is due to a medically diagnosed pathology or who have the aim of avoiding the transmission of a serious disease to their offspring (Article L. 152-2). The couple must consist of a man and woman who are of reproductive age, and are either married or can prove that they have lived together for at least two years. Both members must be alive and consenting. An embryo may only be conceived in vitro for the purposes of medically assisted reproduction, and it must be conceived with the gametes from at least one member of the parental couple. In exceptional cases a couple may decide to donate, via written consent, an embryo to another couple who qualifies for medically assisted reproduction and for whom the technology has not been successful (Articles L. 152-4 and 5). In this case the whole process is mediated by a judge. The donating and receiving couple remain anonymous; only for therapeutic reasons may the doctor have access to non-personal medical data regarding the donating couple and no payment may be received for the embryo (Article L. 152-5).

The excess embryos are kept frozen for a period of five years, and each year a letter is sent to the parent couple regarding the options of using the embryos for reproductive purposes and the continuation of embryo storage. The involvement of a third party as a gamete donor is only permitted in cases where the use of the couple's own gametes has not been successful, and the donation and utilization of gametes is also controlled by law. The donor must be part of a couple who has already procreated, and written consent for the donation is required of the donor and his or her partner (Article L. 673-1). Written consent is also required by both members of the receiving couple. Artificial insemination with fresh sperm or a mixture of sperm samples is prohibited, and the use of gametes coming from one donor is limited to the creation of five embryos (Articles L. 673-3 and 4).

The establishments practicing medically assisted reproduction technology must be authorized by decree. Authorization is accorded by the CNMBRDP and is for five years only. All establishments and laboratories authorized to practice medically assisted reproduction technology or prenatal diagnosis must present an annual written report of their activities to the Minister of Health, and must register the gametes and embryos that they hold in storage (Article L. 184-2).

The Code Pénal imposes a range of fines and prison terms, along with the temporary or permanent revocation of the right of the establishment to practice medically assisted reproduction or related activities. In addition any professional personally involved in a violation is subject to a maximum of 10 years suspension from professional activity. The nature of the sanctions include the taking of gametes from a living person without her consent, which is punishable by five years imprisonment and a fine of 500,000 FF. (Article L. 675-9). The same sanctions apply in obtaining gametes by payment of any kind, except that authorized to prepare and properly conserve the cells (Article L. 675-10).

Divulging information relative to the identity of a person or couple who have donated gametes and the couple who received them is punishable by two years imprisonment and a fine of 200,000FF. (Article L. 675-11). The same penalty is applied for the procurement of gametes from a living person without performing the required sanitary tests for transmissible diseases and proceeding with artificial insemination with fresh or mixed sperm samples.

Revision of "Bioethics Laws"

Taking into account the evolutionary nature of science, the bioethics laws include a legal exception requiring that they be reviewed every five years, thus underscoring the difficulty of reconciling the progress of the life sciences

and respect for the essential values necessary for the preservation of human dignity.

The philosophical principles that form the basis of the laws were formulated with scientific progress in mind, and thus to guide the evolution of the bioethics laws along with the evolution of science. Consequently the bioethics laws were to be reviewed in 1999, and a certain number of modifications are in view to take effect during the year 2000.

Today several basic ethical questions still remain unresolved regarding for instance the legal status of the embryo. The debate in France today is at the level of whether an embryo should or should not, from the first step of development, be considered as a potential human being. While considered, this question was not answered in full with the creation of the bioethics laws, and it has been raised once again with the first five-year revision of the laws, notably concerning the embryo research issue, nowadays forbidden in France. On this issue the French Conseil d'Etat proposes, in a recent report of November 1999, to permit research on frozen embryos no longer intended for a parental project. In performing such the research, the embryo in question cannot be used later for reproductive purposes (9).

On human reproductive cloning, all proposals that have been submitted to the French government, for the revision of the bioethics laws, consider that existing French rules do not suffice to make clear the prohibition in France of such practices. In this context it is clear that new legislation will be taken to forbid human reproductive cloning activities. However, therapeutic cloning could, in certain conditions, be allowed.

COMITE NATIONAL CONSULTATIF D'ETHIQUE POUR LES SCIENCES DE LA VIE ET DE LA SANTE

A highly significant reference body for decisions in bioethics is the National Consultative Ethics Committee for Health and Life Sciences (CCNE) established in 1983 via presidential decree and enacted with the law of July 29, 1994. CCNE is an independent body with the aim of forming opinions and publishing recommendations on ethical issues arising with technological progress in the fields of biology, medicine, and health. The group us comprised of 40 individuals. The president and 5 members of the main philosophies and religious faiths are chosen by the president of the republic, 19 members are engaged due to their qualifications and interest in ethical issues, and 15 are research scientists. Topics may be referred to the CCNE by government officials, a university or establishment for higher education, a public institution, or foundation working in research, technology development or health issues, other persons, or committee members themselves. So far the committee has revisited topics as necessary, in keeping with scientific developments or social issues. CCNE opinions are considered by legislators both in the development and drafting of new laws (10).

Such has been the case with the CCNE Opinion 1 on *the sampling of human embryonic tissue for therapeutic, diagnostic, and scientific purposes*. CCNE has revisited it under various subheadings in several Opinions since 1984. This was shortly after its formation, and CCNE then expressed the Opinion that the human embryo must be considered a potential human person and therefore must never be subjected to in utero experimentation, CCNE further prohibited the commercial or industrial use of living embryos and any remuneration for embryo tissue samples. The same sentiments appeared in Opinion 8 (1986) *on the research and use of in vitro human embryos for scientific and medical purposes*. The emphasis again is on human dignity, with the conclusion that fertilization should not be done for research purposes alone and must exclude any form of industrial or commercial use. These conclusions, among others, are reflected in the bioethics laws of 1994.

Of further interest, the CCNE developed, in Opinion 1, the policy that the use of embryonic tissue for therapeutic purposes must be exceptional and, *considering the present state of scientific knowledge*, justified by certain criteria including the rarity of the disease to be treated, the absence of alternative, equally effective therapies, and a benefit to the recipient, such as survival. An important factor of this policy is the suggestion that the successful development of scientific knowledge may alter the ethical issues at hand. As exemplified by the bioethics laws, this emphasis on the ethical consideration of new technologies has been incorporated directly into legislation.

This policy was applied in a CCNE report in 1997 in which there was re-evaluated the use of embryos in research in light of new technology: the creation of embryonic stem (ES) cell lines from human blastocysts obtained by IVF. Although such research was earlier prohibited under the bioethics laws, with agreement with previous CCNE reports, the CCNE concluded that the therapeutic possibilities of developing this technology actually weighed in favor of a modification of law. This modified opinion is a part of the CCNE's report on the five-year evaluation of the bioethics laws.

CCNE has addressed many other ethical issues and through its Opinions has presented a rich discussion of national position on ethics in technology and health issues. Among the other topics discussed by CCNE are AIDS, drug abuse in the workplace, local ethics committees, human genome research, the utilization of placebos in therapeutic trials involving antidepressants, the care of autistic children in France, and xenotransplantation to name only a few (5).

RESEARCH ETHICS

The Use of human subjects in research is addressed by the law Huriet-Sérusclat of December 20, 1988, modified in July 1994, on the protection of persons participating in biomedical research. The two main objectives of this law are the regulation of biomedical research on humans and the creation of a system of regional Consultative Committees for the Protection of Medical Research Subjects (CCPPRBs). CCPPRBs are charged with reviewing proposals for biomedical research projects as regulated under the law Huriet-Sérusclat. Established by the Minister of Health according to regional needs, CCPPRBs are independent and have legal

standing. Members are chosen by the regional government representative or the head regional committee from various specialities in the biomedical field to ensure the committee's independence with regard to ethical, social, psychological, and legal issues.

Under the law, the institution supporting the research project (whether public or private), called the sponsor, is responsible and held liable for any harm that comes to the subject as a result of the experiment. The investigator, the person directing or supervising the experiment, must be a medical doctor. Two kinds of research projects are distinguished under the law: those with direct therapeutic benefit to the subject and those without. The evaluative criteria for the two classes of research differ in order to better protect the subjects. In the case of research without direct therapeutic benefit to the individual, for example, there must not be any foreseeable serious risk to the participants health, and the research must be aimed at people of the same age group, or with the same disease or handicap as that of the of participant. There must further be no alternative way of performing the studies.

In research proposed to have direct therapeutic benefit to the subject, a limited risk to the subject is permissible. However, any risk must be weighed against the potential benefit.

CCPPRB, in their evaluation of research proposals, applies the ethical principle of protection of human dignity which are specifically codified in the law. CCPPRB may consider, for example, the recruitment methods used for the study and check that the methods respect the confidentiality of the subject. In keeping with noncommercialisation and nonobjectification of the human body, and noncoercion of the subject, there must be no volunteer remuneration, save the cost (or loss) to the subject in participating. The recruitment of certain subjects is restricted by law, such as pregnant women, people without health insurance, children, prisoners, the mentally ill, and people in critical medical conditions. Informed consent is required; volunteers must be informed by a medical doctor of the objective, methodology, and duration of the research and of the possible benefits or risks. Consent must be expressed clearly in writing.

CCPPRB also review the scientific validity of proposals. It adheres to three general principles: that all biomedical research on humans must be founded on the current state of scientific knowledge and sufficient preclinical experimentation, that the foreseeable risk of a project must be proportionate to the benefit to the subject or the research interest, and that biomedical research must be aimed at raising scientific understanding of the human being and the methods for ameliorating its condition. CCPPRB shall consult with local and institutional ethical committees if necessary.

REGULATION OF GENETICALLY MODIFIED ORGANISMS

The development of genetically modified organisms (GMOs) destined for food products and for pest control in agriculture has been a big topic of public debate in France over the past decade. The government's position has shifted concerning the different uses of GMOs and changed several times. This is essentially due to very strong public protest against GMOs. Although administrative regulatory groups (Commission Nationale de classement des recombinaisons génétiques in vitro, later Commission de génie génétique) and specifical rules (decree of July 30, 1985, and AFNOR guidelines), have been in place since 1975 and 1985, respectively, the controversy in France, as in several other European countries, erupted around 1990 with the enactment of the European Council Directives 90/219/EEC on the Contained Use of GMOs and 90/220/EEC on the Deliberate Release into the Environment of GMOs. It should be mentioned that the contained use of GMOs refers to use within the laboratory, and this use has not been a subject of controversy. However, the release and marketing of GMOs affects a larger public which is concerned about the accidental release of such organisms into the environment and other risks associated with their use in food and agriculture.

The two European Directives provide the foundation for regulations regarding GMOs in France; they were transcribed into a single French law 92-654 July 13, 1992, related to the control, utilization, and release of GMOs. This regulation falls under the rubric of environmental protection and modifies a preexisting law 76-663 of July 19, 1976, related to classified installations for environmental protection.

Overall, the European Directives defined two major points. The first was the clear separation within the regulatory framework of the contained use of GMOs and their release into the environment, which includes marketing in the later stages. The second was the establishment of national authorities in all European member states to deal with application procedures as defined under the Directives. In France, although the two Directives have been transcribed into national law under one regulation, two separate national authorities have been established, one dealing with contained use and one dealing with release into the environment and marketing.

Contained Use

The national authority for the contained use of GMOs is the Commission on Genetic Engineering (CGG), which operates under the State Secretary of Research. The purpose of the CGG is to determine measures of confinement suitable to the risks of the use of GMOs, the processes used to obtain them, and the utilization of genetic engineering technologies.

CGG itself is comprised of experts in genetic engineering, public health, and environmental protection plus a member of the parliamentary office for evaluation of scientific and technological options (OPECST). It may also bring in any necessary experts.

Deliberate Release

The deliberate release of GMOs into the environment is mediated by the Commission du génie biomoléculaire (CGB, the Commission on genetically bioengineered organism field releases, or the Commission on Biomolecular Genetics). CGB is responsible for risk assessments related to the release of GMOs into the environment, and advises

the Minister of Agriculture and the Permanent Technical Committee of the Selection of Plants and Cultivars (CTPS). CGB is composed of scientists in genetic engineering, public health, and environmental protection, a member of OPECST, consumer groups, and concerned professionals. It is required to produce an annual report.

The conditions for the release and marketing of genetically modified plants, seeds and seedlings, are given by decree 93-1177 of October 1993 and European Council Directive 94/15/EEC on the data needed for the application. The Directive was transcribed into French law by an Order of September 21, 1994.

The authorization of genetically modified plants must also come from the Minister of Environment and the Minister of Agriculture. Notification of releases are done at the local level, at the town hall. In addition, as with all plants, GMOs must be authorized and recorded as official varieties by CTPS. Marketing authorization comes from CTPS and generally requires two years of agronomic testing which must be accompanied by the opinion of the Superior Counsel of Public Hygiene.

Additional legislation relevant to GMOs includes decree 94-46 of January 6, 1994, which sets the conditions for the deliberate release of cleaning substances or products containing GMOs that may enter human or any animal diet through contaminated tools that come into contact with any drink or food substance. The planned release of pharmaceuticals composed of or containing GMOs is regulated under decree 94-359 of May 5, 1994, related to phytopharmaceuticals control. Finally, decree 95-487 of April 25, 1995, sets the conditions for the deliberate release of genetically modified animals.

Public Debate

In France, even with protective legislation in place, the introduction of GMOs into agriculture has met with much public resistance. The main problem is that there are still too many unknowns with regard to the safety of GMOs to humans, animals, and the environment.

For the French public a series of recent technology-related European biological disasters, namely the scandal of HIV contaminated blood and the "Mad Cow" disease problem, have brought home the importance of risk assessment and the precautionary principle. In addition the central place of the French cuisine in the food culture demands more vigilant regard of the process by which foods are produced and of the differences between industrial and nonindustrial produce. The tendency is to see genetically modified food as an aberration or degradation of the natural product. Besides the alarm about the safety of GMOs and the active movement by the agricultural community against GMOs, a real problem of acceptability based on food aesthetics exists in France.

The Precautionary Principle

Precaution is born of the concern the public, has about the risks that may exist in the uncertainty of scientific knowledge and the possibility of blunders creating grave and irreversible damage. New forms of risk are being considered starting with the environment and the effect on food. The issues focus on how to anticipate the risks and resolve them. Human safety is considered a civil right; this logic is used in decision making to regulate conduct in situations of uncertainty and to guide the regulatory process with prudence.

The precautionary principle is therefore enscribed at the heart of democracy. It must not be taken as a matter of fear but rather as a basic right. Society exercises precaution as it learns from past experiences and uses that information in mitigating future risks. The realm of the precautionary principle is at the interface of science, politics, and law. The principle responds to a priority of our century: safety.

Covering the domain of the environment in French law is law 95-101 of February 2, 1995 (the Law Barnier) on environmental protection. The idea behind it is that in the absence of certainty, scientific knowledge and cutting-edge technology must not defer the adoption of effective and proportionate measures designed to prevent the risk of serious and irreversible damage to the environment at an economically acceptible cost. A report about the precautionary principle from Pr. Philippe Kourilsky and Pr. G. Viney, requested by the Prime Minister, is in progress.

Several structures have been put in place with law 98-535 of July 1, 1998, relative to the reinforcement of sanitary surveillance and the sanitary control of products destined for use by humans. The agencies replace the Agence du Médicament (January 1993) (Medicines Agency) and the Agence Nationale du Médicament Vétérinaire (the National Agency for Veterinary Medicine).

In the surveillance and sanitary control system we find the following groups: the Institut de Veille Sanitaire (the Institute for Sanitary Surveillance), and two new agencies charged with alerting the public when there appear any menace to public health, of any origin. These are the Agence Française de Sécurité Sanitaire des Produits de Santé (AFSSAPS, or French Agency for Sanitary Security of Health Products) and the Agence Française de Sécurité Sanitaire des Aliments (French Agency for Food Sanitary Security). Finally, an institution named the Comité National de la Sécurité Sanitaire (National Committee for Sanitary Security) governs the functioning of these structures.

Unlike the developments leading to the bioethics laws, the regulation of GMOs was not preceded by a large body of government papers and public debate. Rather, the government and public involvement has mostly occurred after the fact, and its effect can be seen in the changes in French policy regarding GM plants over the past decade. In 1995 authorization by the Commission du Génie Biomoléculaire made France the first country to propose the marketing of transgenic corn from Novartis, which received wide European approval in December 1996. The cultivation of Novartis corn in France was authorized in February 1998, and in September of that year the authorization to market Novartis corn cultivated on French soil was revoked by the Conseil d'Etat. In December 1998 the Conseil d'Etat brought the problem of the cultivation of Novartis corn before the European Court of Justice, and in June 1999 the French Minister of the

Environment joined with Ministers of the Environment of several other European member states to declare a moratorium on new authorizations for the marketing of GM plants.

Highlighting the recent public debate on GMOs, and government involvement, is the Conférence de Citoyens (or the Citizens Conference on the Utilization of GMOs in Food and Agriculture), a public discussion event orchestrated by the Parliamentary Office for Evaluation of Scientific and Technological Options (OPECST) in 1998. OPECST'S administrative structure is devoted to technology assessment, with the goal "to inform Parliament of scientific and technological options in order, specifically, to make its decisions clear," as established by Act 83-609 of July 8, 1983. OPECST is an independent organization to which only members of Parliament may refer matters for study. OPECST then "collects information, launches study programmes and carries out assessments." In the Conférence de Citoyens, a representative sample of 15 people was chosen through a survey to participate in a panel of citizens in a public discussion with a panel of experts and government officials with competence in GMOs issues related to food and agriculture. The citizens had two weekends of training prior to the conference, in order to furnish them with an understanding of the principles and issues at hand. The conference was a two-day public media affair, where the panel of citizens posed questions to and entered into discussion with the panel of experts. The conclusions of this panel of citizens is published in annex to the OPECST report on biotechnology in food and agriculture: *From Understanding Genes to Making Use of them*.

Other OPECST reports relevant to biotechnology include topics of biotechnology and the agro-food industry, biodiversity, the life sciences, human rights, and the landmark review of the Lois Bioéthiques, *Application of Law No. 94654 of 29 July 1994 concerning the Donation of Human Body Parts and Products, Medical Assistance with Reproduction, and Prenatal Diagnosis*, by Alain Claeys M.P. and Senator Claude Huriet, 1999 (11).

Despite the strong resistance to the marketing of GMOs, several releases, mostly in the form of field trials, are underway. According to a report published by the European Commission Joint Research Centre, there have been 485 summary notifications for the release of GMOs in France between October 21, 1991, and October 1, 1995. Since each notification covers either the release of one GMO at several sites or an ensemble of GMOs at one site, this number does not represent the total number of test sites in France since 1991, which is much larger. The tests have included 188 strains of transgenic corn, 106 strains of oilseed rape, and 61 strains of sugar beet, among several other species of plants, and they have been conducted by a number of international companies and included French companies and French public research institutions. So, while negative public opinion has had a large influence on the marketing and cultivation of GM plants in France, the number of summary notifications for the environmental release of GMOs in France for research purposes is actually the highest in Europe.

SUMMARY

As a summary we present a list of the principal official French regulatory and consultative groups in the field of biotechnology:

- Commission Nationale de Informatique et des Libertés (CNIL) (law of January 6, 1978)—National Data Protection Commission
- Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale (CCPPRB) (law of December 20, 1988)—Consultative Committee for the Protection of People in Biomedical Research
- Commission de Génie Génétique (CGG) (decree of May 11, 1989)—Commission on Genetic Engineering.
- Commission d'étude de la dissémination des produits issus du génie biomoléculaire (decree of February 23, 1993)—Commission on genetically bioengineered organism field releases, or the Commission on Biomolecular Genetics
- Comité Consultatif National sur le traitement de l'information en matière de recherche dans le domaine de la santé (law of July 1, 1994, and decree of May 9, 1995) Consultative Committee on the treatment of data in research in the health sciences.
- Comité Consultatif National d'Ethique pour les sciences de la vie de la santé (CCNE) (decree of February 23, 1993)—National Consultative Ethics Committee for Health and Life Sciences.

CONCLUSION

To conclude this overview on Biotechnologies and Bioethics in France, we note that France has had 15 years of reflection concerning the ethical and legal issues generated by new technologies in the life sciences. By instituting the bioethics laws of 1994, the government took a step that brought the message to the Minister of Social Affairs, Simone Weil at the time, that the law established "the primacy of ethics over technique." The French politic has thus had the objective of prohibiting eugenic or commercial derivations and sharing with society the choices that the scientific world cannot assume alone. Must ethics be legislated? This is the route that France has chosen; it has preferred a strict legal framework to less constraining guidelines. However, France has inaugurated an unusual statute with this legislation, the provision of a five-year review; a wise decision that will hopefully help avoid conflicts between the progression of scientific research and the general ethic of our society, which is concerned with prevention of eugenics and the appropriation of human by human.

BIBLIOGRAPHY

1. Biotechnologies/France, Available at: *http//biotech.education.fr*
2. Code de la Santé Publique, Dalloz, 1999.

3. G. Braibant, *Sciences de la vie, de l'éthique au droit, Conseil d'Etat, Notes et études documentaires*, La Documentation française Ed, Paris, 1988.
4. N. Lenoir, *Aux frontières de la vie. Une éthique biomédicale à la française, rapports officiels*, La Documentation Française Ed, Paris, 1991.
5. F. Sérusclat, *Les sciences de la vie et les droits de l'homme: bouleversement sans contrôle ou législation à la française? Questions clefs et réponses contradictories, Economica*, Coll. Office parlementaire d'évaluation des choix scientifiques et technologiques, Paris, 1992.
6. J.F. Mattéi, *La vie en questions: pour une éthique biomédicale, Rapport au Premier ministre*, La Documentation française Ed, Rapports officiels, Paris, 1994.
7. Comite National Consultated d'Ethique pour les sciences de la Sante (CCNE), Available at: *www.ccne-ethique.org*
8. C. Chabert-Peltat, *Les Biotechnologies l'Ethique Biomédicale et le Droit, Mémento-guide Alain Bensoussan*, Hermes Ed, Paris, 1995; Alain Bensoussan, Informatique et Telecoms, Editions Francis Lefebvre 1997, reprinted in 1999.
9. A. Claeys and C. Huriet, *Rapport sur l'application de la loi du 29 juillet 1994 relative au don et á l'utilisation des éléments du corps human, à l'assistance médicale à la procréation et au diagnostic prénatal*, OPECST, Paris, 1999.
10. *Cahiers du comité Consultatif National d'Ethique pour les sciences de la vie et de la santé* (CCNE), Biomed Ed, Paris, quarterly data.
11. J.-Y. Le Déaut, *De la connaissance des gènes à leur utiliisation, Premiére partie: L'utilisation des organisms énétiquement modifiés dans l'agriculture et dans l'alimentation*, vol. 1, Assemblée Nationale No. 1054, OPECST, Paris, 1998.

See other INTERNATIONAL ASPECTS entries; INTERNATIONAL INTELLECTUAL PROPERTY ISSUES FOR BIOTECHNOLOGY.

INTERNATIONAL ASPECTS: NATIONAL PROFILES, GERMANY

MATTHIAS KETTNER
Cultural Studies Center of Northrhein-Westfalia (KWI)
Essen, Germany

OUTLINE

INTRODUCTION

In Germany biotechnology is expected to be a prominent future industry ranking as high in perceived economic importance as today's information and communication technologies. However, public opinion is sharply divided between advocates of strong public sector involvement in this emergent field of cutting-edge industrial and scientific research and those who fear, for various reasons, that biotechnological "progress" means above all else overrated promises, technoeconomic determinism, and more "colonialization of the lifeworld" (Jürgen Habermas) at ever deeper levels of somatic existence. The resistance against biotechnology has for a long time been focusing on possible risks unique to genetic engineering. Since no strong evidence for this supposition has come forth, lately the "risk debate" has lost much of its momentum. The arguments that appeal to economic shareholders, and to patients and other stakeholders of medical and pharmaceutical progress, are winning as far as the public is concerned. Recent innovations in Germany's legal culture, partly owing to legislation on the level of the European Union (EU) of which Germany is a member state, support the trends on the scope, ease, and speed of patenting. In this sense, one could say that biotechnology in Germany is well on its way of becoming "normal" business.

GOVERNMENTAL SUPPORT OF BIOTECHNOLOGY

In Germany, in the recent past, advancement of biotechnology was a topic of strong political controversy. Many studies were made of public opinion in the 1980s and 1990s and showed that the public perception of biotechnology, specifically the perception of cutting edge genetic technology, tended to be more negatively biased in Germany than in other countries within the EU. At the same time, spokespersons from the economic sector and ministerial representatives of the federal government designated biotechnology as one of the "key technologies," as a field of promising research activity and economic interest for the future of Germany's economy. The federal government's perspective on biotechnology is richly documented in a 1985 official declaration by the Ministry for Research (1). (For a sceptical appraisal of this perspective from a sociological vantage point, see Ref. 2, for an optimistic one from the vantage point of a large life sciences enterprize, see Ref. 3.)

The tension that shaped the state administration's political discourse on biotechnology during the 16 years in which the Christian Democratic Party (CDU) was the ruling majority party in Germany did not ease after the Social Democrats (SPD) 1997 return to power. It also shaped the discourse on biotechnology of nongovernmental organizations that either promoted the advancement and acceptance of biotechnology or a control of developments in biotechnology which they perceived as politically dangerous or ethically unsound.

Political Economy of Biotechnology in Germany

Because of the long product development times in biotechnology, experts claim that the technical application

of knowledge produced by the biological sciences will not become a palpable reality until sometime after the year 2000. Forecasting studies of the commercial use of biotechnology have estimated that in Germany it will reach around $2 billion by the year 2000 (compared to $1.25 billion in 1995), with a growth rate of up to 25 percent a year depending on specific products. For the state governments, such estimates serve mainly (*1*) to justify support measures for research and development in biotechnological sectors and (*2*) to provide hope against the prevalent impression that economic globalization will mean a growth of joblessness for Germany's economic system. Current estimates are that the biotechnological employment potential in Germany will be at well over 100,000 jobs in 2000. Commercial biotechnology alone is expected to account for up to 40,000 jobs (as compared to 19,000 in 1992), with up to 50,000 additional jobs in the supplier and services sector and another 20,000 jobs in universities or related to academic research. In 1999 roughly 500 firms with substantial or basic biotechnological profiles were counted in Germany (4).

The Federal Ministry for Education, Science, Research, and Technology (Bundesministerium für Bildung, Wissenschaft, Forschung, und Technologie, BMBF), has long supported the biotechnology industry in Germany. BMBF has consistently worked at improving the framework for biotechnological research and development in Germany by giving suitable incentives to industry and start-up companies, and by maintaining the quality of research institutes through state aid in appropriate core areas. In biotechnology the government has effectively assumed the role of a supervisor aiming systematically at creating productive economic alliances, among scientists and promoting a favorable political environment for such alliances. This policy was the focus of an important report (5) on future technology that was sponsored by BMBF and produced by the Karlsruhe-based Fraunhofer-Institut für Systemtechnik, a leading German institute for systems analysis.

One such remarkable initiative is the BioRegio competition (*BioRegio-Wettbewerb*) launched in 1995. This competition enables the best among the biotechnology-regions in Germany to come up with economic strategies to reach new goals. Incentives are given that focus on existing financing, funding, and investment potentials. Out of all proposals a jury selects the three regions with the most convincing biotechnology concepts. In the second phase these three regions receive priority in the appropriation of funds from the Biotechnology 2000 program of BMBF. Currently, the annual funds earmarked for the project are around $110 million. The BioRegio contest has become a way of establishing priorities that helps allocate BMBF funds for biotechnology. The total funding was about for the first phase starting in 1996 $0.6 billion, and this amounts to a clear signal of the priority given the development and implementation of new technologies in the biosciences and molecular medicine. The financial resources of BMBF are envisaged to support basic research, nonuniversity research institutions, and special federal government projects. In the first round, 17 German regions competed. In late 1996 a commendation was given to the Rhine-Neckar Triangel, Munich, and the Rhineland, and the Bioregion of Jena received a special commendation from the jury for its BioInstruments initiative. The BioRegio competition has fueled the regions' desire to be able to boast as many company start-ups as possible. It has cleverly used the traditional federal nature of Germany and the independence of the regions to initiate an upswing in biotechnology in many locations simultaneously. Moreover the project grants provided by the BMBF's BioFuture program (with an allotted $8 million for the year 2000) are, in principle, available to all applicants regardless of region.

Another development of the government's BioRegio contest was the creation in 1996 of a "virtual enterprise" located geographically in the triangle of the university cities of Braunschweig, Göttingen, and Hanover. This has become the largest self-contained research region for natural and engineering sciences in Germany. This virtual enterprise, called BioRegioN, is designed as a network of dense communication and extensive cooperation among more than one hundred scientific institutes, industry-related facilities, and administrative offices (5). A special consulting Web site for handling issues of financing and siting is available to entrepreneurs who want to start up new biotech firms. Support is available on legal matters concerning the use of genetially modified plants and permits for genetic engineering work as well as for marketing and project organization. Moreover BioRegioN offers research institutes and companies in the region the opportunity to make presentations at German and international trade fairs.

A 1995 study by the federal Minister for Education, Science, Research, and Technology revealed that almost half of the surveyed small and medium-sized companies saw a lack of equity as an essential impediment to innovation. Germany seems to have much ground to cover in catching up in the use of venture capital. In 1995 only 6 of the more than 100 capital investment companies were fully venture capital companies specializing in nascent technological businesses. Partly, this poor showing is due to the absence of large institutional investors, for example, pension funds, in Germany. Recently this problem has been debated as due to Germany's underdeveloped shareholders' culture. However, in 2000 it is becoming easier for biotechnological start-up firms to attract venture capital because a number of banks have set up special services in response to this conspicuous problem.

In the legal realm the federal government has been urging a more flexible legal framework for biotechnology in Germany. In accordance with European directives, the federal government has moved in favor of standardizing the safety and application regulations that open up scope for more efficient methods, with more flexible structure and organization while maintaining risk protection. Germany also pioneered in the implementation and commercialization of biotechnological research and development by an initiative in 1996 that facilitates patenting procedures for research findings. The federal government has further regulatory responsibility for the German contribution to the Human Genome Project.

Since the late 1980s expenditure on research and development (R&D) has declined in Germany in relation to

the gross domestic product (GDP). Between 1988 and 1994 R&D expenditure decreased from 2.88 to 2.33 percent. As a results there were considerable declines in research. In expenditure on R&D in the industrial sector declined even more, and public funds have not been able to compensate for the loss of funds. There has been much cause for political concern, particularly in view of the fact that Germany's industries depend heavily on R&D. Moreover, funding for R&D in the biotechnology sector has remained comparatively low in Germany in relation to corresponding activities in other European countries. As a result there is much concern about the efficiency and selectivity in the government's discourse on biotechnology. Funds and human capital need to be managed specifically in prominent future-oriented and pioneering aspects of R&D. The same applies to state aid. Correspondingly the bioindustrial sector has articulated concerns on how to improve the conditions for commercializing results of "pure" biological research in Germany.

Furthermore there are continual concerns about identifying the appropriate policies in order to correct institutional and financial inflexibility currently prevailing in research establishments. Increasingly academe, especially biology and chemistry departments, has had to confront criticism, voiced in the economic sector as well as by curricular commissions and higher education authorities, as to obsolete training and career structures, outdated conceptions of autonomous research ("freedom of science"), and lack of openness for teaching inter- and transdisciplinary skills and methods. An often-cited fact is that German universities have a share of less than 1 percent in biotechnological patent applications compared to about 15 percent in the United States.

To sum up, the record of the federal government's efforts reflects a fairly robust political majority consensus on creating an innovation-friendly framework for biotechnology. This policy encounters opposition for ideological reasons mainly from the "fundamentalist" wing within the Green politial party (DIE GRÜNEN). Apart from the somewhat special case of the Green party, what opposition there is to public policies in support of "biotechnological progress" is scattered across the other three major political parties and seems to be independent of party-line orientations. Owing to the majority political consensus, the government's budgetary preferences for R&D in the biosciences and biotechnology, despite falling overall rates for R&D funding, has not encountered serious political roadblocks.

Technology Assessment and Public Participation

In German law, administrative procedures for high-risk technologies generally involve an element of public participation. The 1990 Genetic Engineering Act, for example, made a public hearing (*Erörterungstermin*) a legal prerequisite for every license application to release genetically modified plants. The former public knowledge condition stipulated that the public notification of the project involve display of the submitted documents and provide for public opportunity to lodge objections and a public hearing. Anyone could make written objections, and these were then discussed at the hearing. In 1993 there were three hearings on applications for the release of genetically modified potatoes, sugar-beets, rape, and maize. The hearings each received between a few hundred and 20,000 written objections. The procedure, it turned out, frustrated both sides. For this reason, the provision of direct public participation was removed and substited by a much less demanding requirement for a written submission with the first amendment to the Genetic Engineering Act in 1994 (6).

Among the German people there is the feeling that industry and politics, and to some extent even the biological sciences, lack credibility, transparency, and democratic accountability in biotechnical matters. Neither concerted effort for citizen debate by the government nor various public relations campaigns by the biotech-industry have so far been able to dispell this attitude. Sociologists are apt to explain this as indeterminate suspicion caused by a knowledge gap. Nevertheless, no number of publications in every media form has translated into a corresponding degree of knowledge among the population. This knowledge gap itself requires more clarification. Political scientists concerned with technology assessment point to the fact that with certain exceptions, such as the citizens' forums in Baden-Württemberg (8) managed by one of Germany's leading institutes for technology asessment (Akademie für Technikfolgenabschätzung, Stuttgart), there has been too little open public discussion of the opportunities and risks of biotechnology. Pounding out information packages is one thing. Promoting attitudinal change through free and open public dialogue is quite a different thing. Government and industry have been strong on the former but weak on the latter communication policies.

Observations on cases where the general public was invited to take part in decision-related discussions (e.g., about sites for genetic engineering plants) indicate a number of recurring deficits and problems: (*1*) Those in positions of responsibility often fail to answer the questions that really concern critics and the public at large. This might be so because such questions tend to be of a fundamental nature. Frequently they concern consequences of technology on a broad cultural and societal scale. (*2*) Sometimes exaggerated expectations are aroused regarding the technical and economic potential of genetic engineering. The much-hyped gene therapies for cancer are a case in point regarding "red" (i.e., medically related) biotechnology. Other prominent examples are furnished by unrealistic pronouncements about expected positive impacts of biotechnology on employment and economic growth figures. (*3*) The depiction of biotechnology in the mass media is disturbingly polarized. Industry-financed advertising campaigns are apt to present progressive biotechnology as a safe panacea, whereas professional journalists writing in the mass media tend to capitalize on hypothetical risks of biotechnology (9,10). Consequently any balanced reporting that is both interesting and beyond obvious partisan interests is structurally disadvantaged and blotted out.

Between 1991 and 1993 a technology assessment (TA) of crop plants with genetically engineered herbicide resistance in agriculture was organized by the Wissenschaftszentrum Berlin, Germany's leading sociological research

center (11). Herbicide resistance was chosen as the subject for TA because it seemed to be sufficiently relevant and controversial. It could be expected that a broad spectrum of developmental problems of modern biotechnology would be considered in the process, such as (*1*) the possible risks of transgenic plants, (*2*) the toxicological and ecological effects of the use of nonselective herbicides, (*3*) the future of genetic resources, (*4*) the advantages and disadvantages for farming, (*5*) the long-term safeguarding of world food supplies, and (*6*) the ethics of plant manipulation.

This pioneering TA project was not merely a forum of experts for evaluating the state of knowledge on the possible consequences of a technology; rather, it provided an arena in which the social conflicts related to the introduction of a new technology could be articulated and discussed in an exemplary manner. The procedural scheme was somewhat conventional in taking an emerging technology-induced development as its starting point for an analysis of possible desirable and undesirable consequences. The goal of technology-induced TA is to determine the political actions that might be necessary in order to cope with that technology. Critics of the Berlin herbicide resistance TA project called instead for a "problem-induced" approach. The starting point would shift to the social problem the technology purports to help solve (e.g., the agricultural problem of weed control). In problem-induced TA, various ways of tackling the problem would be compared (e.g., solutions by intensive industrial farming in comparison to solutions provided by ecological farming). Any comparisons would take questions of larger social context and fundamental political issues into account. The decisive question for transgenic herbicide resistance technology when considered from a "problem-induced" TA-perspective would have been whether the technology was needed and what kind of farming was socially desirable and ecologically acceptable. Problem-induced TA was found to be more demanding in terms of time, money, and other resources than the technology-induced approach. In retrospect, the Berlin project aimed at a broad and demanding interpretation of TA with resources appropriate only for dealing with narrow TA.

Participants from research institutions, industry, environmental groups and other NGOs, and governmental agencies (totaling 48 groups) reflected the interests and positions of the ongoing political concerns over biotechnology. Among them were outspoken advocates and critics. The debates that normally take place outside the TA procedure and only heat up as soon as results of "closed" TA are made public, were internalized in the procedure. TA became a social dialogue between representatives of opposing positions. After a series of meeting, the participants had to define a study framework, evaluate the results of expert reports, and discuss any conclusions. The idea was that a dialogical framework would promote a rational form of discussion. Whether this expectation was fulfilled remains an open question. As a matter of fact, shortly before the final meeting was to take place, the environmental associations announced their withdrawal.

The Berlin TA project reenergized the scientific and public debate in Germany on the proper form, scope, and aim of assessment of biotechnology and other new technologies. There is now emerging a consensus that a narrow TA is necessary but not sufficient for the democratic governance of technological development to florish. The narrow TA is essentially an investigatory strategy aiming at the production of information where the validity of such information is conditioned by truth-claims and not by the factual acceptance by a majority of participants. The narrow TA is not to be conflated with purely political dialogues or consensus conferences where discussion focuses on goals and criteria of desirable development for society. The narrow TA contributes factual information about potential risks and expected advantages. Citizens have a right to identify the state of knowledge on politically controversial subjects. The TA procedures can at least advance answers as to whether publicly declared risks actually exist, whether claimed advantages exist, and so forth. Criticism of the methods, scientific and otherwise, through which the procedure arrives at its conclusions must also be submitted to the scrutiny of the public. Consequently technology-induced assessments, though narrow, must not be so narrow to involve only scientific experts. All stakeholders are to be considered in any political conflict over a new technology for a rational assessement of factual and other validity-claiming and procedural fairness.

LEGAL ASPECTS: SALIENT STRUCTURES IN THE REGULATION OF BIOTECHNOLOGY

Germany's Genetic Engineering Act (Gentechnikgesetz, GenTG) adopted in 1990 seeks to regulate the approval and registration procedures for genetic engineering facilities and for genetic engineering work geared to research and commercial purposes on microorganisms (viruses, bacteria, fungi, parasites) and macroorganisms (plants and animals). The GenTG does not cover reproductive medicine nor the use of somatic gene therapy.

The GenTG, in its amended 1993 version (12), is intended to safeguard the life and health of humans, animals, and plants; to protect the environment as an integrated system; to shield property from possible risks of genetic engineering methods and products; to prevent any such risks from emerging, and to create a legal framework for research, development, promotion and use of the scientific, technical and commercial possibilities of genetic engineering.

The GenTG legally defines a consensus-building processes within a network of agencies that share decisional power, political responsibility, and scientific competence concerning the issues that arise within the scope of the GenTG's regulatory framework. The network defined by the GenTG comprises four principal actors on the national level: (*1*) the Biological Federal Institute for Agriculture and Forestry (Biologische Bundesanstalt für Land- und Forstwirtschaft) that is situated within the Federal Minstry of Food, Agriculture and Forestry, (*2*) the Federal Environmental Protection Agency (Umweltbundesamt), in cases of animal applications, (*3*) the Federal Research Center for Virus Diseases in Animals (Bundesforschungsanstalt für Viruserkrankungen der Tiere).

In accordance with relevant European Community Directives, Germany's GenTG determines that a scientific body, namely (*4*) the Robert Koch–Institute (RKI) in Berlin, integrate the relevant consensus-building processes into resulting decisions. The RKI's Department of Genetics and Genetic Engineering serves as the institutional base of special advisory committee set up by the RKI, the Central Advisory Committe for Biological Safety (Zentrale Kommission für Biologische Sicherheit). This committee comprises a broad range of scientific and socially relevant points of view. Its 30 expert members come from micro- and cell biology, genetics, hygiene, virology, ecology, the trade unions, occupational safety, economics, research-promoting organizations and environmental protection organizations.

European Community Law and German Law

On the national level of law, the German GenTG implements directives of European law. Mainly three directives of the European Council are relevant: (*1*) Directive 90/219/EEC (13) stipulates joint measures to be implemented for the application of genetically engineered microorganisms in contained situations in order to safeguard human health and the environment. (*2*) Directive 90/220/EEC (14) on the intentional environmental release of genetically modified organisms (GMOs) serves both to promote the formation of the single European market for methods and products of genetic engineering and to constrain this process by the observance of suitable environmental and health protective considerations. (*3*) Directive 90/679/EEC (15) contains minimum requirements for the protection of employees in countries of the European Community against exposure to safety and health hazards caused by biological substances.

Unlike Germany's GenTG, the European Community directives include no special liability provisions. Approvals based on the directives focus on the genetic engineering product and not on the genetic engineering facilities, whereas the German GenTG prescribes extreme hazard liability for the operator of the facility concerned. In the European Community directives, hearings are only an optional feature. In the German GenTG, public hearings remain a legal requirement in approval procedures for operating sites, notwithstanding that the 1993 amendment of the GenTG narrows this requirement to commercial projects classified as "risky" or "very risky." (There are four legally recognized safety levels ranging from 0, no risk, to 4, very risky.) Despite these comparatively more demanding national requirements, the GenTG in its 1993 amendment fully exploits the scope for simplification and acceleration of approval procedures contained in the directives on the legal level of the European Union (EU).

In specific, the legal framework that regulates genetic engineering in Germany contains the following legal ordinances:

1. Genetic Engineering Safety Ordinance (*Gentechnik-Sicherheitsgesetz*) (16). This Ordinance is the most important legal ordinance. It consists of safety requirements for genetic engineering work, risk assessment of organisms, and safety measures in laboratory and production facilities, greenhouses, and animal facilities. Also covered are occupational safety, and sewage and water treatment, requirements that have to be met by project managers and biological safety engineers.
2. The amended Genetic Engineering Hearings Ordinance (*Gentechnik-Anhörungsverordnung*) (17). This Ordinance restricts public hearings to approval procedures for safety levels 3 and 4 of commercial projects. It regulates the formalities of hearings. It allows simplifications and modifications that favor the operator while still serving the purpose of protection.
3. Amended Genetic Engineering Records Ordinance (*Gentechnik-Aufzeichnungsgesetz*) (18). This amendment specifies the content, structures and time period for genetic engineering records. The amended Ordinance for the first time included environmental release in the range of requirements. The amendment introduced simplifications for safety level 1 projects. The safekeeping periods prescribed for records amount to 10 years for safety level 1 and 30 years for levels 2 to 4 and for environmental release. Laboratory logbooks are not an admissible substitute for proper records.
4. Amended Ordinance on Genetic Engineering Procedure (*Gentechnik-Verfahrensverordnung*) (19). This amendment regulates the formal documentation requirements to be met by facility operators in the registration and approval procedures for genetic engineering facilities, genetic engineering work, and environmental release and introduction of GMOs. The amendment has again enabled simplifications, clarifications, and accelerations. A distinction is made between genetically engineered "minor" and "major" plants.
5. Amended Ordinance on the Central Commision for Biological Safety (*Verordnung über die Zentrale Kommission für die biologische Sicherheit*, ZKBS) (20). This amendment describes the tasks, capabilities, and internal structure of the ZKBS.
6. Federal Cost Ordinance to the Genetic Engineering Act (*Bundeskostenverordnung zum Gentechnikgesetz*) of 1991. This Ordinance stipulates the fees and charges for offical acts performed by the Robert Koch–Institute as a senior federal authority. Fees for granting approval for environmental release range from $3,000 to $80,000, and fees for obtaining approval of introduction into circulation from $4,000 up to $170,000 for extremely expensive procedures.
7. Genetic Engineering Participation Ordinance (*Gentechnik-Beteiligungsverordnung*) (21). This Ordinance regulates participation of the European Council, the European Commission, and the authorities of the member states of the EU and those of the European Economic Area in the approval procedures for environmental release and introduction as well as in procedures for any supplementary measures based on the German Genetic Engineering Act. It is the

legal basis of the Robert Koch–Institute's obligation to observe set time limits in transferring received applications (for environmental release or introduction) to the European Commission. It also details the institute's right to information if such applications are submitted in other EU member states.

Intellectual Property Rights and Patenting

Industrial property rights are protected in Germany by the 1980 Patent Act (*Patentgesetz*, PatG) (22,23), by the 1973 Agreement on Granting European Patents (24) and the Species Protection Act (*Sortenschutzgesetz*) (25). PatG provides comprehensive regulation of all aspects regarding the patented item, such as patenting preconditions, effects of the patent, patent-granting procedures, revocation proceedings, patent infringements and their consequences. PatG is a major economic and political instrument for promoting innovation. The granting of exclusive rights over a limited time for the exploitation of inventions provides incentives. Inventors are rewarded, early revelation and dissemination of technical findings and know-how are safeguarded.

The German PatG, in keeping with European agreements, excludes from patenting any method invented exclusively for surgically or otherwise treating or diagnosing maladies of the human or animal body. Gene therapies, according to this criterion, are not eligible for patenting, whereas biopharmaceutical products produced by gene therapies are eligible. Ruled out from patentability are discoveries that are not inventions. By this criterion, nucleic acids, and more or less complete indications of sequences of nucleic acids, do not qualify for patent protection if the availability of the substance is merely based on a discovery. Likewise the PatG excludes "essentially biological" methods for breeding plants or animals, or creating new plant or animal species. Microbiological methods and their products, however, are eligible for patenting. Until recently this implied that there was no patent protection at all for generic creations like transgenic plants and animals. In comparison to the United States and Japan, this put the German bioindustry at serious disadvantage with regard to plant and animal biotechnological creations.

Things have changed somewhat with a recent amendment to PatG. The new rules of the General Agreement on Tariffs and Trade (GATT) took effect in Germany in 1995. All member countries must now modify their national intellectual property right laws so that they match with the new agreement on trade-related intellectual property rights (TRIPs). Since the issue of TRIPs and their discretionary leeways continue to be notoriously controversial, the GATT negotiations stipulated a review period in 1999 for the enforcement of the agreement on TRIPs. In September 1999 the GATT–TRIPs agreement was accepted by the European Patenting Organization in the format of the European Biopatents Directive, which in turn was enacted by the European Patent Office in Munich. Whereas former article 53b of the European Patent Agreement clearly ruled out the patenting of plant varieties and animal species, the recent Biopatents Directive allows for a much more patent-friendly legal interpretation on these crucial points. However, the acceptance of the Biopatents Directive in Germany and the EU generally is so far only provisional, pending a ruling by the Supreme Chamber of the European Patent Office on a number of complaints about would-be patented genetically modified plants and, moreover, pending the European Supreme Court's decision on complaints against "the patenting of life" that have been filed by Italy and the Netherlands.

In the issue of deregulating European (and hence also German) patenting restrictions the parliamentary assembly of the European Council (*Europarat*) in Strassbourg is pitted against the bureaucracy of the European Commission (*Europäische Kommission*) in Brussels. In September 1999 the former organization, representing 42 European states, has recommended that patenting of genes, cells, tissue, and organs be prohibited altogether, whether they be of human beings, animals, or plants (the key argument is that these entities are discoverable but not inventable). Interestingly the former organization is perceived in the political public sphere as having far more democratic standing than the latter. This makes it easier for biotechnology critics to capitalize on the issue of patenting by interpreting it as an antagonism between a democratically incorporated humanistic ethos and the organized commercial interests of business corporations and their lobbies.

Consumer Sovereignty and "Novel Food"

Big agro- and food companies began promoting in the second half of the 1990s the prospect that within a few years a new generation of genetically modified novel food products would attract consumers through the appeal of both superior quality (longer lasting and tastier fruits and vegetables) and lower prices as compared to natural food. Commercially, this confident vision continues to prevail.

Public outrage at biotechnologically produced or modified "novel" food is reflected in German food legislation. This legislation took the form of the Food and Consumer Act (*Lebensmittel- und Bedarfsgegenständeverordnung*), numerous follow-up ordinances, and regulations pertaining to specific products (e.g., milk legislation), plus directly applicable EU law. The aim of this legal framework is to safeguard human health and to protect the population against fraud and deception. All provisions apply equally to standard products and products manufactured by genetic engineering methods. Specific differences, however, remain. The GenTG, with its special regulations on introducing products into circulation, refers to food products or food ingredients that contain or consist of GMOs. However, these regulations do not apply to food products or food ingredients that are manufactured from GMOs without containing any such organism. These foods remain subject to the "general food law."

Thus regulatory uncertainty exists about the proper implementation of the EU novel food regulation (NFR). As far as genetic engineering is concerned, NFR refers to (*1*) introducing novel food products and novel food ingredients into circulation within the EU; (*2*) food products and ingredients that are "novel" in the sense that they were previously not widely used for human consumption in member states; (*3*) products and ingredients that contain

or consist of GMOs, or (*4*) products and ingredients that were manufactured from but do not contain GMOs.

Exempt from the regulation are food additives, flavorings and extraction solvents. They are subject to other special EU regulations. This exemption has recently become the focus of much public concern. In one interpretation at least, it creates a gap in the generally desirable principal goal of NFR, namely to ensure that novel food products and food ingredients do not pose a hazard for the consumer. According to NFR, approval decisions for novel food check for environmental compatability in order to prevent any harmful effects of GMOs on the nonhuman environment. In addition the Scientific Food Committee of the EU must give an opinion on all aspects of food that relate to public health.

One considered hazard is that novel food differs so greatly from comparable products and ingredients it intends to replace that normal consumption patterns would result in nutritional deficiency in people. However, the hazardous nutritional deficiency argument, though frequently marshaled by ecological consumer organizations and other critics of "Frankenfood," fails concerning mere additives, flavorings, and the like. Another critical argument — call it the consumer sovereignty argument — has better logic. It has also drawn more public support. It is a second explicit goal of NFR to ensure that nutritional improvements (biotechnological or otherwise) do not lead to consumer confusion. This goal is subverted, according to the consumer sovereignty argument, if consumers who (for whatever reasons) repudiate novel food completely lack the necessary information for making fully informed choices in following their convictions.

The appeal of the consumer sovereignty argument becomes lost if companies begin to discharge their requirement for full consumer information in a way that turns labels for food products into something like the patient information leaflets that accompany prescription drugs. Legally, how much leeway biotech-food companies have for their labelling policy is defined by conditions that are already quite demanding: (*1*) Food products containing or consisting of GMOs must be labelled under all circumstances. (*2*) They must also be labelled if the use of genetic engineering methods means that they are no longer equivalent to an existing food product or ingredient. "Not equivalent" means that a scientific assessment based on appropriate analysis of existing data indicates that the tested features differ from conventional food products or ingredients. Accepted limited values for natural fluctuations in these features are to be taken into account. If a product or ingredient is no longer seen as equivalent, the label must indicate the changed features or characteristics along with the method used to bring about this change (so-called method labeling). (*3*) The label must also indicate substances that are not present in existing equivalent food products and that could affect people with certain health problems like allergies or that could foreseeably violate ethical or religious dietary restrictions.

The preamble of NFR states that a label may declare that a food or ingredient is not a novel product in the sense of the regulation. A label "free of genetic engineering" or "without genetic engineering" is allowed, pending proper definition. Germany submitted a proposal to the EU in 1998 for defining this label, and still is awaiting approval. The proposal aims to require that labels state whether raw materials from transgenic plants, enzymes, or additives for flavoring, and the like, were derived from GMOs or used in food production. This would include animal feed and feed additives derived from transgenic organisms, even if the end product (e.g., milk from cows fed on transgenic soy) is chemically indistinguishable from the product of animals fed conventionally. What is decisive is that the labelled products must be garanteed not to have been in contact with genetic engineering. To give an example, though genetically modified barley, transgenic hops, or genetically modified yeast are not commercially available in Germany, most beers cannot be labelled "without genetic engineering" because conventional yeasts are fed GMOs during their fermentative reproduction.

ETHICAL ASPECTS: CATEGORIES, ASSESSMENTS, AND CONTROVERSIES

Genetic engineering and the underlying molecular biological research is viewed, in Germany as elsewhere, as cutting-edge biotechnology. In the public's mind genetic biotechnology comes in three colors: It is red if its application is mainly medically related (in the diagnosis and treatment of diseases), green if mainly related to agriculture and food, and gray if used for purposes of environmental conservation and the protection of natural resources, ranging from pollutant disposal to environment-friendly bio-mining processes.

If a 1995 survey of public opinion is any indication, the percentages of positive (+), neutral, and negative (−) attitudes toward certain applications of biotechnology are nearly the converse of each other in terms of the red and green ends of the spectrum of applications. The most positive results were for the diagnosis of incurable diseases, 75+ and 7− out of 100; negative attitudes increased for treatment of cell diseases, production of vaccines, use of modified bacteria to reduce oil pollution in the soil, modification of crop resistance against insects or plant diseases, breeding laboratory animals for pharmaceutical research. The most negative result was for the modification of the taste, keeping quality or appearance of food, 8+ and 76− (26,27). The basis for these attitudes is probably a mix of conceptions. Clearly, the big differences in the attitudes are to some extent due to the fact that the different practices to which biotechnology is applied are loaded with different ethical background beliefs. Biotechnology and genetic engineering in particular continues to be perceived in public opinion as something special resisting easy assimilation to "normal" new technologies (e.g., energy, information, and communication technologies). The new biotechnology is mysterious and evokes a pandora's box. The negative attitude persists despite an increasing evidence that genetic engineering biotechnology keeps dispelling all initially raised claims of associated high and unusual risks compared to other large-scale new technologies. Public risk perception, even if well informed, and moral public opinion are surprisingly independent of each other.

Recent analyses (28) of public debate about and public perception of genetic engineering in Germany point to the fact that it would be unrealistic to expect broad acceptance of such biotechnology to be achieved by media campaigns, improved distribution of information, inclusion in school-curricula, and the like. All that can be realistically expected is that a consensus on the acceptability of a limited number of specific applications where the application is, and is perceived to be, juridically regulated and politically controlled in ways that ensure that relevant ethical convictions (whether rooted in rational and secular views or in religious comitments) that are held by substantial groups of citizens are being respected.

In reaction to the strategically insurmountable momentum of ethical beliefs concerning biotechnology, the European Association for Bioindustries (EuropaBio), a business interest nongonvermental organization in Europe representing 45 multinational corporate members and 4 national associations that total around 600 small and medium-sized biotechnology enterprises, created an Advisory Group on Ethics (AGE) in 1997. The first task set for AGE was the crafting of a document on *Core Ethical Values*. This code of ethics has achieved adherence among EuropaBio members independent of their respective national legislation, and it contains commitments to ethical values and ethically important goals in all major fields of biotechnological applications. It is less specific and less stringent than the *Convention on Human Rights and Biomedicine* (29) that was crafted by the Council of Europe. The convention was officially adopted in 1997 and already ratified by six member states, though not by Germany. At present the hub of distinctively ethical normative regulation of biotechnology in Germany is informal in nature (by nonmandatory codes of ethics) rather than formal (by actual law).

The use of philosophical arguments of applied ethics in Germany is still very tentative. Four well established centers dominate the research: The Academy for Ethics in Medicine located in Göttingen, the Center for Ethics in the Sciences in Tübingen, the Center for Ethics and Science in Bonn, and the European Academy for the Study of Technology in Bad Neuen-Ahrweiler (30). Each of these major centers has invested a considerable part of its capacity into biotechnological research. The growing literature on bioethics generally in Germany is replete with attempts to appropriate arguments that have already come to enjoy some currency in contexts of applied ethics in England or the United States. This fact could attest to a level of similarity in the ethical problems raised by global progress in biotechnology independent of the cultural contexts. Such similarity could also reflect a level of homogeneous academic discourse layering itself over heterogeneous cultural and moral inclinations at an intuitive level. At the level of theory, it is noteworthy that one of the above-mentioned centers, the Center for Ethics in the Sciences, addresses social ethics. Unlike prevalent notions of principlism (e.g., the bioethics principles addressed at the Kennedy Institute in the United States), a social ethics approach emphasizes all reliable moral considerations that have already been worked out in cultural traditions and affect the target practices. Moreover a social ethics approach works with what is, technically speaking, a "wide reflective equilibrium methodology" and takes into account a wide spectrum of religious, juridico-legal, and other relevant normative textures and theories in modern society, modern law, technological progress, risk asessment, and so on. As the qualifier "social" indicates, a social ethics approach has its integrative frame in taking seriously the impact that the foreseeable social change is likely to make a on moral conceptions that shape a society (30).

Gray, Green and Red

Gray Biotechnology. Although gray biotechnology is frequently hailed by the government as a dynamic and commercially rewarding field (32), there is at present hardly any ethical controversy about gray biotechnology. Speculating about the deeper reasons for the comparative moral inconspicuousness of gray biotechnology could be interesting, but it is safe to say is that a good reason is that no dramatic accident has occurred up to now. Absent something like gray biotechnology's Chernobyl, public attention will continue to focus on the more salient perplexities in red and green biotechnological developments (33). A recent development in public debate is an attempt to add ethical merit to gray biotechnology by pointing out its potential for political strategies of sustainable development. Such strategies enjoy much political enthusiasm in Germany and they have large support among the public. Interpreted optimistically, sustainable development presents itself as a viable response to most of today's pressing environmental problems.

Green Biotechnology. In public debate the following points provide most of the standard argument in favor of green biotechnology: Because of global population growth, and limited agriculturally utilizable terrain, the food problems of future generations must be solved by an environmentally friendly increase in yield and by the production of high-yield, low-cost food by means other than conventional methods in agriculture and in the foodstuff industry with new biotechnological methods and products. This argument is invoked for justifying genetic engineering based strategies for maximizing agro-industrial turnover as well as for associated strategies for minimizing the deployment of herbizides and insectizides in combination with suitably modified crops.

Despite efforts to provide an ethical underpinning for green biotechnology by linking it to moral responsibility for future generations, the vision of green biotechnology has not fared well in Germany (34). Many people have reservations or even reject its use in agriculture and in the food industry. Surveys indicate that the public sees no adequate benefits that could compensate for existing negative expectations and anxieties. For the individual consumer, it is obviously of major significance that products of green biotechnology have concrete and palpable advantages. By contrast, most advantages of biotechnologically "improved" food that have been touted by the biotechnology industry do not refer directly enough to the final product. Public relations campaigns designed

to swing consumer attitudes in favor of bio-food are widely recognized as poorly concealing the fact that the would-be advantages are essentially commercial advantages on the side of bio- and agro-industrial entrepreneurs, and not on the side of the Consumer. At the moment, consumers do not perceive the show-cased advantages (e.g., the Flavrsavor tomato) as personally benefiting. Even tangible product improvements like longer shelf lives for fruit and vegetables do not automatically result in greater consumer appreciation in the face of the ample range of food products offered in Germany and in advanced capitalist countries generally. Also there is an as yet small, but in economic terms increasingly important, segment of the population that cherishes "natural" food products notwithstanding its comparative cost disadvantage.

General knowledge about green (genetic) biotechnology, especially about food stuff production, is not very good on the average in the population. Attitudes toward food are frequently characterized by culturally validated consumption patterns and a good deal of wishful thinking concerning the quality and the production process of preferred food. These attitudes are not shaped by rational choices in terms of cost–benefit calculations in everyday shopping. This, the green biotech-industry has been very slow in realizing. Many German consumers are afraid not only that gene technology is creeping surreptitiously into their food without their consent, but also that they are being misused as guinea pigs by agro-industrial corporations. Highly sensitive analytical techniques (e.g., PCR) can detect even minute quantities of transgenic ingredients, like soybeans used in the production of soya oil. Soya products are used in thousands of different foods. The fact, much publicized in the media, that Monsanto's "Roundup Ready" genes can now be detected even in tofu made from soya that was grown on controlled ecological farms engenders in consumers an insidious feeling of powerlessness. On the other hand, with the attractiveness of modern lifestyle preferences mounting, more people are becoming interested in consuming "functional" (e.g., high energy) foods. It will be interesting to observe how prevailing patterns of resistance to green biotechnology in Germany will change when vitamin enhanced vegetables and fruits become available in the supermarkets (35).

Red Biotechnology. Advocacy for red biotechnology in Germany mainly follows the conventional argument of medical utility: "Our rapid increase in understanding molecular genomics and genetic engineering has brought us to the brink of a health care revolution. Biotechnology brings the tools (gene therapy, recombinant proteins, and cellular therapies) for not only curing common diseases but also many rare diseases." The medical promises of biotechnology have materialized so far mainly in the pharmaceutical sector. For instance, in 2000 there were 42 (compared to 29 in 1998) medical drugs with genetically engineered components fully licensed in Germany (36).

Unsurprisingly, genetic engineering technologies and diagnostic techniques based on molecular genetics are generating most of the moral perplexity found in red biotechnology. In Germany as everywhere, great anxiety is being raised about the benefits of somatic cell gene therapy. Any original enthusiasm was dampened by the medical profession failing to deliver on the goods that had been promised by biotechnological advances. Ironically, somatic gene therapy has turned out to be technically more intractable than first thought but ethically less perplexing. At first arguments likening somatic gene therapy to other forms of transplant therapy won the day in the German debate. A trail-blazing government report on the chances and risks of gene-technology concluded as early as 1987 that "somatic gene therapy is a special form of transplant therapy" and that the "transfer of genes must be evaluated in the same way as the transfer of living material" (37). The German Association of Physicians (*Bundesärztekammer*) followed with similar policy statements in 1989. If successful therapies were available, many people would view the prospect of somatic gene therapy as less problematic than organ transplants.

Germ line gene therapy is ethically taboo in Germany, and it is legally prohibited by the 1990 Embryo Protection Act (*Embryonenschutzgesetz*) (38,39). This act prohibits (*1*) the sale, use, or acquisition of in vitro fertilized eggs for all purposes other than for an intended pregnancy. It also prohibits (*2*) the generation of more than three embryos per cycle in in vitro fertilization (IVF) procedures, (*3*) the selection of sex, (*4*) fertilization involving gametes from dead persons, (*5*) deliberate altering of the genomic information of gametes intended for procreation, (*6*) embryo cloning, and (*7*) the production of animal-human chimeras. These restrictions (with the exception of 2) in Germany's national law overlap more or less with relevant articles of the aforementioned European Convention on Bioethics (Articles 13, 14, and 18). However, the Convention has not yet been, and probably will never be, juridically implemented in Germany. The Convention has been politically attacked by relevant groups of stakeholders (e.g., by organizations representing the interests of disabled people) as being unduely liberal especially in matters foremost concerning research on subjects incapable of giving informed consent. The ethical arguments against germ-line gene therapy, cloning, and enhancement genetics that have gained most currency in Germany are (*1*) the violation of human dignity in identity-altering genomic manipulations and thus the integrity of human beings, and (*2*) the slippery slope of socially amplified genetic discrimination which recalls the outragious practices of Third Reich "eugenics." The metamorphosis of "healing" into "enhancement" and the interpretation of "eugenics" as "breeding" easily evokes the image of human beings treated as cattle. Many people in Germany associate eugenics not only with mass murder policies by totalitarian states but also with sexism and a patriarchal attitude toward women (40,41). This dramatic subtext of eugenics in Germany tends to distract from the disquieting fact that socially conditioned value-of-life judgments and selective decisions enter often into the "private" realm of parental responsibility as more and more prenatal diagnostics becomes a matter of course in "normal" pregancies.

The German Research Association (*Deutsche Forschungsgemeinschaft*, DFG) cooperates with the BMBF in managing Germany's part within the global Human

Genome Project. The German Human Genome Project (DHGP), established in 1995, has a budget of about $30 million a year (42). Integral to DHGP is a program for the study of ethical, legal, and social issues. Between 1996 and 1999, during the first phase of this ELSI-program, seven interdisciplinary conferences were sponsored with a budget totaling about $0.3 million. Theses conferences ran the gamut of interesting topics, but very little overarching direction was achieved besides a set of general topics that function as preferences when new project grants are offered. Conferences were held on (*1*) Genetic Knowledge and Human Self-Understanding, (*2*) Talking Human Genetics: Verbal Communication, Knowledge and Genetic Make-Up (43), (*3*) The New Genetics: From Research into Health Care. Social and Ethical Implications for Users and Providers (44), (*4*) Predictive Genetic Tests, (*5*) Ethical and Legal Problems in the Patenting of Genetic Information, (*6*) Postgenomics? Historical, Techno-Epistemic and Cultural Aspects of Genome Projects, (*7*) The Human Analyzed (45). For the second phase of the DHGP's ELSI-program, these preferences are (*1*) ethical, legal, and social aspects of human genome research in practice (e.g., informed consent, protection of privacy for genome-related information, questions of patenting and of the commercialization of research), (*2*) the application of genetic testing (e.g., quality of genetic counseling and diagnotics, genetic testing beyond the confines of medical human genetics, aspects of health care economics), and (*3*) the cognitive apraisal and the social context of human genome research (e.g., social and cultural differences in the handling of genetic knowledge, impact of genetic knowledge on the concepts of prevention and malady, and sociological metaresearch on the human genome project itself).

The DHGP has not raised much protest in comparison with other contested technologies in Germany (like nuclear energy). What moral outcry there is applies indiscriminately to the global Human Genome Project. Prominent are arguments on the subversion of human dignity in critical debates about bioethics, and this is perhaps more pronounced in Germany than else where. There is a mixed coalition of small groups of diverse political orientation that oppose any research with a potential for furthering prenatal selection of human traits and other forms of what they perceive as genetic discrimination. Some of these groups are radically suspicious of the integrity of bioethics as a scholarly discipline (46). Besides such insignificant "fundamentalist" opposition against biotechnology, major nongovernmental organizations, such as *Gen-ethisches Netzwerk*, have developed a highly differentiated dissentious culture of counterexpertise and civic mobilization, and brought forth excellent journals (e.g., *Gen-ethischer Informationsdienst*, *Wechselwirkung*, *Politische Ökologie*) for critically monitoring developments in biotechnology (47).

Arguments Against "Patenting Life"

Patents, trademarks, and copyrights, are forms or intellectual property protection. The practical point of such protection is to ensure, for a certain period of time, that the individual or a corporation that rightfully claims to have invented that product or technology maintains an exclusive right to make, use, and sell a new product or technology. Any intellectual property rights regime can be construed as a purportedly legitimate compromise between industry's desire to capitalize on its investments in technological development and the prima facie justified claim of society to benefit from the knowledge and resources of its members in a terrain of activities that are only made possible by that society. However, embedded in the normative texture of the (amended) German patent law is a moral component that goes beyond those moral considerations of commutative justice that are pertinent to societally useful inventions and investments.

This moral amendment can best be understood by considering which exclusions from patentability of biotechnological inventions are claimed specifically on ethical grounds. As it turns out, not patentable in Germany (and generally in those EU member states that have implemented Article 6 of the Directive 98/44/EC) are processes (*1*) for cloning human beings, (*2*) for modifying human germ-line genetic identity; (*3*) for use of embryos in industrial or commercial capacities; (*4*) for modifying the genetic identity of animals that are likely to cause them suffering without any substantial benefit to human or animal, and (*5*) for creating new animals from genetic modifications.

The popular claim against permitting the patenting of genetic sequences in living organisms ("no patents on life") makes the following key arguments:

1. Patenting would blur the conceptual distinction between discovery and invention for the sake of vested private interests. It upsets the carefully established balance, controlled by that distinction, between a monopolistic commercial privilege and an associated benefit for the common good.
2. Patenting would regionally and globally threaten genetic diversity. For pharmaceutical, food, and seed companies, and the biotechnology firms behind them, the ability to scan, pick from, and patent the world's biological diversity harbors prospects of great new sources of revenue. But the emphasis on finding and isolating plants and other living matter with the most marketable traits leads to the decline of other plant species, since only the cultivation of those species that are required for the creation of new varieties becomes constantly reenforced. Tailoring property rights to the privatization of genetic resources that have been engineered and patented also promote crop monocultures. A study by the German Parliament's Office of Technology Assessment concludes that in comparison to conventional methods, no substantial risk of loss of biodiversity obtains for the use of genetic biotechnology in plant and crop cultivation (48).
3. Patenting of genetically modified seeds would cause farmers individually and in developing countries to incur extreme economic strain. The economic

incentive of royalties set by intellectual property rights would benefit the technically advanced countries (being the principal producers of economically valued biotechnologically modified seeds and crops).

4. Patenting would encourage "biocolonialism" and "biopiracy." Genetic resources should be treated as a common heritage of humankind, with the moral implication that any commercial use of a genetic resource whose origin can be determined to belong to a certain country must also benefit the people living in that country. (A recent example is the Indian Neem tree and its pharmaceutical exploitation by an English medical drugs corporation.) This argument overlaps with argument 2 in that the ability of business companies to gain monopolies over what were formerly freely available community resources (seeds, plants, and even microorganisms) is assumed to have devastating effects on both human communities and the protection of biodiversity.
5. Patenting would in the long run slow down the development of medical drugs; similarly it would stifle the advancement of very successful conventional techniques of animal breeding and plant cultivation.

In order to underscore the ethical meaning of these five issues, they can be summarized as concern over (*1*) greed, (*2*) loss of biodiversity, (*3*) exploitative commercialization, (*4*) biotechnological colonialism, and (*5*) misplaced utility. There is one further argument that is harder to categorize because it taps into a more general resentment against economic globalization. This is that (6) permissive biotechnological patenting engenders an ethically undesirable shift away from treating people as bearers of human dignity toward an image of people as biological material. This attitude is a possible outcome of colonialization of the nature by economic interests that turn genetic and cellular materials and human (and nonhuman) organisms into potential sources of revenue.

Interestingly, the last argument is not confined in its scope to genetic engineering. It can be applied to patenting in many areas of biotechnology. The argument comprises complex cultural consequences. The idea of life being something not of human making, or something beyond what human ingenuity may try to control, is deeply engrained in European culture. The outcry against patenting life captures a concern that, theologically speaking, scientists might be attempting to play God. Biotechnological patenting would give them license to play God. Yet this resistance to patenting could also be spelled out in less elevated terms. The more commonsense attitude is that whatever can be done is going to be done. This attitude which used to be a flippant remark on uncontrollable technical progress is now gradually being displaced by the realization that whatever pays off is going to be done. The anxiety over biotechnology patenting bespeaks widespread fears over uncontrollable market forces unleashed by economic globalization.

BIBLIOGRAPHY

1. Der Bundesminister für Forschung und Technologie, *Biotechnologie. Programm der Bundesregierung*, Bonn, 1985.
2. U. Dolata, *Politische Ökonomie der Gentechnik. Konzernstrategien, Forschungsprogramme, Technologiewettläufe*, Sigma Verlag, Berlin, 1996.
3. G. Vita, *Internationale Politik* **8**, 7–14 (1998).
4. Schitag Ernst & Young, *Germany's Biotechnology Takes Off in 1998*, Stuttgart, 1998. Survey available at: *www.dechema.de/deutsch/isb/firmen/sme.htm*
5. Available at: *www.gbf-braunschweig.de/bioregio/home.html*
6. H. Grupp, ed., *Technologie am Beginn des 21. Jahrhunderts*, Physica Verlag, Heidelberg, 1993.
7. A. Bora, *Law Policy* **20**(1), 113–133 (1998).
8. T. von Schell and H. Mohr, eds., *Biotechnologie — Gentechnik. Eine Chance für neue Industrien*, Springer, Heidelberg, 1995.
9. H.M. Keplinger, in P. Stadler and G. Kreysa, eds., *Potentiale und Grenzen der Konsensfindung zu Bio- und Gentechnik: vom 11. bis 14. März 1996 in der Evangelischen Akademie Schloss Tutzing am Starnberger See*, DECHEMA, Frankfurt, 1997, pp. 15–50.
10. K. Menrad, in P. Stadler and G. Kreysa, eds., *Potential und Grenzen der Konsensfindung zu Bio- und Gentechnik: vom 11. bis 14. M,,rz 1996 in der Evangelischen Akademie Schloá Tutzing am Starnberger See*, DECHEMA, Frankfurt, 1997, pp. 93–124.
11. W. an den Daele, A. Pühler, and H. Sukopp, *Transgenic Herbicide-Resistance Crops, A Participatory Technology Assessment, Summary Report* FS II 97-302, Wissenschaftszentrum Berlin (WZB), Berlin, 1997.
12. *Bundesgesetzblatt* (Federal Law Gazette), part I, December 16 (1993), p. 2033.
13. *Official Journal*, No. L 117/1 (1990).
14. *Official Journal*, No. L 117/15 (1990). European directive on release of GMOs, Available at: *www.bbba.de/gentech/genelaw.htm*
15. *Official Journal*, No. L 374/1 (1990).
16. *Bundesgesetzblatt* (Federal Law Gazette), part I (1995), p. 297.
17. *Bundesgesetzblatt* (Federal Law Gazette), part I (1996), p. 1649.
18. *Bundesgesetzblatt* (Federal Law Gazette), part I (1996), p. 1644.
19. *Bundesgesetzblatt* (Federal Law Gazette), part I (1996), p. 1657.
20. *Bundesgesetzblatt* (Federal Law Gazette), part I (1996), p. 1232.
21. *Bundesgesetzblatt* (Federal Law Gazette), part I (1995), p. 734.
22. *Bundesgesetzblatt* (Federal Law Gazette), part I (1981), p. 1. Comparison of Germany with 21 other countries available at: *www.oecd.org/ech/tradedoc/htm*
23. *Bundesgesetzblatt* (Federal Law Gazette), part I (1996), p. 1546.
24. *Bundesgesetzblatt* (Federal Law Gazette), part II (1976), p. 826.
25. *Bundesgesetzblatt* (Federal Law Gazette), part I (1985), p. 2170.

26. Akademie für Technikfolgenabschätzung in Baden-Württemberg, ed., *Bürgergutachten Biotechnologie/Gentechnik — Eine Chance für die Zukunft?*, Stuttgart, 1995.
27. J. Hampel et al., in J. Durant, M.W. Bauer, and G. Gaskell, eds., *Biotechnology in the Public Sphere*, Science Museum, London, 1997, pp. 63–76.
28. Akademie für Technikfolgenabschätzung in Baden-Württemberg, ed., *Verbundprojekt Chancen und Risiken der Gentechnik aus der Sicht der öffentlichkeit*, Stuttgart, 1998. Reports available at: *hampel@afta-bw.de*
29. Council of Europe, *Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine*, European Treaties (ETS), No. 164, Oviedo, 1997.
30. Available at: *www.gwdg.de/~ukee; www.uni-tuebingen.de/-zew; http://ibm.rhrz.uni-bonn.de/iwe/iweframe.htm; www.-europaeische-akademie-aw.de*
31. D. Mieth, in T. vonSchell and H. Mohr, eds., *Biotechnologie — Gentechnik. Eine Chance für neue Industrien*, Springer, Berlin, 1995, pp. 505–530.
32. The Council for Research, Technology and Innovation (*Technologierat*), *Biotechnology, Genetic Engineering and Economic Innovation. Making Responsible Use of Existing Opportunities. Assessment and Recommendations*, Bonn, 1997. Available at: *www.technologierat.de/vdi/*
33. J. Radkau, in M. Bauer, ed., *Resistance to New Technology: Nuclear Power, Information Technology, and Biotechnology*, Cambridge University Press, Cambridge, UK, 1995, pp. 335–356.
34. W. Bender et al., eds., *Gentechnik in der Lebensmittelproduktion: Wege zum interaktiven Dialog, Dokumentation eines Workshops an der Technischen Hochschule Darmstadt*, Darmstadt, 1997.
35. W. Bender, K. Platzer, and K. Sinemus, *Sci. Eng. Ethics* **1**, 21–32 (1995).
36. European Agency for the Evolution of Medical Products, *www.eudra.org/en_home.htm*
37. Deutscher Bundestag, *Bericht der Enquete-Kommission 'Chancen und Risiken der Gentechnologie' des 10. Deutschen Bundestags*, Bonn, 1987, p. 83.
38. *Bundesgesetzblatt* (Federal Law Gazette), part I (1990), p. 2746.
39. E. Deutsch, in D. Evans, ed., *Conceiving the Embryo*, Kluwer, Dordrecht, The Netherlands, 1996, pp. 343–346.
40. U. Wessels, in A. Dyson and J. Harris, eds., *Ethics and Biotechnology*, Routledge, London, 1994, pp. 230–258.
41. S. Graumann, *Biomed. Ethics* **2**(1), 12–16 (1997).
42. German Research Association (Deutsche Forschungsgemeinschaft), *www.dhgp.de*
43. *med. Genet.* **4**, 618–624 (1997).
44. I. Nippert, H. Neitzel, and G. Wolf, eds., *The New Genetics: From Research into Health Care, Social and Ethical Implications for Users and Providers*, Springer, Berlin, 1999.
45. *Ann. Rev. Law Ethics* **7**, 3–340 (1999).
46. S. Boshammer et al., *J. Med. Philos.* **23**(3), 324–333 (1998).
47. Available at: *www.gen-ethisches-netzwerk.de*
48. R. Meyer, C. Revermann, and A. Sauter, Gentechnik, *Züchtung und Biodiversität*, TAB, Bonn, 1998, Available at: *www.tab.fzk.de*

See other INTERNATIONAL ASPECTS entries; INTERNATIONAL INTELLECTUAL PROPERTY ISSUES FOR BIOTECHNOLOGY.

INTERNATIONAL ASPECTS: NATIONAL PROFILES, JAPAN

DARRYL R.J. MACER
Institute of Biological Sciences, University of Tsukuba
Tsukuba Science City, Japan

OUTLINE

INTRODUCTION

Japan spends a high percentage gross domestic product (GDP) on research relative to other nations, and biotechnology spending is a high priority (1). The public in Japan is well educated, and is aware of biotechnology, perceiving both benefits and risks of most applications, and has a reasonable degree of bioethical maturity. Most of the regulation of modern biotechnology is through guidelines and directives issued from numerous government ministries. Among the many ethical issues that have been discussed, trust in authorities is one of the central public policy issues that must be dealt with in the future policy toward biotechnology in Japan.

GENETIC ENGINEERING AND BIOTECHNOLOGY IN JAPAN

Japan is one of the world leaders in modern biotechnology, producing about half of the world's antibiotics, building on a long history of fermentation technology (2). Biotechnology itself, as the use of living organisms to produce goods or services, has a history as long as the humans who have shaped the environment (3). While some may consider biotechnology to be a term more suited to genetic engineering and cell manipulation, to consider its ancient origins is important when we look at the ethical issues it raises and the legal approaches that have evolved. This is especially apparent when we look at the origin of bioethics in Asia, because the links to the past are more emphasized there than in Europe or North America, where terms like "gen-ethics" have arisen (4).

There have been several surveys of the progress of the policy decisions behind Japanese biotechnology research (5–7). The government and industry promoted biotechnology throughout the 1980s, and it was then predicted that by the year 2000 bioindustry would represent 10 percent financially of the Japanese economy (8), with 90 percent of this in traditional industries such as fermentation of food and drink. There have been some joint government and industry efforts to promote biotechnology, including the Bioindustry Development Centre (BIDEC), now called

the Japan Bioindustry Association (JBA), a private think-tank of the Ministry of International Trade and Industry (MITI). The Science and Technology Agency (STA) also has invested in public acceptance of biotechnology. The prefix "bio" has been attached to many new words in the spoken Japanese language, like biocandy or biocosmetics, but perhaps not more so than in the argot of most other countries (9).

There is a very positive view of the contribution of science to improving the quality of life and economy. Research spending in Japan, at 3 percent of GDP, is the highest level in the world, with U.S.$120 billion being spent in the 1996 fiscal year (10). The United States spends about 2.5 percent of GDP of which 0.5 percent is earmarked to the defense industry. About 15 percent of the funding in Japan is from the government, which is close to that of the United States (16 percent) but less than that in the UK (23 percent) or France (37 percent) (11). Between 1996 and 2001 the Japanese government increased spending on research by almost 50 percent.

Agricultural applications were slow to develop, with few field releases of genetically modified organisms (GMOs) in Japan. Although the Ministry of Health and Welfare released guidelines to assess applications for foods and food additives made from GMOs in 1992 (12), at present most of the 30 foods accepted by the regulatory committee are from foreign imports. Despite the efforts to promote biotechnology, there appear to have been some bottlenecks caused by strict or bureaucratic regulations. Up to 1994 there had been 13 GMO field trials compared with over 1000 in the United States by that time (13). However, Japan leads the world effort in sequencing the rice genome.

The Human Genome Project (HGP) had its origins in Japan in the early 1980s, and the genetics programs have been on the rise but without much thought given to ethical, legal, or social impact (ELSI) issues (14). In the 1998 government budget, U.S.$149 million was allocated for genome research (15), and industry also made significant contributions as in other research areas. For example, Takeda Chemical Company was the leading patent claimer in a 1995 survey of world patents on human gene sequences (16), and it also obtained exclusive rights to use of the genetic database bought by SmithKline Beecham from the Institute for Genomic Research in 1992.

PUBLIC ACCEPTANCE OF BIOTECHNOLOGY

The public acceptance of biotechnology in Japan is reasonably high, somewhere in between the attitude in the United States and that in Germany (3,9,17). A number of studies on public acceptance since 1991 have described the bioethical concerns that different groups within Japan have toward biotechnological applications. It may be that the Japanese have the highest familiarity with the word "biotechnology" in the world. In 1991 two surveys found that 97 percent had heard of the word (9), and close results of 94 percent in 1993 (3) and 89 percent in 1995 were confirmed by Hoban (18). Clearly, there is at least high recognition of the word, since programs on genetics and biotechnology are seen on Japanese television almost daily and developments in the field are covered by most major newspapers. There are many science magazines, though they are more in the style of the English language *Scientific American* than of the *New Scientist*.

The importance of medical care, agriculture, and aquaculture to human life is generally acknowledged among the peoples of large societies. The question, is, however, to what extent are the attitudes toward the use of organisms to provide these goods, universal, and the relationships with the organisms and ecosystems that provide the organisms universal, as well as the attitudes to the consumption of the products? To answer this question we need to consider a number of strategies. First, we can look at the use of organisms and new products by different groups with in a society and compare the results. For example, do the people eat meat? Although in Japan meat was not eaten widely a few decades ago, thought to be due to Buddhist influence, it is very difficult to find meals which are strictly vegetarian now. Do the people farm animals in open spaces or in factory farms? In Japan land constraints mean that most animals are in factory farm situations, except in Hokkaido during the warm seasons. We have further to standardize for environmental and economic conditions, and look at the religious traditions.

The religious traditions include guidance on ethical issues, answers to problems that are faced around the world. In one sense looking at the end result of choices, the adoption of science and technology products by consumers, is the best description of acceptance of science and technology. However, if we only look at the consumption statistics we may still not understand the reasons behind the choices, and whether, for example, there was really much choice for the consumer in the home environment and society. The ideal model would say that a consumer determines what products are best, but this may not be apparent in a world dominated by large commercial interests, trade groups and associations, and connections among producers, retailers, and regulators. Since Japan is not self-sufficient in any major species of food (even rice is imported for processed rice products), it is going to be dependent on exporters. However, purchasing power means selection among suppliers, and some new practices in air freight have been introduced in Japan that enable the import of live seafoods and many fresh fruits and vegetables. The market proportion of organic foods and pesticide-free foods has been increasing, though surveys find this to be more for reason of interest in health than for environmental concern (3). Some supermarkets provide nongenetically modified soybean products, like tofu (bean curd), and significant resistance to the products of GMOs emerged in late 1998 and continued through the year 2000.

Another strategy that is used to judge public acceptance is to look at cultural tradition in determining what could be adopted. Schmid in 1991 observed that public acceptance of biotechnology is high, "reflecting a high level of education and information within Japanese society, and the specific way of reaching decisions, which usually involve lengthy discussions with all groups" (7). However, if one asked the Japanese public if they had been involved in the decisions associated with the promotion of biotechnology, it is very doubtful that anyone would say yes. Decision-making in Japan tends to exclude

public participation (13), although certain applications of biotechnology like organ transplantation technology have introduced the idea of individual choice in use of the technology (19), as will be discussed below. In Japan the strong public support of biotechnology is something that cannot be conclusively obtained from public opinion surveys. There is a predominant cultural attitude among the Japanese not to create friction with those they disagree with. Therefore protest movements and oppositions toward biotechnology are small and lack unity. The major outlet for dissent is the media, through discussion forums in magazines, newspapers, and television chat-shows.

The survey strategy that allows us to look at what individuals will accept, and their reasons for this, must be supplemented in Japan by topics covered in small group discussion forums. The survey results from Japan, and also from other Asia-Pacific countries (3), compared with the rest of the world, reveal an important distinction in the main concerns people have about biotechnology. They want to protect nature, not because of its value or property, but simply because it is there. The bioethics part of risk assessment, which elsewhere takes precedence over the analysis and prediction of risks, is combined with the value of avoiding harm and the benefits of doing good or beneficence. The assessment of risk in biotechnology involves both the potential to change something and the potential to do harm (20). The extent to which a change is judged to be a subjective harm depends on human values, whether nature should be "intransient" or modified. Open survey questions in all countries reveal that the major determinant of moral acceptability of a technology is whether it is perceived to be unnatural, or morally acceptable (3,9,17,21).

In the 1993 International Bioethics Survey (3), when asked about specific developments of technology, including in vitro fertilization (IVF), computers, pesticides, nuclear power, biotechnology, and genetic engineering, both benefits and risks were cited in open comments by many respondents in Japan as in Australia, Hong Kong, India, Israel, New Zealand, the Philippines, Russia, Singapore, and Thailand. In Japan 74 percent saw biotechnology as worthwhile, which was less than the 85 percent in the 1991 survey (9) but still a high positive response. In both years 37 percent said that they had not worried about its development. Thirty percent of those who cited a benefit in 1993 said it would help humanity, 19 percent said agriculture in general, and 15 percent said science. About a half did not describe any benefit or concern. In addition to those who saw it as unnatural, another common response was of human misuse.

In the 1997 telephone surveys (17), a single question on the perceived impact of seven areas of science and technology was used. Comparisons with the data from the European Commission, Eurobarometer 46.1, reveal that there is more optimism about solar energy, new materials, and space exploration in Japan (and New Zealand and Canada) but similar optimism toward computers, information technology, and telecommunication in the EU. There is less optimism about biotechnology and genetic engineering in Japan (with New Zealand being even lower). Only 62 percent in Japan thought that biotechnology would improve the way we live in the next 20 years, 12 percent thought biotechnology would make things worse, and 4 percent said they perceived no effect, with 22 percent saying they do not know. For genetic engineering 54 percent saw it as worthwhile, 12 percent as making things worse, and 7 percent as having no effect. In 1991, 76 percent in Japan said that genetic engineering would be a worthwhile area to explore, while 20 percent were extremely worried about the consequences (9). In 1993, 57 percent said that genetic engineering was worthwhile scientific research, while 15 percent still had a lot of worries about it. In Japan there does not appear general trend against genetic engineering over time, unlike the situations observed in Europe (22) or New Zealand (17).

When people were asked in 1997 to say what came to mind on hearing the term "biotechnology," 8 percent expressed a concern and 4 percent expressed a positive view of science, but most people just said "something technical" (17). In all Asian countries there is strong support for certain kinds of environmental release of GMOs (3). Plant genetic engineering is regarded more favorably than microbe, animal, or human genetic modifications (23), except for gene therapy for diseases like cancer, which is received very positively in Japan (24,25). Despite the concern expressed about genetic engineering, in 1997, 35 percent in Japan said they would buy genetically modified fruits if they tasted better, suggesting a positive image of the products. However, only 8 percent said current regulations are sufficient to protect people from any risks connected with modern biotechnology.

JAPANESE LAW AND BIOTECHNOLOGY

There are few specific laws on modern biotechnology in Japan but rather a series of regulations by the different Ministries. Many scientists and people in industry claim that these regulations have inhibited research development, which is a different view from the public as shown above.

The modern Japanese Constitution was drafted by the occupation forces after the Second World War, was reviewed by the Japanese government, and voted into force by the Japanese Diet (Parliament) in 1948. It has had almost no changes since then, reflecting a trend for laws to become fixed. It includes 31 Articles on the rights and duties of the people. The right of people to life, liberty, and the pursuit of happiness, to the extent that it does not interfere with public welfare, should be the supreme consideration of the law and government (Article 13). Two articles are relevant to access to biotechnology:

1. *Article 11*. The people shall not be prevented from enjoying any of the fundamental human rights. The fundamental human rights guaranteed to the people by this Constitution shall be conferred upon the people of this and future generations as eternal and inviolable rights.
2. *Article 25*. All people shall have the right to maintain the minimum standards of wholesome and cultured living. In all spheres of life, the State shall use its

endeavors for the promotion and extension of social welfare and security, and of public health.

Article 25 assumes a welfare state but does not have much legal meaning. It does not vest in each individual a concrete right that can be enforced by the judicial process, as such type of right comes into force only through implementing legislation. There are six major Codes in addition to the Constitution. Many medical procedures and medical protocol are regulated by specific legal Acts. The Civic Code is concerned with family and inheritance and was completely amended after the Second World War. The Criminal or Penal Code includes some relevant Articles. When there is a more specific law and it conflicts with a general code, the specific law usually takes precedence. Administrative guidance by government agencies and local authorities plays the more significant role in many of the biotechnology issues. The power in the guidance is that the Ministries have the power to grant licenses or permissions. There is little legislation on recent bioethical issues.

The basic philosophy of the Japanese health care system is universally mandated, government-provided health insurance coverage. There is little choice over which insurance scheme a person must join. Employees must join the one statutory plan offered by their employers, and self-employed persons must join the plan administered by the local government or by their trade associations (26). In the year 2000 a public long-term care insurance program will provide for some extra services for the elderly or chronically sick, such as home help, visiting nurses, or day care (27).

The Preventive Vaccination Law established a national program for influenza vaccination in 1976. The Law was weakened in 1987 by removing its obligatory nature, and was further weakened in 1993 with the broadening of exceptions and the removal of provisions that penalized parents who failed to have their children vaccinated. Influenza vaccination is performed annually in children aged 3 to 15 years in certain target groups. However, it has not been very effective, and has recently been recommended for young children only (28). Since 1987 it has been easy for parents to refuse influenza vaccinations, resulting in large differences among the various kindergarten, primary, and junior high schools. Between 1951 and 1965, 169 persons died because of reactions to vaccination. Articles 16 to 19-4 provides for a system of national compensation, under which there are set reimbursements for injuries. For example, in April 1992 a death was compensated by 20.5 million yen, with funeral costs of 140,000 yen. A pension after 18 years of age was 2,925,900 yen for the first category of disability and 1,910,500 yen for second class of disability (29).

A controversy erupted in 1993 over the high incidence (1 in 400) of side effects from a MMR (mumps, measles, rubella) vaccine made and used in Japan. It was withdrawn after the media released unpublicized government risk data. In 1994 no MMR vaccine was offered to children because the government refused to use the U.S. vaccine, which has a 20-year history of safe use with almost no side effects. The scandal reveals that the Japanese Ministry of Health and Welfare has been attempting to encourage Japanese industry by not using a foreign vaccine, while risking public health with a vaccine with 100 to 200 times more side effects. Parents who want their children vaccinated had to pay about U.S.$80 for a vaccine that was previously free (29).

Another sign of the support for the biotechnology industry is the system of drug reimbursement, and the overuse of antibiotics in Japan. One of the embarrassments of the Japanese health care system is the corruption that is implicit in the way drug prices are set and reimbursement is made, and the contributions from pharmaceutical companies to doctors who use their drugs. The Japanese are the world's highest spenders on prescription drugs (30). Almost all general practitioners and hospitals have their own pharmacies for outpatients. Every two years the Ministry of Health and Welfare sets the "official" prices for all drugs. These prices are used to determine the charges to patients and the national health insurance systems. However, pharmaceutical companies offer drugs to hospitals at a discount. The permitted discount is 10 percent, which means there is even official sanction of the scheme to have financial reimbursement for dispensing prescription drugs. In practice, the current discounts are 20 to 30 percent or more in competitive markets. This means that hospitals and doctors benefit from prescribing drugs, and it explains why the consumption of drugs is so high.

The average use of antibiotics is 3 times the U.S. average and 20 times the United Kingdom average (29). There is concern over methicillin resistant *Staphylcoccus aureus* (MRSA) infections and deaths, which are relatively common in Japan. In one old person's home in Chiba, Japan, 12 residents out of 80 were infected with MRSA in 1992. Many products are allowed to support the presence of a large local pharmaceutical industry.

Before a new drug is approved for use, the Central Pharmaceutical Affairs Council within the Ministry of Health and Welfare must examine the results of toxicity tests, animal tests, and clinical trials. The basic policy is outlined in the Notice No. 645 of 1967, on "Control of the manufacture of and trade in pharmaceutical products: approval of manufacture and import of medicaments." Drugs must pass three phases of clinical trials:

1. Phase I is to check a drug's safety in humans.
2. Phase II is to access its therapeutic index (the response rate and severity of side effects) in selected populations of patients for whom the drug is intended.
3. Phase III is to determine whether the new treatment is better than existing ones.

All these tests are done at universities or research organizations at the request of pharmaceutical companies, and there have been cases of bribery (29). There is pressure today for more of the trial results to be presented in scientifically refereed journals, but until now much of the data has been either in internal documents or in the drug companies' own journals. International standardization of tests in 1991 in the United States, Europe, and Japan avoided some duplication of tests and halved the length of long-term studies. In practice, when a drug has passed phase two trials, it may be widely prescribed in "trials,"

and there are some top-selling drugs that are said to be anticancer drugs which are only prescribed in Japan (30). The Phase III trials can become part of the pharmaceutical companies' marketing plan. There are also widespread issues of lack of informed consent in such clinical trials at the Phase II and Phase III level because most patients do not know what medicine they are being given.

The government can suspend production by a company if it fails to report details of clinical trials, especially if these involve deaths. On May 20, 1994, the chairman of Japan's fourth largest pharmaceutical wholesaler resigned to take public responsibility for insider trading among employees, and for the sale of sorivudine, which has led to the deaths of 18 people. The company had reached settlements by May 1994 with 10 of the 23 families of patients who died or had serious consequences as a result of the using drug. The penalty of a 105-day suspension of production at the Nippon Shoji Kaisha, Ltd. Okayama prefecture factory in response to the company failing to report the deaths of two persons during clinical trials of sorivudine occurred on September 2, 1994. This is the longest suspension of products that has ever occurred because of an infringement of the Drugs, Cosmetics, and Medical Instruments Law (31). It also gave time for reflection on the system of clinical trials in general (32).

The journals that publish the results are often the same journals that sponsor the trials or make the drug (32), and all major pharmaceutical companies are also involved in research on production of new drugs using modern biotechnology methods. Another feature of Japanese biotechnology is that most research is conducted in large established companies, usually multinational, rather than the small biotechnology companies that are a feature of North America. However, it could be that the sponsorship links are just more obvious than elsewhere.

Despite the blood donation system, which in 1991, saw 8,861,137 persons donate blood, 6.5 percent of the population, there is also a large import of blood, since Japan uses more blood per person than any other country in the world. The Product Liability Law passed unanimously by the National Diet on June 22, 1994, came into effect from July 1, 1995. It virtually excludes any liability on transfusion products. "Complications of blood transfusion such as those caused by contamination of viruses whose complete removal by existing technology is impossible cannot be considered as product defects" (33). There have been compensation claims paid to some of the blood transfusion victims of delayed implementation of heat treatment procedures to eliminate HIV from imported blood products (34) and some officials were sentenced to prison.

Patent claims on products in Western countries are recognized in Japanese patent cases, and there have recently been cases involving the use of recombinant DNA products. The approval of such products is independent of the patent claims. A Japanese court rejected a claim by Hoffman-Roche that a Japanese company infringed its patent on interferon, but the case will be appealed (29). The sales of interferon has rapidly been rising since it was approved for use against chronic hepatitis C. An Osaka local court in January 1991 decided that the company Toyobo cannot market tissue plasminogen activator (TPA) because it conflicts with the Japanese patent given to Genentech in January 1991. Genentech licensed two other companies to sell TPA in Japan. TPA has been sold in Japan since May 1991. The case involving the rights to sales of erythropoietin (EPO) in Japan was solved out of court. EPO has also a very large market in Japan where kidney transplant rates are low.

Privacy of communication is guaranteed in the Constitution. Article 21 of the Constitution guarantees freedom of assembly and association as well as speech, press, and all other forms of expression. Censorship is prohibited, and secrecy of any means of communication must not be violated. There is the Law on the Protection of Computer Information on Individuals, which states that government agencies are prohibited from using the information on individuals for purposes other than the original purpose for which the files were compiled. Any person may require a government agency to disclose the information on themselves that is stored in the computer, and if necessary, demand its alteration. It could be interpreted to mean the truth of any health check information entered into a computer must be revealed following a person's request.

If someone informs others of the medical data of a person, for example, the result of genetic screening test to an employer, Section 134-1 of the penal code could apply. If the person who leaked the information is a national employee, he or she will be punished by the Law on Government Employees. There is still debate over how to control life insurance companies' questioning on the results of genetic tests, but other family history data and smoking are currently used in deciding policies.

For any medical intervention, physicians are required to obtain consent to medical treatment according to the Medical Practitioner's Act, Article 23. The Supreme Court, in 1949, (Decision 3.1) said that the obligation for treatment is based on assessing what can reasonably be expected in view of the knowledge and experience that ought to characterize the average physician. However, in practice and in court cases in the 1980s and 1990s, the doctrine of informed consent has yet to be fully recognized (35,36) as a right for patients to be told all information. Nevertheless, more doctors are starting to use informed consent, and truth telling in cancer cases has also been increasing (37).

Organ transplantation using cadaver donors, especially those that are determined to be dead by brain death criteria, is rare. A law permitting such transplants, and allowing whole brain death criteria to be used for determination of death if patients themselves have signed a donor card to that intent and if family members do not object to it, was passed in 1997. Until then, Law No. 64 (Article 4.17) enabled cornea transplantation since 1958, while in 1979 the Act Concerning the Transplantation of Cornea and Kidney was passed allowing kidney transplants. Kidney transplants since 1979 have been from both live and dead donors, with prior consent and family approval. In the mid-1990s there were few transplants from brain dead donors while the new organ transplant law was being debated. There was much

discussion of the issue and whether it was related to any particular Japanese ethos or just suspension of the medical profession (14,19,38). From 1968, when the first heart transplant was performed, until 1997 there had been wide debate on the question in Japan, which was a rare, if not unique, occasion for extensive public debate on a biotechnology process. Live liver donations were possible but most patients had to seek heart and liver transplants overseas.

Until 1996 the 1948 Eugenic Protection Law governed the use of abortion services in Japan. The number of abortions conducted is declining, but it is still high among developed countries, and in 1996 the title was changed to Mother's Body Protection Law (39). After World War II, the Japanese government changed the population policy into "to stabilize and not to increase" from "to increase." How to popularize family planning became the primary policy in health care of postwar Japan. At the same time the Eugenic Protection Law was promulgated in 1948, and Japan became the second largest populated country after the Soviet Union in the liberalization of induced abortion.

The Eugenic Protection Law was a modification of the Preventive Law of Offspring with Hereditary Diseases (*das Gesetz zur Verhuetung erbkraanken Nachwuchses*), 1933 of Germany under Hitler, combined with a liberal view of induced abortion (39). In June 1996, however, some inappropriate parts of this law were amended by the omission or elimination of the eugenic articles, and the title of this law changed from Eugenic Protection Law to Mother's Body Protection Law.

Fetal diagnosis and selective abortion, while not explicitly allowed under these laws, is however, widely practiced when the pregnancy could cause psychological distress to the mother or economic hardship, both of these being acceptable reasons under the law for induced abortion. The marketing of genetic diagnosis and triple marker biochemical tests is governed under existing laws for pharmaceutical products and devices. Abortion is restricted to the period in which the fetus is not viable outside of the uterus, and this period is determined by the notification from the Ministry of Health and Welfare, currently being 22 weeks. The Japan Society of Human Genetics has voluntary guidelines on use of genetic screening (40).

IVF and assisted reproductive technology for married couples are guided by the voluntary guidelines of the Japan Society of Obstetrics and Gynecology (JSOG), and moves to introduce a law have been resisted by the medical community (41). The first baby was born after IVF and embryo transfer (ET) in 1983, and there was much media attention. There are significant differences in attitude among infertile couples, medical practitioners, and the general public on the technology (3,9,42). Most Japanese obstetricians belong to JSOG, and after the first baby was born they rushed to form an ethical committee concerning IVF ET. This committee consisted of 14 members of JSOG. While listening to the opinions of representatives from the mass media and highly educated laypersons, in October 1983, they wrote and announced a statement to the Society of Obstetrics and Gynecology. Even though it was a little too late, this was probably the first time that a Japanese medical society made and announced ethical guidelines for its members. The statement is as follows (29,39):

1. The Method should only be used for women who are judged unable to become pregnant by any other medical method.
2. The individual implementing this Method must be a qualified doctor who has mastered a high standard of knowledge and technology in the field of reproductive medicine. Every procedure and treatment should be carried out with the utmost care. The procedures and expected results of the Method should be sufficiently explained to the applicants concerned prior to implementation of the Method. Upon obtaining consent from the applicants, and acknowledgment should be filled out and signed by the applicants and retained by the doctor.
3. The applicant receiving the Method should be married, have a strong desire for a child, and be in satisfactory mental and physical condition for pregnancy, delivery, and raising of a child. It must be possible to successfully conduct retrieval of mature ova, implantation into the uterus and maintenance of pregnancy.
4. The fertilized ovum should be carefully handled in respect to the basic moral values of life.
5. When implementing the Method, no gene manipulation is permitted.
6. The privacy of the couple and their delivered child should be respected and protected according to relevant laws and regulations.
7. Considering the importance of the Method, the organization using it should provide opportunities to hear opinions from individuals other than those directly concerned.

The number of babies born through IVF/ET has now reached an annual total over 10,000 since the first case in 1983 in Japan. Surrogacy, however, is not permitted, though foreign surrogacy agencies have been used by Japanese clients, and at least two agencies operate for U.S. surrogacy businesses in Japan (29).

Artificial insemination by donor sperm (AID) is conducted largely through the Obstetrics and Gynecology Department of Keio University, Tokyo. AID has no law to regulate it, and it started in 1948 (39). Now there are about 500 attempts at AID a year at Keio University, and about 250+ births per year. Each sperm donor is used for up to 15 pregnancies, and only married women are accepted. Keio University is the most public about its program. Other institutes do not admit having a program. The guidelines used are those of Keio University and Japan Society of Obstetrics and Gynecology. Since a conference discussion of the Japanese Association of Civil Law in 1953, many legal scholars have construed the law to allow the AID baby of a married woman to be a legitimate child of her husband, so long as the procedure was carried out according to the current practice, but there is still no specific law.

Preconception sex selection has been investigated in Japan, but in a 1993 survey, 76 percent said that if they had only one child they would want a girl, suggesting

that traditional ideas of family inheritance are discounted by many people (29). The reason why more people wish to have a girl than a boy, which is in contrast to many other Asian countries, may be because girls are considered cuter, or better caregivers for elderly parents. Following its announcement in 1986, the Ethical Committee of JSOG, and at the same time, the newly founded Ethical Committee of the Japan Medical Association (JMA) came to about the same conclusion in September 1986. That is, it was decided that this procedure should only be adopted to prevent the creation of conceptuses with severe sex-linked recessive genetic disorders.

The details of the Ethical Committee of JSOG's statement are as follows (39):

1. Any individual implementing the Method must be a qualified physician who has mastered a high standard of knowledge and technology in the field of reproductive medicine.
2. A physician intending to implement the Method must be previously registered with the Society according to the specified format. It is also desirable that the results be reported to the Society.
3. Before application of the Method, the physician should sufficiently explain the procedures and expected results to the individual(s) concerned, and should obtain their written consent.

It is also against the guidelines of the Ministry of Health and Welfare to generally inform parents the sex of the fetus during routine prenatal ultrasound diagnosis.

The Ministry of Health and Welfare in Japan set up a special Ethics Committee to assess applications for gene therapy in 1994. The Ministry of Education also made guidelines and set up a separate committee (with seven overlapping members). In university hospitals, drugs already need the approval of both Ministries, and so does gene therapy. The first protocol was approved by Hokkaido University in 1994 allowing research on one child with ADA deficiency. Approval was given by both ministries. The guidelines are basically those of the National Institutes of Health (NIH) in the United States. The guidelines rule out germ-line therapy and until 1999 limited cases to terminal illnesses without effective therapy. However, they only require verbal informed consent, not the written consent that may be determined by local hospitals policies. Japanese scientists and public strongly support the use of gene therapy (24,25), but progress has been slow due to regulatory delays. There is also a lack of domestic vector production, and many trials that are considered are in collaboration with U.S. companies. Japanese people disapprove of use of enhancement genetics in surveys, unlike tendencies seen in China, India, or Thailand (25).

In Japan by 1994, DNA fingerprinting had been used in 180 criminal investigations, but it had only been used 12 times as evidence (29). It is more common to use blood typing and other methods, but it is being introduced in the same manner as other modern forensic techniques. By 1997 advertisements to fathers to check their real genetic relationships to children had appeared in popular magazines, without apparent regulation.

JAPANESE CULTURE, BIOTECHNOLOGY, AND BIOETHICS

The data from modern public opinion surveys needs to be interpreted in the context of the cultural heritage of Japan. The relationships of human beings within their society, within the biological community, and to nature and God are a fact of prehistory; therefore we cannot precisely define the origins of bioethics (3). One of the major elements that needs to be considered in Japanese bioethics is the history of polytheism and animism. However, during the expansion of agriculture and paddy fields over 500 years Japan has seen similar disregard for the environment as have Western countries, suggesting that religious belief does not overcome economic or self-interest (43). The decision to burn a forest and plant a crop is a bioethical decision, and we can see that almost all possible land has been utilized for agriculture, industry, or urban life, with wilderness area remaining only in those regions that proved difficult to exploit.

Japanese ethics is a mixture of Buddhist and Confucian influences combined with a later Shinto influence, and more recently Western influences. From the fifth and sixth centuries the medical profession was restricted to care for the privileged classes. With the centralization of government in the seventh and eighth centuries, there was established a bureau of medicine, and by the Yoro penal and civil codes, there was created an official physician class. The shortages of doctors opened the profession to others. After the Heian period (800–1200) the government-sponsored health service was replaced by a council of professional physicians. In the sixteenth century a code of practice was drawn up the physicians that is very similar to the Hippocratic code and is called the Seventeen Rules of Enjuin (44).

In all areas of public policy, committees of experts work together with bureaucrats to issue reports and guidances. There is a Ministry for the Environment that attempted to introduce a law to govern genetic engineering in 1992, but it was blocked in a power struggle between resistant academics and the Ministries for Agriculture, Forestry, and Fisheries, the Ministry of Education, Science, Culture, and Sports, the Ministry for International Trade and Industry, the Science and Technology Agency and the Ministry of Health and Welfare — all who had their own regulations and committees on biosafety and release of GMOs (9). The first two guidelines had been introduced in 1979 by the Ministry of Education, Science, Culture, and Sports, and the Science and Technology Agency. In 1986, following the OECD recommendations, the other three Ministries also introduced guidelines. Given the interministry division of duties, it is not surprising that the smaller, then Agency for the Environment, could not push through a law claiming it had jurisdiction for all GMO releases over the other Ministries. Each Ministry has revised its guidelines gradually but has kept control over its traditional areas, the same as other biotechnology applications and research.

For the medical discipline, in addition to the Ministry of Education, Science, Culture, and Sports and the Ministry of Health and Welfare, there is a Council of Medical Ethics established under the provisions of Article 25 of the Medical Act. It is an advisory body

supervised by the Minister of Public Welfare, consisting of the presidents of the Japan Medical Association, the Japanese Dental Association, and scholars and staffs from related administrative departments. It functions to take administrative measures to eliminate physicians and dentists who commit acts of malpractice or unethical acts. If the media exposes a scandal, then usually top officials or Ministers must resign, an investigatory committee may be established, and then there is a proclamation of policy change. However, Japanese politics is dominated by long-term stability, as the ruling parties and coalitions have been in power except for a year since the Second World War.

Currently Japanese medical ethics is under change (35), recognizing that Japanese society contains people with a similar diversity of views to that of Western countries. The hesitant introduction of bioethics is more related to the structure of Japanese society than to any difference in a person's attitude in Japan or the Western countries (13,19). This fact emerged from opinion surveys where individuals were asked to give their reasons for their responses to bioethical issues regarding genetic manipulation or screening. There was at least as much variety expressed by members of the general public in Japan as there has been in other countries (3).

It must be noted that in terms of equity of access to biotechnology applications, apart from access to prenatal genetic screening and gene therapy, and novel treatments for rare diseases, most other applications of modern biotechnology are accessible to all under the universal medical coverage system. There have been criticisms of the health research system in Japan (45). Open debate is still not common, and this may be the greatest public policy need in biotechnology.

THE FUTURE AND TRUST

In summary, considering the two sides of bioethics, descriptive and prescriptive, a key issue is trust. If we describe the ethical issues that people think are associated with biotechnology in Japan, we find great diversity, the same as in other countries. One common feature, however, is a lack of trust in the process and the policy makers. The prescriptive ethics, or processes that can be used to make decisions and/or the range of decisions that can be made, has been influenced by the relativism that is perceived as correct in Japan. This means rather than one absolute view being right or wrong, we should respect the view of others and not challenge them. This is enshrined in the Constitution and also is in the spirit preserved in the choices given over acceptance of brain death for organ donors (though the family can override the decision of an individual).

While people should not judge, policy must be formulated. What is good for one person may not be good for the broader society, and the global nature of agricultural economics and environmental impact mean we have to think far beyond the small field trial of a GMO. Prescriptive bioethics not only calls for certain factors to be included in decision making but that certain groups of people be involved. Different groups of people may call for different levels of risk assessment, and of what constitutes a significant risk. Therefore a central question is who can be trusted, and how the public in Japan can regain trust in authorities.

In the 1997 survey, when given a range of bodies, international organizations like the United Nations and WHO were considered the best placed bodies to regulate modern biotechnology by 62 percent in Japan compared to 34 percent in Europe (17), reflecting another cultural value that the opinions of those outside the country may be more trusted than the opinions of policy makers within the country. Europeans and North Americans prefer their local authorities. Therefore the legal tolerance limits of acceptable risk and harm as broadly outlined in international covenants such as the Declaration of Human Rights, the UNESCO Universal Declaration on the Human Genome and Human Rights, and international treaties on environmental protection, have been well accepted as a cultural norm and something to aspire to. Japan is less likely to break with world opinion than, for example, the United States, which regards national autonomy as a higher ideal.

In the same question, industry was more trusted to regulate modern biotechnology than public authorities, the reverse of other countries surveyed (17). This may reflect that by 1997 there was very little trust left in the government in Japan, not that industry is particularly well trusted compared to consumer organizations or university scientists. In the 1991 survey, when asked how do you think biotechnology should be regulated, 62 percent of the public chose the option standards and practices agreed upon jointly by industry and government, with 19 percent saying by the government alone, and 2 percent by industry alone, and 5 percent by individual researchers (9). There was more support for government expressed by scientists and high school teachers who were included in the same survey.

The distrust of authorities by the Japanese does not stem from lack of knowledge. Surveys of the different groups in Japan showed that educated people have as much concerns; in fact biology teachers considered there to be more risk from genetic engineering than the general public (9,17). The risk perceptions among scientists had a tendency to be more concrete than among the public, but all groups expressed a wide variety of concerns. In related questions on the risks of genetic engineering to animals and humans, only 16 percent expressed concerns that it meant interfering in nature (9). However, as in most countries among Japanese academics, industry scientists, and public authorities, there are still claims that increased knowledge is correlated to decreased perception of risk. This is not supported by empirical studies. The balancing of benefit and risk is necessary for bioethics, and it is the most effective indicator of the bioethical maturity of a society (46). The use of surveys can provide us with knowledge of the degree to which a society can make well-reasoned, "mature" judgment rather than offer impulsive, "childish" views based on immediate gain.

In late 1997 one of the government agencies, the Council for Science and Technology of the Science and Technology Agency, established a bioethics committee and

a special committee to consider legislation on human cloning (47). The Agency, however, has been criticized for allowing meetings to be closed to public participation, despite claiming as its mission the universal response of all Ministries to a problem where "open and nationwide debate and public consensus is needed." Many Japanese people, of course, are aware that debates do not reach the public, so they rely on opinions expressed by selected experts. Several Ministries have started to open to the public meetings on bioethical issues like brain death and gene therapy, but otherwise most meetings stay closed.

Bioethical decision making involves recognition of the autonomy of all individuals to make free and informed choices provided that they do not prevent others from making informed choices. This is consistent with democratic principles, and the extent to which a society has accepted this is the one criteria of the success of bioethics. However, the structured paternalism of Japanese society is built on the idea that only the views of so-called experts (*sensei*) should be heard (13). It also means that their views should not be questioned, in accordance with the traditional paternalistic Confucian ethos. Medicine is "an Article of Jin," the expression of loving kindness (*Jin*) by the health care professional (34). The main theme of Confucianist ethics was the maintenance of moral discipline for the nation, society, and the home, and it was to the benefit of rulers and family leaders. Therefore it is not surprising that many of the authorities in Japanese society share this ideal because it means respect for them, and hence the rejection of an autonomous development of bioethics (13). They may promulgate the idea that Japanese are different as an attempt to prolong the Confucian ethic.

The bioethics debate may be the catalyst required to transform Japan from a paternalistic democracy. People of any country have at times resisted rapid change, and the globalization of ethics, ideals, and paradigms (may cause) ethnic and national identities to be changed, or lost, especially those of countries with unique cultural histories. How countries approach globalization is a fundamental question, but many individuals in countries with access to common news media have already answered the question by their converging lifestyles and values. To the extent that human rights and the environment are more respected, this trend is to be encouraged.

When Japan opened its doors to Western society in the nineteenth century, it was introduced to a newly emerged science and scientific paradigm that was only part of the fabric of Western society. Meanwhile Western society has continued to evolve, and bioethics has emerged. There has been a series of meetings on bioethics initiated in Japan, both through the Japan Association of Bioethics founded in 1987 and through international seminars on topics such as the Human Genome Project (48–51). There are several bioethics centers and university departments at which it is possible to do research in bioethics, but no degree course is specialized only in bioethics. An early part of this development has included importing and debating ethical approaches, but the current phase has opened up a multidisciplinary dialogue that has included the public in the discussion and development of its diverse, indigenous ethical traditions. Modern biotechnology may be the stimulus to transform Japanese public policy to better encourage people's involvement in technology decisions.

BIBLIOGRAPHY

1. Committee for Scientific and Technological Policy, Working Party on Biotechnology, *Economic Aspects of Biotechnologies Related to Human Health*, OECD, Paris, 1998.
2. U.S. Congress Office of Technology Assessment, *Biotechnology in a Global Economy*, U.S. Government Printing Office, Washington, DC, 1991.
3. D.R.J. Macer, *Bioethics for the People by the People*, Eubios Ethics Institute, Christchurch, New Zealand, 1994.
4. D. Macer, *Nature* **365**, 102 (1993).
5. M.V. Brock, *Biotechnology in Japan*, Routledge, New York, 1989.
6. R.T. Yuan and M.D. Dibner, *Japanese Biotechnology: A Comprehensive Study of Government Policy, R&D and Industry*, Macmillan, London, 1990.
7. R.D. Schmid, *Biotechnology in Japan. A Comprehensive Guide*, Springer-Verlag, Berlin, 1991.
8. BIDEC, *Impact of Biotechnology on Industrial Structure in the Year 2000*, Japanese Fermentation Association, Tokyo, 1986 (in Japanese).
9. D.R.J. Macer, *Attitudes to Genetic Engineering: Japanese and International Comparisons*, Eubios Ethics Institute, Christchurch, New Zealand, 1992.
10. Editor, *Science* **278**, 1887 (1997).
11. J. Grant and G. Lewison, *Science* **278**, 878–879 (1997).
12. Ministry of Health and Welfare, *Guidelines for Foods and Food Additives Produced by Recombinant DNA Techniques*, Ministry of Health and Welfare, Tokyo, 1992.
13. J. Kinoshita, *Science* **266**, 11845–1185 (1994).
14. D. Macer, *Nature* **359**, 770 (1992).
15. D. Normile, *Science* **278**, 1700–1792 (1997).
16. S.M. Thomas et al., *Nature* **380**, 387–388 (1996).
17. D. Macer et al., *Eubios J. Asian Int. Bioethics* **6**, 137–151 (1996).
18. T.J. Hoban, *Nature Biotechnol.* **15**, 232–234 (1997).
19. D. Macer, *Politics Life Sci.* **13**, 89–90 (1994).
20. D.R.J. Macer in A. van Dommelen, ed., *Coping with Deliberate Release: The Limits of Risk Assesment*, International Centre for Human and Public Affairs, Tilburg, The Netherlands, 1996, pp. 227–245.
21. D.R.J. Macer, in D. Brauer, ed., *Modern Biotechnology: Legal, Economic and Social Dimensions, Biotechnology*, vol. 12, VCH, Weinheim, Germany, 1995, pp. 115–154.
22. Biotechnology and the European Public Concerted Action Group, *Nature* **387**, 845–847 (1997).
23. D.R.J. Macer, in K. Watanabe and E. Pehu, eds., *Plant Biotechnology and Plant Genetic Resources for Sustainability and Productivity*, R.G. Landes, Austin, TX, 1997, pp. 87–99.
24. D.R.J. Macer, *Hum. Gene Ther.* **3**, 511–518 (1992).
25. D.R.J. Macer et al., *Hum. Gene Ther.* **6**, 791–803 (1995).
26. N. Ikegami, *Science* **258**, 614–618 (1992).
27. N. Ikegami, *J. Am. Med. Assoc.* **278**, 1310–1314 (1997).
28. S. Morio et al., *Epidemiol. Community Health* **48**, 46–51 (1994).

29. D.R.J. Macer, in *International Health Care Law*, VistaMedia Corp: USA, & Center for International Legal Studies, Salzburg, Austria (for CD-ROM and print), in Press.
30. M. Fukushima, *Nature* **342**, 850–851 (1989).
31. C. Ross, *Lancet* **343**, 1418–1419 (1994).
32. M. Fukushima, *Nature Med.* **1**, 12–13 (1995).
33. M. Yawata, *Lancet* **344**, 120 (1994).
34. T. Hannay, *Nature Med.* **1**, 396 (1995).
35. R. Kimura with L. Bishop, in W.T. Reich, ed., *Encyclopedia of Bioethics*, rev. edn., Simon & Schuster Macmillan, New York, 1995, pp. 1496–1505.
36. N. Tanida, *Lancet* **345**, 1176 (1995).
37. H. Hattori et al., *Soc. Sci. Med.* **32**, 1007–1016 (1991).
38. M. Lock, *Cult. Med. Psychiatry* **19**, 1–38 (1995).
39. S.N. Shinagawa, *Eubios J. Asian Int. Bioethics* **6**, 158–160 (1996).
40. Japan Society of Human Genetics, *Eubios J. Asian Int. Bioethics* **6**, 137–139 (1996).
41. K. Bai, Y. Shirai, and M. Ishii, *Hastings Center Rep.* (special supp.), 18–20 (1987).
42. Y. Shirai, *J. Law, Med. Ethics* **21**, 43–53 (1993).
43. K. Friday, *Eubios J. Asian Int. Bioethics* **8**, 46 (1998).
44. J.M. Kitagawa, in W.T. Reich, ed., *Encyclopedia of Bioethics*, rev. edn., Simon & Schuster Macmillan, New York, 1995, pp. 1491–1496.
45. K. Imamura, *Lancet* **342**, 279–282 (1993).
46. D.R.J. Macer, *Soc. Sci. Med.* **38**, 23–33 (1994).
47. A. Saegusa, *Nature* **391**, 313 (1998).
48. N. Fujiki and D.R.J. Macer, eds., *Human Genome Research and Society*, Eubios Ethics Institute, Christchurch, New Zealand, 1992.
49. N. Fujiki and D.R.J. Macer, eds., *Intractable Neurological Disorders, Human Genome Research and Society*, Eubios Ethics Institute, Christchurch, New Zealand, 1994.
50. M. Okamoto, N. Fujiki, and D.R.J. Macer, eds., *Protection of the Human Genome and Scientific Responsibility*, Eubios Ethics Institute, Christchurch, New Zealand, 1996.
51. N. Fujiki and D.R.J. Macer, eds., *Bioethics in Asia*, Eubios Ethics Institute, Christchurch, New Zealand, 1998.

See other INTERNATIONAL ASPECTS entries; INTERNATIONAL INTELLECTUAL PROPERTY ISSUES FOR BIOTECHNOLOGY.

INTERNATIONAL ASPECTS: NATIONAL PROFILES, SCANDINAVIA

MATS G. HANSSON
Uppsala University
Uppsala, Sweden

OUTLINE

INTRODUCTION

The development of biotechnology is intense in the Scandinavian countries, Denmark, Finland, Iceland, Norway, and Sweden. In a sense this development is a natural continuation of a long tradition originating with the Swedish botanist and professor of medicine, Carl von Linné (1707–1778). Linnaeus was eager to learn everything there is to know about nature; he and his disciples traveled the world to collect information about both natural and human resources, information that he then systematized and described in detail. Linnaeus lived at a time when the significant economic potential of science was first being recognized in Swedish political life, and he furthered this development. The Swedish historian of science, Tore Frängsmyr, has described this contribution as part of the Linnaeus heritage:

> He himself helped spread the new economic thinking and often maintained that natural history, his own science, formed the basis of any sound economy. Both agriculture and industry used nature's products, so knowledge of the three realms was fundamentally important. . . .Linnaeus was very willing to put his branch of science at the service of the economy. At the request of the Riksdag, he made three long journeys in the 1740s to different provinces — to the islands of Öland and Gotland in the Baltic, to Västergötland in the west, and to Skåne in the southern Sweden. The aim was to inventory and list the "utilities," i.e., natural resources, ores and tree species, edible plants and berries, animals that were suitable as food or had good pelts, waterways that could be used to power mills, anything that could be put to economic use (1, pp. x–xi).

Linnaeus wrote of the wonders of wide open landscapes, and of plants, and insects that filled him with awe and religious inspiration, but he was able to combine his feelings of reverence with a commitment to putting these wonderful natural resources to use. Linnaeus was a biotechnologist of mind in a very modern sense. In a way he was also a forerunner of the contemporary interest in social aspects of the implementation of biotechnology in different sectors of society. He said that it was not enough to understand the wonders and utilities of natural resources, one must also understand the society in which the new science is going to be implemented. Accordingly

Linnaeus devoted a great deal of his time to describing the local customs and traditions he encountered on his journeys. He was interested in how people lived, what kind of food they ate, their housing conditions, and their cultural values.

The Linnaeus tradition of combining passionate interest and erudition in biology with concern for social and cultural values is not only a source of inspiration for Scandinavian scholars and scientists, but it has international significance. One might add the value of also acknowledging the limits of biological understanding as an explanation of complex human behavior and social phenomena. This, however, is not part of the Linnaeus heritage. He was too much of an optimist and an ardent adherent to the faith in progress that characterized science in the eighteenth century. The respect for the limits of science and circumspection with regard to what can be rightfully claimed as knowledge is Kantian. As Frängsmyr observes, Linnaeus was not a man of the Enlightenment in the philosophical sense (2).

Scandinavia is rich in natural resources, with long-standing traditions in fishing, forestry, and agriculture. As may be expected, plant and animal breeding have become focal points for strategic developments in biotechnology. Strategic research programs in many areas of biotechnology range from basic research in gene technology and protein engineering to the application of the new biological tools in agriculture, forestry, fishery, and medicine. As the international projects for the mapping of the human genome approach completion, major efforts are being made to develop technological platforms for understanding the function of proteins encoded by the discovered genes. Bioinformatic tools are being developed in order to help scientists at universities and in the biotech industry make use of the results in molecular biology and information technology.

Finland, Norway, and Sweden are well known for their large forests and their timber exports. Accordingly there is a special emphasis in these Scandinavian countries on the development of forest biotechnology. There is a chain of research groups in the universities and in the forest industry working on the biological production of forest biomass and tree fibers in cooperation with groups interested in the utilization of tree fibers for pulp and paper production. Plant biotechnology is the field in which the application of the research results most often meets the general public. The development of photosynthetic starch utilizes familiar products such as barley and potato. The use of transgenic tools in order to produce plant oils and fats that can replace fossil products are other examples generally accepted by the public.

Genetic research and the general use of biotechnologies have become intrinsic to many fields of medicine, in both clinical and nonclinical contexts. Genetic intervention has become a continually expanding concept referring to medical genetics as practiced by clinical geneticists, to laboratory pathology and clinical chemistry, and to many medical subspecialties in between. A number of initiatives in Scandinavia in genetic medicine with solid links to industry and academy have been taken in fields such as genetic diagnosis, gene therapy, pharmacological genomics, drug development, and nucleic acid research. One goal has been to open new avenues for the development of vaccines against viral, bacterial, and parasitic infections and their immunopathological consequences.

For centuries the populations of Finland, Iceland, Norway, and Sweden have been relatively homogeneous. Until the latter half of the twentieth century, immigration to these countries was rather limited. As a result many parts of these countries may be characterized as genetically isolated. This fact, together with the long-standing tradition of keeping records of the citizens in church books, in national health registers, and through social security numbers, has provided a strong impetus to research in population genetics and genetic medicine.

Universities in Scandinavia are governmentally funded, but close collaboration with industry in joint research and development structures is well established. Many companies take part in cooperative efforts with scientists and scholars at the universities. The figures vary among the Scandinavian countries, but it is estimated that as much as 30 to 60 percent of research funding at the universities comes from nongovernmental sources. This situation has given rise to an ongoing ethical discussion in the scientific community with industry and public authorities on the conflict of values related to vital research in medical and pharmaceuticals. The scientific community firmly upholds a principle of openness, arguing that research results should be made public for two reasons. First, it should be available to other scientists for them to build upon in their work. Second, accessibility to the results ensures a critical scrutiny of methods and scientific claims. With regard to clinical research, however, the universities lack the commercial means necessary for bringing scientific results into medical application in the form of new drugs and new treatment. In the interest of patient seeking cures, therefore, it seems that the much-celebrated principle of openness must be reconciled with commercial interests related to the seeking of patents (or immaterial rights). Accordingly, it has become increasingly common for university departments to sign contracts with industry in which they agree to postpone publication of results in order to secure funding for clinical and nonclinical research in biotechnology. The issue is not settled, however, and from time to time heated debate occurs over how to balance public and commercial interests in an ethically acceptable manner.

PUBLIC CONCERNS: WHY? OR WHY NOT?

Several surveys in the Scandinavian countries have revealed a skeptical attitude toward biotechnology among the general public (3–5). The average Scandinavian citizen has lower expectations of positive effect on everyday life of biotechnology compared with the average citizen in Europe. Significantly, however, public perception varies depending on the kind of biotechnological application involved. Genetic medicine receives the highest support, followed by the use of biotechnology for food production and plant and animal breeding.

There may be several explanations for Scandinavian skepticism toward biotechnology. There is a long-standing tradition of public environmental concern that may explain public resistance, at least in part. The general public wants justification in terms of likely benefits for society at large, as well as assurances concerning safety issues, before they approve of new technology. There have been rather frequent calls for moratoria on implementation, pending investigations into the consequences of biotechnological applications for different sectors of society. Recently a moratorium was proposed in Sweden regarding xenotransplantation. Proofs of estimated benefits and well-founded assurances of safety claims were demanded before a new technology could be accepted. This was what had happened with nuclear energy technology, and it seems to apply in biotechnology as well. While people in other regions of the world may be more accepting and open-minded, Scandinavians are inclined to ask why instead of why not, when confronted with this new technology. In the United States it might be the other way around (6).

The public surveys conducted in Scandinavia demonstrate that the common assumption that more information about biotechnology and its applications will result in a more favorable opinion toward this technology is false. Both scientists and industry often consider lack of knowledge to be the big factor behind low acceptance figures for biotechnology. A European survey from 1993, however, followed by a similar survey done in Norway in 1995, found high levels of knowledge together with low degrees of acceptance (3,7). More information, according to the analyses of these surveys, means more nuanced opinions among people, who are able then to make their own judgments about the various biotechnological applications. Applications within medicine are often highly regarded, whereas applications in the animal-breeding industry or in the production of genetically modified food products are met with much skepticism. According to the surveys, people have fundamental values that remain unchanged despite the availability of more information. Those who are skeptical from the start might well change their arguments, but they will not so easily change their attitudes. Nonetheless, it remains to be proved that there is no correlation between basic knowledge in biology and attitudes toward biotechnology. A common finding in the Scandinavian surveys is that within the population knowledge of biology, in general, and of molecular biology or genetics, in particular, is low. This may not be unique to the Scandinavian countries. Research and development in biotechnology challenges traditional biological concepts of the educated person. The skeptical attitudes toward biotechnology may be tied to old ideas about the biological world, ideas that have been made obsolete by new findings.

In Scandinavia opponents of biotechnology are often found among political and activist groups working with environmental issues. There are two issues that have galvanized resistance toward biotechnology and may be related to an insufficient knowledge of biology or to an outdated understanding of biology. The first deals with the view that natural ecological systems are fragile. The second takes up the idea of the sanctity of a species.

Value surveys among young people show a strong inclination toward environmental concerns (8). Fundamentally the concern for protection of the environment is connected with a resistance to biotechnology. Behind the resistance, there is a belief that the creation of transgenic plants and the release of genetically modified organisms will destroy the natural ecological balance, a balance that is said to be very fragile. Opposition to biotechnology is then an intrinsic part of a concern about changes to the natural ecological systems that must be based on a careful examination of likely consequences to the conditions of a specific ecosystem. This is not a controversial point. However, the more fragile the natural ecological system is, the more strictly this principle must be interpreted and applied. Anyone who has seen the local ecological effects of a discharge of an oil-tanker or who has seen how several plant and animal species have disappeared from the flora and fauna will notice that the ecological balance is very fragile. Against this background, the principle of caution must be applied in a very strict sense. But it is not clear that this is always true in the larger perspective. Seabirds die, and sea plants and sea microbes die from oil contamination, but the ecological system has recuperated rather quickly. Ecological systems may be vulnerable in one sense, but there is good evidence that they are robust even after devastating attacks on their balance (9). At the level of the ecosystem, it seems really not to be a matter so much of whether one species or another disappears but that the overall life-maintaining capacity must prevail. Ecological systems have further proved to be self-preserving in the sense that they do not easily admit new species to be created nor allow modified species to survive the complex steps of an organism's reproductive cycle. Against the background of this biological evidence, one can argue that nature as such is not fragile, and therefore the principle of caution should not be applied in a very strict sense. Depending on which stand one takes regarding the fragility of nature, one could have a different opinion regarding the release of genetically modified organisms in nature. If the biological knowledge of the dynamic capacity of ecological systems is communicated, acceptance figures related to biotechnology at large may improve, even if specific applications might be resisted for good biological reasons.

Respect for the natural boundaries among species seems to be the basis of the line of argument taken by opponents of the production and use of transgenic animals and plants for research, new food, or new medical substances. This view also sees nature as fragile and may be likewise based on a limited knowledge of biology. Aristotle was the first person to suggest that each species is determined by a specific idea. Aristotle's starting point was the Platonic theory of ideas, but he rejected the theory that ideas exist independently of the sense world. For Aristotle ideas are found in the phenomena of nature. They are present within the phenomena as teleological forces. Each species carries a fundamental purposiveness present as a formative power, a purposiveness that determines its characteristics and its relationships with other species. This fundamental purposiveness cannot be changed. Human beings must instead be attentive to this inner formative power of nature. Here one might recognize a certain Aristotelian inspiration in the opposition toward biotechnology that asks the scientists to

keep the boundaries of species sacrosanct (10). Research in plant and animal biotechnology entailing the creation of new transgenic organisms does not respect natural borders such as species barriers.

The Aristotelian conception of a species, however, has long been obsolete. After Gregor Mendel, the essentialist concept of species was replaced by a purely statistical concept. A species is the sum of an arbitrary selection of characteristics. A species does not carry an inner purposiveness but has a basis in a statistical description of its biological or practical purpose (11). This biological fact about life challenges the popular view held by many Scandinavians that the differences among species are sacred and must not be transgressed.

However, the more prevalent view of nature is that the present state of balance must not be disturbed. Humanity is not entitled to manipulate the genetic basis of life in cells. Among the general public, the ethical arguments are seen largely to respond to the idea of a normativity in nature. Particularly among the younger population the belief is that there is a "natural order of things" to which humanity ought to adjust (12). The values of the younger population reflect concerns about animal welfare and frequently refer to animal rights. Animals are believed to have natural rights the same as human beings. This attitude represents a shift in the value system of Scandinavia. Skepticism about biotechnology can be explained therefore, at least in part, by reference to beliefs of the kind described. Surveys in other parts of Europe, however, indicate that distrust of the biotech industry is based on its unwillingness to assume responsibility for environmental and safety concerns (13,14). Consumer organizations have pointed to the reluctance of the food industry to provide information to consumers on genetically modified products. Added to that is the distrust among the general public of the capacity of regulatory authorities to monitor these developments.

BIOBANKS AND THE DUAL INTERESTS OF THE CITIZEN

The Scandinavian National Health Service has increased dramatically the use of registers. Health information has long been available in the form of cancer registers, disability registers, cause-of-death registers, and health service registers. With the aid of social security numbers, these registers have proved to be valuable in social medicine and in epidemiological research. Developments in DNA research have brought to the register concept now further health information in the form of tissue cultures, tissue sections, and blood samples. The rise in the number of biological information banks, or biobanks, has become the focus of several governmental investigations in the Scandinavian countries. Denmark has taken the lead in proposing legislation to protect sensitive information and the value of privacy that are at stake in this practice (15).

The Scandinavian PKU (phenylketonuria) registers and biobanks are familiar to the general public, since a blood sample is taken from every newborn baby. During a workshop organized by the Nordic Committee on Biobanks in 1997, Bent Nørgaard-Pedersen from Statens Seruminstitut in Copenhagen described the development and potential of this kind of register (16). In Denmark, a nationwide screening of newborns for PKU has been carried out since 1975. Filter paper blood samples are taken 5 to 7 days after birth, and Guthrie blood tests are used for the analysis. The parents are informed about the sampling and the storage of the samples. They have the option of declining on behalf of their newborn babies, essentially to exercise an informed refusal. The agenda of the biobanks is to store laboratory and clinical data as well as the filter samples for future research. The biobank provides the valuable data needed for (*1*) diagnosis and treatment of phenylketonuria, (*2*) control of previously performed analyses, (*3*) quality assurance for diagnosis, (*4*) new analyses, in order to check for other diseases, and (*5*) research projects using biochemical, genetic, and environmental marker analyses. The register has been approved by the Health Ministry. All research projects using material or information from a biobank in Denmark must be approved both by a research ethics committee and, in special cases, by the national data surveillance authority.

Ethics committees play a central role in the regulation of the use of the biobanks for research purposes in all Scandinavian countries. They have a difficult task to strike a balance between the values that are at stake. To understand this function, a distinction must be made between two fundamental citizen concerns. On the one hand, the individual citizen as a potential patient has interest in the efficient storage and use of the biobanks as tools for new medical treatment and for the development of new drugs. From this viewpoint it is not good if one pharmaceutical company obtains exclusive rights to, and therefore a monopoly on, the information in a biobank. This would prevent other scientists and companies from working with the material and posing other scientific questions of potential interest for the health and well-being of the citizens. On the other hand, if a company is denied exclusive rights, it may not find it worthwhile to invest in the first place. Since the cost of research and development in the field of biotechnology greatly exceeds what most governments can afford to invest on their own, such a policy would not be beneficial to actual or future patients.

It is a difficult and delicate task for Scandinavian governments to find an ethically acceptable balance between these two extremes. The costs involved are also a challenge for the scientific community, health professionals, and the health authorities. They have yet to agree on efficient procedures for storing, sharing, and distributing biobank samples. Without rules and guidelines that take into consideration the interests of patients (actual and future), scientists, universities, and industry, there is a great risk that the doctors and scientists who control the freezers and the drawers containing the samples will not be able to cooperate and coordinate their efforts, both within and across national borders.

Against the interest of the citizen in efficient stewardship there is a no less important interest in protecting his or her integrity. The individual citizen must have sufficient safeguards to guarantee that the information contained in

the samples is not used in a way that is harmful to him or her. A problem is that no one yet fully understands the potential for the abuse of knowledge yielded by the medical information contained in tissue samples combined with hereditary and environmental factors. A particular problem is that the information acquired and processed is of relevance not only for the individual who is the source of the sample but also for genetic relatives of this individual. A stumbling block in the ethical discussion of biobanks is the formulation of necessary and sufficient rules of informed consent.

Article 22 of the Convention on Human Rights and Biomedicine by the Council of Europe from November 1996 states: "When in the course of an intervention any part of a human body is removed, it may be stored and used for a purpose other than that for which it was removed, only if this is done in conformity with appropriate information and consent procedures." This decision of the Convention has been the focus of intense discussion in Scandinavia in the effort to create ethically responsible legislation on biobanks. If the rule of informed consent is taken too rigidly, much of the epidemiological research will be precluded, with the consequence that advances in health improvement dependent on that knowledge will be lost. On the other hand, it must be acknowledged that the information gathered about the individuals might violate their integrity.

There is reason to believe that Scandinavia's long experience in providing medical benefits for patients by using health information registers will facilitate the application of a nuanced rule of informed consent that is sensitive to the values at stake for all concerned parties. In cases where no issues of integrity are at stake, or are inconsequential to the individual who is the origin of the sample, and the risks of harm are negligible, the consent procedures may be conducted in accordance with health and well-being as the primary objectives. The Danish model of informed refusal might be appropriate for some research protocols. In other cases, when more is at stake for the individual from which the sample has been taken, stricter rules, including written informed consent procedures, might be appropriate. If information and consent procedures are formulated too rigidly, they may be detrimental to the individual that the convention seeks to protect. There is a need for different information and consent procedures for different research and medical practices (17).

THE CAUTIOUS LEGISLATIVE APPROACH

Legislation affecting biotechnology is not the same among the different Scandinavian countries. Norway has enacted strict regulation. Sweden appointed a Gene-Ethics Committee as early as 1982, and its task was to conduct an inquiry into ethical, humanitarian, and social issues arising from genetic engineering. The report was completed in 1984 (18). After several years of public discussion, a law restricting research on fertilized human eggs was passed in 1991 (19). It stated a time limit of up to 14 days when research can be done on fertilized human eggs, and it allows eggs to be frozen and stored for up to one year. The storage time limit was later extended to five years (20). With some variation, there are now similar regulations in the other Scandinavian countries. The Swedish law is unique in one important aspect. The 1991 Swedish law not only prohibits the implantation of a fertilized egg that has been subject to experimentation, it also rules out all kinds of experiments directed toward altering heritable characteristics. Germ line gene therapy cannot be practiced; neither can it be studied for a potential future application. The law has prohibited the development of this research, and as such it is unique both in Scandinavia and presumably the rest of the world.

Laws have also been passed which regulate plant biotechnology and the release of genetically modified organisms (GMOs). On the whole, however, there have been relatively few attempts to regulate the field through legislation. One is given the impression that Scandinavian parliamentarians and governments do not want to create special legislation for biotechnological research and development. They first try to apply existing regulation for the safety of patients within the health care Acts or, with regard to ecological concerns, in the environmental acts (21). The Scandinavian countries also have long traditions of ethical review provided by specially appointed scientific-ethical committees and of public ethical debates. The scientific-ethical committees are well regarded and they are often asked by the legislators to take on a large responsibility in analyzing and judging the acceptability of research. Finland has until very recently relied exclusively on a system of ethical committees, instituted on a voluntary basis by the scientific community (the Finnish proposition RP 229/1998 on medical research). Denmark has legislation that supports the ethical committee system; a characteristic is the large presence of ordinary citizen representatives on the committees, although the legislators do not direct the review process normatively. Both Norway and Sweden are fundamentally dependent on a voluntary system. In Sweden, however, the recent parliamentary commission on research ethics has proposed that there should be a law stating that research on human subjects or human tissue should be the object of examination and approval by a research ethics committee from the university in question (22). In summary in the Scandinavian countries, legislation is believed to be too blunt an instrument for use on biotechnology, since developments occur at a rapid pace and new scientific facts are produced in an unending flow.

SPACE FOR SELF-REGULATION

Now and then there is public demand for stricter regulation. It turns out that the balancing mechanism is most often more moderate legislation and self-regulation on the part of scientists and industry. Part of the explanation behind this mechanism in the Scandinavian countries is the readiness of scientists to go public and express their own moral concerns. Until very recently, there have been few animal rights activists engaged in these activities such as in other parts of the

world. The American scientist Andrew Rowan (personal communication), a writer on the science and ethics of animal research, has suggested a plausible explanation for this difference (23). Rowan has noted that scientists in Scandinavia went public and expressed their moral concerns with regard to animal research almost from the start. They invited representatives of nongovernmental organizations to a public dialogue on the means and ends of scientific research. These animal ethics discussions led to review by specially appointed ethics committees that included a substantial number of lay members (24). The scientists expressed concern about animal welfare, but they could also explain the necessity of using animals as experimental models in order to provide cures for both humans and animals suffering from diseases. Thus the scientists achieved two goals: they brought ethical and policy issues related to research into the public debate, and they secured self-regulation for the scientific community.

The willingness of scientists and the biotech industry in Scandinavia to assume the moral responsibilities associated with a certain latitude for self-regulation can further be seen in the area of animal breeding and husbandry. Over many years, there has emerged a Nordic profile with regard to breeding goals in which the focus is not only on production traits but also on efforts to stabilize or improve the genetic level of functional traits related to animal welfare (25). According to this breeding ideology, increased birthweight of calves is not a value in itself, but must be related in a significant way to the health and well-being of the animals. Thus it comes as no surprise that the "Belgian Blue" (known as the monstrous bull because of its extremely large muscles) has met with great resistance not only from the public but also from the breeders in Scandinavia. Health, calving performance, quality of udder and teats, and fertility have long been vital breeding goals in addition to production traits.

Many scientists active in biotechnology research and development in Scandinavia seem to act in accordance with a maxim that has a certain Kantian ring to it (10). The maxim is general in nature, and does not specify any concrete goals of action. It is here suggested as the codification of a long tradition of moral thinking in bioscience and is also reflected in regulations created by the public authorities monitoring development in this field. Its role is as an aid in sorting out the value conflicts related to specific proceedings and applications of bioscience. The maxim may be formulated thus:

> Act in your biotechnology research and development so that you protect the health and well-being of human beings and animals; minimize their suffering; protect biological diversity; and make use of natural resources so that justice prevails and a contribution is made to a sustainable development.

PUBLIC CONSENSUS CONFERENCES IN DENMARK AND NORWAY

Among the Scandinavian countries Denmark and Norway have created their own ways of including public opinion in dialogues on biotechnology. The usual form a dialogue takes is for the bioscience community to engage the public in different activities related to popular science. These activities are important means for bridging the confidence gap between scientists and the public, but they are not enough. What is needed is a dialogue where the questions of the public are allowed to set the agenda and direct the discussion. This is what has been established in Denmark through the Danish Board of Technology for many years, and has recently emerged in Norway under the auspices of the National Committees for Research Ethics and the Norwegian Biotechnology Advisory Board. The early experiment of Denmark have been described and discussed at length within the European consensus conference context (26). These conferences are described as meetings that enable technology assessments to be made by an expert panel and a panel of concerned citizens (27). Issues like gene therapy, the handling of genetic information, and genetically modified food products have come under intensive discussion. The panel of citizens directs the proceedings and decides what questions to ask and what experts to engage. This panel then puts together a consensus report of their opinions regarding the questions discussed.

THE BILL ON MEDICAL DATABASES IN ICELAND

Iceland has passed an Act on a Health Sector Database and a Bill on Medical Databases that are particularly interesting for Scandinavia. The Icelandic population (270,000) is considered to be optimal for such a database for three main reasons. It is a homogenous population, health records are good and reliable, and the Icelanders are cooperative and positive toward research. These three features taken together are, to a great extent, shared by Finland, Norway, and Sweden.

The Health Sector Database, which draws on information from the entire Icelandic population, may be consulted for the purpose of discovering new drugs, developing new or improved methods for prognostic or diagnostic purposes and treatment of diseases, seeking the most effective solutions in the operation of health systems, or for medical reports or other comparable purposes in the health sector. The Bill elicited severe criticism from medical doctors, researchers, ethicists, and lawyers. The Ethics Council of the National Director of Health analyzed the implications of this Bill for the patient's right to privacy and for the relationship of confidentiality between doctors and patients. The Council has also worked on the ethical aspects of the intended bill on biosamples.

Public acceptance of biotechnology in Iceland is quite high. The criticism comes rather from professionals: the Bill on Medical Databases met with substantial resistance in the scientific community. A national survey in 1998 showed, however, that 75 percent of the Icelandic population is willing to have depersonalized information from their health records in a central database available for biotechnological research (V. Arnason, personal communication) (28).

ETHICAL, LEGAL, AND SOCIAL ASPECTS OF GENOME AND GENE TECHNOLOGY RESEARCH IN SWEDEN

A safe and wise implementation of genetic engineering and biotechnology in different sectors of society requires the

cooperation of scholars working in well-developed multidisciplinary research environments. In 1999 the Swedish Foundation for Strategic Research initiated a national research program with the aim of stimulating research focused on the ethical, legal and social implications of genome research and its implementation in different sectors of society (29). The following areas of research or fields of interest have been identified as examples of areas and fields that will be addressed in the program.

Public Perceptions and Values

It is not enough to be able to master the new gene technology tools. It is also important to understand the values and the formation of norms in the society where the new technology is integrated. As described at the beginning of this article, this perspective is part of the Linnaeus heritage. Values, attitudes, beliefs and worldviews reflect the larger impact of gene technology. Developments on genetic technology have challenged established concepts of health, illness, disease, and human integrity. Changes in concepts and changes in popular notions about these concepts are to be investigated. What are the underlying motives and views supporting certain beliefs about health and disease, about being human, or about the relationship between humans and animals? How do people classify their views? What are the general perceptions and evaluations of risks?

Genetic Medicine

Advances in molecular biology have provided medicine with powerful tools for diagnosis and for monitoring diseases and their treatment. Gene therapy may still be years away from becoming established medical practice, but several clinical trial protocols have been approved. The most urgent problems described in the research programs are related to genetic diagnosis and the handling of genetic information about individuals and families. The communication of risk and risk-related issues in relation to single-gene disorders is also an important area for research. Clinical practice is in need of the contributions of psychologists, anthropologist, ethnologists, and theologians who have the skills to map the wide spectrum of individual responses to the implications of genetic medicine. Polygenic and multifactorial conditions must be addressed besides the problems associated with the development of treatments for common genetic disorders. In this regard there are many questions about the rules and guidelines for the involvement of human subjects in clinical trials and related research, and for the storage and utilization of tissue samples in genetic research. As gene therapy gives way to cell therapy and the use of stem cell biology provides opportunities for the replacement of organs of the body, there will be calls for ethical and psychosocial interpretations of this development. If xeno-transplantation is considered, to what extent is implantation of an animal organ a violation of human dignity? What does it mean for human identity?

Genetic Engineering in Agriculture, Forestry, and Fishery

Agriculture, forestry, and fishery are areas in which gene technology will enable the development of new characteristics and breeding traits with a strong economic potential for big industry. The need is for research that focuses on practical problems encountered by scientists, industrialists, and policy makers in this area. Among the issues in need of attention there are many problems relating to four areas: (*1*) the intentional release of transgenic plants into the environment, (*2*) the introduction of genetically modified foods into the market place, (*3*) transgenic animal research, and (*4*) the impact of agriculture and plant biotechnology on developing countries.

From an ethical perspective, a society free of risks is neither possible nor desirable. Vital values in terms of survival, better food, health, and well-being are at stake, and these values have to be weighed and balanced against each other. The approach to risk assessment is expected to be normalization of risks (i.e., toxicological, allergenic, health hazards, ecological) through a comparison of risks associated with gene technology with the risks posed by conventional and future technologies for breeding and development.

In risk assessment there are six general ways of considering what can be expected. Preferably these questions will be taken together in regard to multidisciplinary projects.

1. What is the risk? (Potential risk identification)
2. How likely is the risk to occur? (Quantifying the probability of occurrence)
3. What is the severity and extent of the effect if it occurs? (Quantifying the effects)
4. What are the expected benefits?
5. Is the risk acceptable? (Normalization of the risk and risk–benefit analysis)
6. How important is risk assessment for different actors?

Implementation of Gene Technology

No one disputes the fact that gene technology will constitute an important economic sector in society. However, before the development of gene technology from science to commercialization can occur, several steps must, be taken. Even knowledge on how to go about this process is incomplete. On the one hand, gene technology could provide economic growth by stimulating the development of new industries. On the other hand, no companies have yet produced services and products of the anticipated quality. The gap between optimism and slow reality points to an equally important question on how companies that can harm the environment to create new products can achieve economic growth. Then there is still the problem of access to and use of genetic information, both commercially and noncommercially. All these concerns involve questions of intellectual property, questions of confidentiality and privacy, and questions of equity.

Within plant biotechnology and within pharmaceuticals, there has been a rapidly growing concentration of control in a few hands. Large corporations have bought out small companies. The issue of labeling is soon becoming a nonissue. Nongenetically modified soya is now more a thing of the past. The need is becoming urgent to clarify what is happening in biotechnology with regard to both

the scientific developments and the economic regulation of corporations taking over the developments.

The funding of plant and animal biotechnology needs to be investigated, likewise the funding in the production of medical products and diagnostics. Pharmaceutical companies have made enormous investments where there are indications of possible breakthroughs. Human diseases are expected to be treated with products not yet foreseen. Who will pay the bill? Who will be able to afford the new treatments? What will be the prioritization scheme in the health care sector? What price will be paid for success in biological and genome research?

GENETICS IN DIALOGUE WITH OTHER DISCIPLINES

It has long been obvious that there is a genetic history to many human diseases. Recent research, however, indicates that there is also a genetic component in more complex human behavior. If it can be proved that language behavior and learning ability express hereditary variability, the consequences will be great for established theories within disciplines such as linguistics and educational research. The old question about nature and nurture that has been so vividly discussed in relation to evolutionary theory is reactivated. Of particular interest is the development in brain research in which complex phenomena such as memory function and emotions may soon be comprehensively described in chemical terms. How is human society going to survive in a culture in which everyone is aware of the genetic components of one another? No doubt, in these new emerging fields of research geneticists, evolutionary theorists, linguists, philosophers and scholars in educational theories need to collaborate on such perplexing issues.

CLOSING REMARK

Clearly, any nation hoping to compete in the international economy must bring itself to the cutting edge of research and development in biotechnology. Without an investigation of the ethical and social implications, there is the danger that political decisions and legislation will be premature and not based on good, strong knowledge. There is also the risk that vital values related to health and survival will be neglected. In all five Scandinavian countries, there is a growing interest in the ethical implications of new technology. A number of ethics committees have been formed at various levels. Questions about the ethical implications of new technology are also recurrent themes in the media and in parliamentary debates. An interesting phenomenon in these discussions is that saying no to new technology is often believed to be of greater moral significance than saying yes. "Better safe than sorry" and "Safety first" seem to be the primary guiding principles of the debate on biotechnology. To be sure, when important issues are at stake such as survival, health, and well-being, there is great moral responsibility assumed in saying no or yes. A society free of risks is neither possible nor desirable, since, when vital values are at stake, these values must be balanced against one another. A "no" to biotechnology may deprive society of enormous benefits from research in bioscience and biotechnology. There may be good reasons to say no to certain problematic innovations. However, whether one says yes or no, this should be done only after a careful risk assessment and a weighing of the pros and cons.

ACKNOWLEDGMENT

I am grateful to Peter Sandoe, Per Sandberg, and Vilhjalmur Arnason for providing valuable background information and texts regarding ethical, legal, and policy issues in their countries. I am also grateful to Knut Erik Trany and Sharon Rider for valuable comments on an earlier version of this article.

BIBLIOGRAPHY

1. T. Frängsmyr, ed., *Linnaeus. The Man and His Work*, Science History Publications/USA, Division Watson Publishing International, Canton, MA, 1994.
2. A. Jeffner, *Biology and Religion as Interpreting Patterns of Human Life, The Idreos Lectures*, Harris Manchester College, Oxford, 1999.
3. T. Hviid Nielsen, in S. Lundin and M. Ideland, eds., *Gene Technology and the Public. An Interdisciplinary Perspective*, Nordic Academic Press, Lund, Sweden, 1997.
4. B. Fjaestad, ed., *Public Perceptions of Science, Biotechnology and a New University*, Mid Sweden University, Östersund, 1996, p. 10.
5. J. Durant, M.W. Bauer, and G. Gaskell, *Biotechnology in the Public Sphere. A European Sourcebook*, Science Museum, London, 1998.
6. P. Sandoe, *K. Skogs- Lantbruksakade. Tidskr.* **136**(20) (1997).
7. B. Nygård, *Ny bioteknologi i Europa*, Rapp. 1/1995, Senter for Bygdeforskning, Trondheim, Norway, 1995.
8. P. Ester, L. Halman, and R. de Moor, eds., *The Individualizing Society*, Tilburg University Press, Tilburg, The Netherlands, 1994.
9. B.G. Norton, *Why Preserve Natural Variety?* Princeton University Press, Princeton, Nd., 1987.
10. M.G. Hansson, *Human Dignity and Animal Well-being. A Kantian Contribution to Biomedical Ethics*, vol. 12, Acta Universitatis Upsaliensis, Uppsala Stud. Soc. Ethics, Uppsala, Sweden, 1991.
11. E. Sober, *From a Biological Point of View. Essays in Evolutionary Philosophy*, Cambridge University Press, New York, 1994.
12. C.R. Bråkenhielm and K. Westerlund, in S. Lundin and M. Ideland, eds., *Gene Technology and the Public. An Interdisciplinary Perspective*, Nordic Academic Press, Lund, Sweden, 1997.
13. S. Lundin and M. Ideland, eds., *Gene Technology and the Public. An Interdisciplinary Perspective*, Nordic Academic Press, Lund, Sweden, 1997.
14. A. Mauron and J.-M. Thévoz, *J. Med. Philos.* **16**(6), 649–666 (1991).
15. *Health Science Information Banks — Biobanks*, The Danish Medical Research Council, the Danish Central Scientific-Ethical Committee and the Danish Council of Ethics, Copenhagen, 1996.

16. B. Nørgaard-Pedersen, in M. Sorsa and J. Eyfjörd, eds., *Human Biobanks — Ethical and Social Issues*, Nord 1997:9, The Nordic Committee on Bioethics, Helsinki, 1997, pp. 59–67.
17. M.G. Hansson, *J. Med. Ethics* **24**(3), 182–187 (1998).
18. SOU, 1984:88, *Genetisk integritet. (Genetic Integrity)*, Betänkande av Gen-etikkommittén, Stockholm, 1984, Swedish Governmental Report 8F.
19. SFS 1991:115, (The Swedish Code of Statutes).
20. SFS 1998:282, (The Swedish Code of Statutes).
21. T. Achen, *Den bioetiske udfordring (The Challenge of Bioethics)*, Linköping Studies in Arts and Science, Linköping University, Sweden, 1997.
22. SOU, 1999:4, *God sed i forskningen (Good Research Practice)*, Stockholm, 1999, (Swedish Govermental Report).
23. A. Rowan, *Mice, Models and Men: A Critical Evaluation of Animal Research*, State University of New York Press, Albany, 1984.
24. B. Forsman, *Djurförsök. Forskningsteik, politik, epistemologi (Animal Research, Research Policy, Politics and Epistemiology)*, Almqvist & Wiksell International, Gothenburg, Sweden, 1992.
25. L.G. Christensen, *Acta Agric. Scand., Sect. A, Anim. Sci. Suppl.* **29**, 77–89 (1998).
26. S. Joss and J. Durant, eds., *Public Participation in Science. The Role of Consensus Conferences in Europé*, Science Museum, London, Sweden, 1995.
27. J. Grundahl, in S. Joss and J. Durant, eds., *Public Participation in Science. The Role of Consensus Conferences in Europé*, Science Museum, London, 1995.
28. J.R. Gulcher and K. Stefansson, *Nature (London)* **400**, 307–308 (1999).
29. Swedish Foundation for Strategic Research, Available at: *www.bioethics.uu.se/elsa*

See other INTERNATIONAL ASPECTS entries; INTERNATIONAL INTELLECTUAL PROPERTY ISSUES FOR BIOTECHNOLOGY.

INTERNATIONAL ASPECTS: NATIONAL PROFILES, SWITZERLAND

FRANÇOISE BIERI
ORESTE GHISALBA
OTHMAR KÄPPELI
HERBERT REUTIMANN
Swiss Priority Programme Biotechnology of the Swiss National Science Foundation
Basel, Switzerland

OUTLINE

INTRODUCTION

The 1980s were a turning point for biotechnology. Prior to 1980, biotechnology applications in Switzerland were limited to the industrial sector and a few, pioneering academic institutions such as the Swiss Federal Institutes of Technology in Zurich (ETHZ) and Lausanne (EPFL). In the 1980s, more and more leading Swiss institutions and Small and Medium size enterprises (SMEs) began using molecular biology and genetic engineering techniques for all types of applications in the life sciences. It soon became clear that the potential of biotechnology to benefit society was immense, if provided with the proper environment for its development.

Various groups and organizations in Switzerland, such as the Swiss Academy of Technical Sciences, the Swiss Coordination Committee for Biotechnology, the Board of the Swiss Federal Institutes of Technology, and the State Secretary for Science and Education, launched several proposals in order to induce national efforts for the promotion and development of biotechnology. In 1989, the Swiss Science Council (SSC) mandated the Swiss Coordination Committee for Biotechnology to perform a comparative study on national and international biotechnology R&D programs, their goals and development strategies. Swiss science policy makers used this document (1) to lay the foundation for the first nationwide biotechnology program, subsequently approved by the Swiss Parliament and initiated in 1992. This was the beginning of the Swiss Priority Programme Biotechnology (SPP BioTech).

PAVING THE WAY FOR THE DEVELOPMENT OF SWISS BIOTECHNOLOGY

Organization and Goals of SPP BioTech

SPP BioTech is financed through the Swiss National Science Foundation (SNSF). Its goal for 1992 through

2001 is to ensure the international competitiveness of Swiss biotechnological research and development (2,3). The program is applications-oriented, and it aims to bring basic and applied research closer to the development stage by encouraging synergistic collaborations among universities, research institutes, and private industry. Fields of biotechnology where Switzerland already holds a strong position are strengthened, while fields that need encouragement are given an opportunity to fortify their bases, through the setting of relevant research priorities that ease technology transfer in Switzerland.

From among the broad range of modern biotechnology applications possible, it would be impossible for a small country like Switzerland to fund all activities equally. Therefore SPP BioTech has created a number of modules based on a thorough assessment of the national research capacity and was able to consolidate applied biotechnology research in Switzerland as listed in Table 1.

In its goal of strengthening biotechnological research in Switzerland, SPP BioTech has not neglected the peripheral activities necessary for bringing technology innovations into society. SPP BioTech supports scientific activities that use modern biotechnology to help achieve sustainable development and efficient use of resources in industrial processes and agricultural systems. The program also recognizes the important role of continuing education in biotechnology for young researchers, and funds are accordingly allocated for Ph.D.s, postdocs, visiting scholars, and junior group leaders. In addition the program includes a unit of study concerning biotech-related issues that have significant interest for the public.

The SPP BioTech program has prioritized addressing public concerns regarding applications in biotechnology in a timely and informative manner. The level of public acceptance for technology applications can determine the speed at which development proceeds in certain critical research areas. It is for this reason that the agencies BATS (Biosafety Research and Assessment of Technology Impacts), BICS (Biotechnology Information Center), and Unitectra (technology transfer) were created under the aegis of the SPP BioTech.

The research activities within the SPP BioTech gradually proceeded from ideas and goal-oriented basic research to practical applications of the achieved results. The program comprised three distinct phases with a total budget of approximately 100 million Swiss francs:

Table 1. Research Modules of the SPP BioTech

Processes for the production and purification of proteins for medical applications
Biotechnology: bioengineering and biocatalysis
Food biotechnology (started in 1996)
Bioelectronics and neuro-informatics
Biosafety research and development of biotechnology
• Biotechnology Information and Communication (BICS Agency)
• Biosafety Research and Technology Assessment (BATS Agency)
• Technology Transfer (Unitectra)
Biotechnology of higher plants
SPP BioTech education program

- *Buildup phase, 1992 to 1995.* Focus on applications-oriented R&D, by introducing collaborative ventures involving universities, research institutions, and industries; begin technology transfer activities in the transfer of products, methods, and services.
- *Consolidation and Extension of Collaboration with Industry, 1996 to 1999.* Continue applications-oriented research and concentrate on successful strategies; extend and intensify contacts with industry; motivate SMEs to join; speed up technology transfer (including the creation of new SMEs).
- *Harvest and termination (outphasing), 2000 to 2001.* Continue the most successful and productive projects; focus on development aspects and technology transfer in order to exploit the achievements.

Participation in SPP BioTech has also helped a large number of research teams find easy access to Framework IV Programs of the European Union (EU). The success rate for Swiss applicants (36 percent for the first call in 1995) was by far above the European average (26 percent).

Achievements and Impacts of the SPP BioTech

Through SPP BioTech there have been created centers of competence and nationwide networks for biotechnology research (see Figure 2). Biotechnological activities at ETHZ and EPFL have been strengthened. An Institute for Neuro-Informatics, is now jointly operated by the University of Zurich and the Federal Institute of Technology, Zurich.

Further SPP BioTech has provided support for bioelectronics research and applications of this technology for the development of biomedical equipment. There has been created a nationwide network for Swiss biosafety research on recombinant and "naturally" occurring organisms. Universities and government institutions are closely collaborating in the field of plant biotechnology in developing a more sustainable agriculture.

Since 1996, SPP BioTech has promoted innovative research food biotechnology for healthier and safer dairy products. Technology transfer between academe and industry has been facilitated and given a priority.

Research Network in Biotechnology

A significant part of modern biotechnology research at Swiss universities occurs outside of the SPP BioTech. A survey carried out by Unitectra in 1997 (4) revealed more than 300 research groups active in various fields. There has been estimated, overall, between 350 and 400 biotech-oriented academic research groups in Switzerland (see Table 2).

Funding of Research in Biotechnology

At present, there are three types of public funding for biotechnology research in Switzerland (see Fig. 1):

- Funding of basic research projects (individual projects) directly via the Swiss National Science Foundation (SNSF).

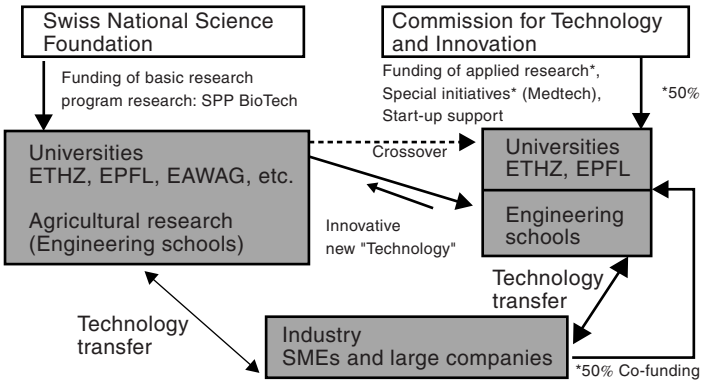

Figure 1. Funding, knowledge and technology flow.

Table 2. Main Areas of Interest of Research

Core area of nucleic acid technology (122)
Pharmaceutical biotechnology (for therapeutics and diagnostics) (69)
Agro/plant biotechnology (58)
Bioengineering, fermentation/reactor design (41)
Environmental biotechnology (27)
Bioinformatics (25)
Bioelectronics (23)
Biotransformation (22)
Biosafety (23)

Note: Number of research groups involved is given in parentheses.

- Funding of target-oriented program research (projects coordinated in units) via SNSF, with a strong emphasis on technology transfer at the precompetition level.
- Funding of applications-oriented R&D via the Commission for Technology and Innovation (CTI) (industry finances 50 percent of these projects).

In future developments of Swiss biotechnology, CTI will play a more important role. In 2001 SPP BioTech will be terminated. It can be assumed that a large number of SPP BioTech research teams will find new SNSF funding within the framework of the newly established National Centres of Competence in Research (NCCR), which are now in the evaluation phase. In the crucial attempts to organize smooth transitions, at present many researchers on SPP BioTech teams have already taken advantage of the extensive research network created by the SPP BioTech to access additional CTI and/or industrial funding.

USING RESEARCH RESULTS TO CREATE NEW JOBS

Technology Transfer — Universities Warming Up to Private Industry

The top researchers at Swiss academic institutions are a significant reservoir for of new inventions. Swiss academe thus provides opportunities for cooperative ventures with the private sector in the creation of start-up companies. This has been an important trend in the biotech industry worldwide, since many innovative concepts have emerged from the academic environment.

An example of the growing interest of companies in such ventures is the agreement signed in the fall of 1999 between Novartis, the Neuroscience Centre of the University of Zurich, and the Federal Institute of Technology Zurich. Under the agreement Novartis will fund research projects for up to 40 million Swiss Frances over a period of 10 years.

All Swiss universities are public. In recent years universities have further been given a high degree of administrative autonomy. In the course of these changes, ownership to all inventions resulting from research performed at the university has also been transferred from the state to the university. Corresponding laws are either in preparation or are enforced already, such as in Berne, Geneva, and Zurich.

Technology transfer has gained a lot of attention at universities in recent years and is strongly supported. Most universities in the meantime have established policies and technology transfer offices that support cooperative activities with the private sector, and provide support for faculty members on issues such as sponsored research agreements, the protection of intellectual property, licensing, and the creation of spin-off companies. Whereas technology transfer in the past was handled with mixed success by individual scientists, it is now being administered more professionally. Pragmatic and flexible guidelines for technology transfer are designed to facilitate interactions between academe and the private sector.

Increasingly academic researchers are considering the creation of spin-off companies as an interesting alternative or complement to their standard career paths. The high degree of entrepreneurial spirit among young academics is mainly due to the reduced job security in the large multinational companies that have all undergone significant restructuring in recent years due to the trends for globalization and the resulting mergers and acquisitions. Entrepreneurship at universities is even encouraged by various successful programs. A good example is Venture 98, a national business plan competition for university scientists organized by McKinsey & Company and the Federal Institute of Technology (ETH) in Zürich. More than 20 percent of the 215 projects submitted were in the field of biotechnology and life sciences, and they have spawned a number of companies. A similar subsequent program called Venture 2000 was launched in November 1999.

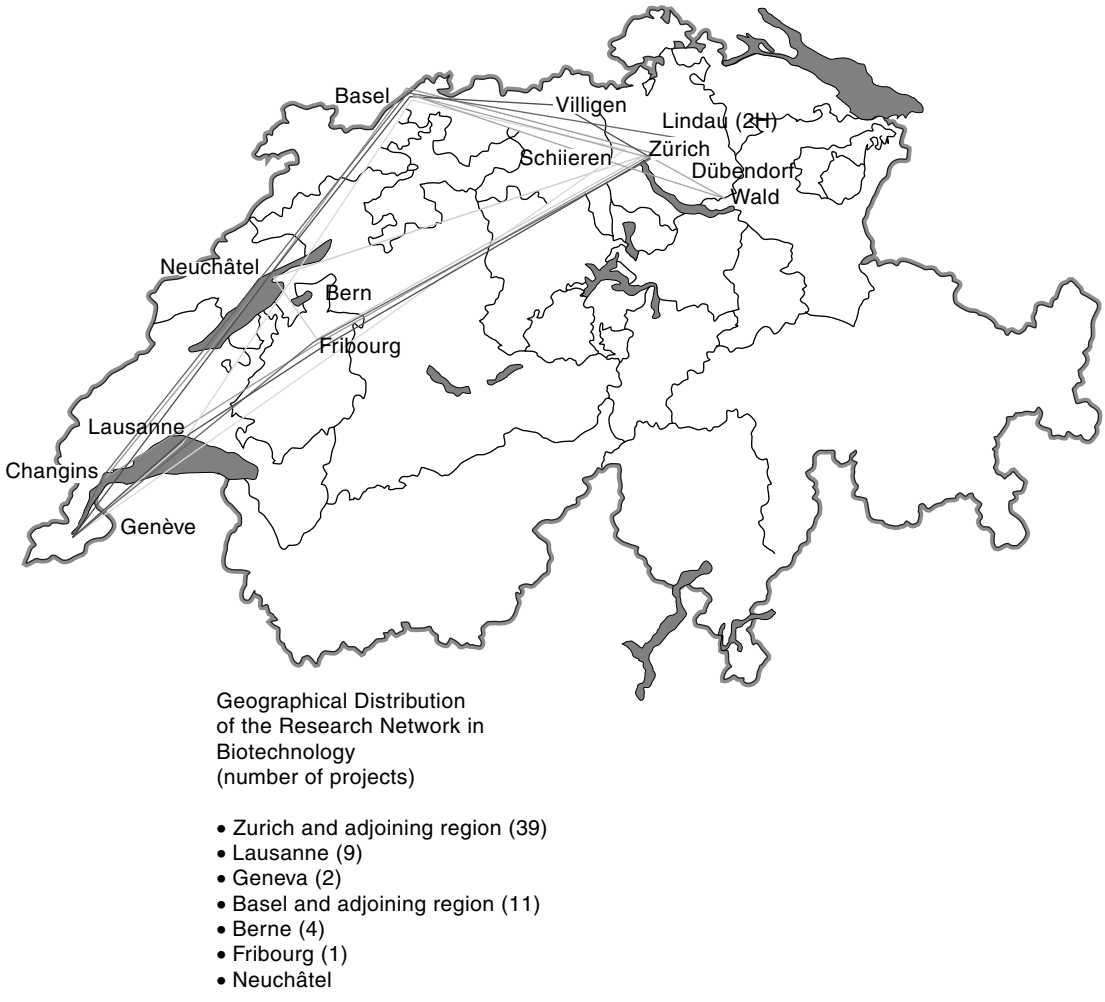

Figure 2. The Swiss Priority Programme. SPP BioTech was launched in 1992 with public funds. Six research areas in biotechnology and complementary activities in continuing education, information, communication, technology assessment, and technology transfer were designated to receive state support over a period of 10 years. The objective of the SPP BioTech is to consolidate strategic, applied biotechnology research in Switzerland.

Diversified and Rapidly Growing Swiss Biotech Industry

With a long tradition of economic strength in chemistry, Switzerland has added considerable expertise to bioscience in the past decades. The leading multinational drug companies Novartis, Roche, and Ares-Serono have, however, rather obscured the view of a very dynamic entrepreneurial bio-industry of small and medium-size enterprises. Unitectra, the technology transfer organization of the universities of Berne and Zurich and of the SPP BioTech, recently conducted the second comprehensive review of modern Swiss industrial biotechnology. This survey was published in September 1999 as *Biotechnology Industry Guide Switzerland* (5). It includes companies that meet the biotechnology definition of the European Federation of Biotechnology (EFB).

> Biotechnology is the integration of natural sciences and engineering sciences in order to achieve the application of organisms, cells, parts thereof, and molecular analogues for products and services.

Overall, the new directory lists 234 companies compared to 177 in the first edition which was published in 1996. Half of the companies (117) are classified as biotech companies, that is their main business focus is on biotechnology. The other half (117) are "other companies," that is, enterprises where biotechnology represents only one segment of their activities.

Forty-five percent of the listed companies are manufacturers of biotech products in Switzerland, 30 percent are suppliers or distributors, and about 20 percent are service companies. The fields of activity of the various companies are listed in Table 3.

The majority of companies can be grouped in three geographical clusters. The Zurich area has 89 companies, the Basel area 74 companies, and the region around Lake Geneva has about 30 companies. The total number of biotechnology-related jobs in these companies is estimated at 6500 to 7000. This is more than three times as many as in the United States on a per capita basis. Behind these figures lies a typical feature of the Swiss biotechnology industry. In Basel companies are based around

Table 3. Number of Companies in Different Fields of Activity

Agriculture	6
Analytical services/quality control	10
Biomaterials	3
Bioreactors/equipment/engineering	31
Bioelectronics/bio-informatics	9
Bioseparations/down stream processing	13
Cell culture	11
Chemicals (specialty/commodity)	8
Consulting	18
Contract R&D/contract manufacturing	16
Cosmetics/health/beauty products	2
Diagnostics	25
Environmental treatment/waste disposal	9
Fermentation/production	4
Food	9
Laboratory equipment	59
Medical devices	4
Pharmaceuticals/therapeutics/vaccines	26
Platform technologies	16
Reagents/biochemicals	29
Veterinary	2

Note: Some companies are active in several fields.

the chemical and pharmaceutical multinationals. In the Zurich and the Lake Geneva area academe provides the main impetus.

The new survey reveals a sharp increase in the number of entrepreneurial spin-off and start-up biotech companies created over the past three years. Without considering the numerous consulting firms, more than 40 new biotech companies were created. Two-thirds of these start-ups have a strong R&D focus mainly in the pharmaceutical area; the rest can be divided evenly into engineering and service companies, respectively. About one-third of the newly formed companies are typical university spin-offs, whereas others are spin-offs from large pharmaceutical companies. The high number of university spin-offs, in relation to the population and the number of universities, reflects the new entrepreneurial spirit among young scientists in academe. Moreover it also is the result of the increased support provided by various start-up programs.

Some examples of recent innovative start-up companies are:

Actelion has its focus of research on the endothelium, which constitutes the innermost layer of blood vessels and plays a role in cardiovascular diseases, inflammation, asthma, and many types of cancer. The aim of the company is the discovery and development of innovative drugs.

Biolytix is a young company located close to Basel providing services in the area of molecular biological analyses. It is specialised in the field of real-time quantitative PCR using state-of-the-art technology.

Biospectra develops and manufactures novel on-line analytical equipment and state-of-the-art automation solutions for fast bioprocess development and bioprocess optimisation.

Cytos Biotechnology, a spin-off company of the ETH Zürich, is developing new process solutions in the area of cell culture technology to optimise protein production. It also develops innovative solutions in other areas e.g. vaccines.

Modex Thérapeutiques is a university spin-off company based in Lausanne. Its focus is on new approaches to cell therapy for chronic systemic diseases, such as anaemia or diabetes. It plans to soon start with clinical trials on its first treatment regimen.

Prionics specializes in the detection of prions which cause BSE ("Mad cow disease") in cattle, Traber's disease in sheep, and Creutzfeld–Jakob disease in humans. It markets a bovine spongiform encephalopathy test for slaughtered cattle and sheep. Future R&D efforts also aim at prevention and therapy of prion diseases. Prionics was created as a spin-off company from the University of Zürich.

Zeptosens has its core competencies in the areas of advanced optical sensor and array technologies, bio-interaction analysis and bioassay design and development. It develops analytical platforms for the detection of analytes at extremely low concentrations. Typical applications include gene expression analysis, investigation of drug-receptor interactions, identification of bioactive compounds, and so on.

The rapid growth and development of the industrial biotech sector is confirmed by the annual *European Life Sciences* report of Ernst & Young (6). Although the absolute numbers in this report differ from the Unitectra survey because of slightly different definitions and inclusion criteria, the recently published report lists about more than 90 entrepreneurial life sciences companies (ELISCOs). This corresponds to an increase of 600 percent in a period of only three years (7).

The Association of Swiss Biotech Companies (AESB) established in March 1998 has already more than 100 member companies, mostly small and medium-size enterprises. It promotes biotechnology in the country and actively represent its members' interests in political and other circles. One of the AESB's essential goals is the facilitation of technology transfer between its members and universities. AESB will also advise foreign biotech companies looking for development opportunities in Switzerland.

FINANCE

Over the past decade Swiss industry has attracted considerable media attention through a series of biotechnology acquisitions and partnerships in the United States. This has been particularly true of large chemical/pharmaceutical corporations.

The new drive in Swiss biotechnology is reflected in the country's financial community. The last few years have seen the creation of a number of funds focusing on private equity and venture capital financing. There are currently more than 60 different funds operating in Switzerland and many of them explicitly seek opportunities in the biotech field.

Investments in venture capital in Switzerland reached 215 million Eurodollars in 1998, an increase of almost

Table 4. Selection of Swiss Investment Funds and Companies

Life science funds in Switzerland
BB Bioventures LP
Clariden Biotechnology Equity Fund
CS Equity Fund Pharma
Global Life Science LP
Lombard Odier Immunology Fund
Lombard Odier Nutrition Fund
Novartis Venture Fund
Pharma wHealth
Pictet Global Sector Fund–Biotech
UBS (Lux) Equity Fund/Biotech
UBS (Lux) Equity Fund/Health Care
Life-sciences investment companies listed at SWX
BB Biotech
BB Medtech
Micro Value
New Venturetec
Pharma Vision
Life-sciences venture capital companies
Alta Berkeley Associates SA
Angel Capital
Apax Partners & Co.
Aventic AG
Castle Private Equity AG
Euroventure-Genevest
Friedli Corporate Finance
Dr. Jürg F. Geigy
Invecor AG
MiniCap Technology Investment AG
New Capital AG
New Medical Technologies
Nextech Venture
Private Equity Holding AG

300 percent over the previous year. About one-third of the 86 projects supported were in the seminal or early stage phase, another third in the expansion phase. Most of the investments went into the high-tech sector. A number of biotech and life sciences companies are listed on the Swiss stock exchange (Swiss Exchange, SWX). SWX has introduced a new market segment in summer 1999 especially designed to meet the needs of young companies (Table 4). This adds another exit opportunity for investors.

However, the main shortage currently is not in finance but in skilled management with experience in setting up and running high-tech start-up companies. A number of initiatives are set to improve this issue such as the CTI Start-up Program and the recent formation of a Swiss Business Angels Club.

SAFETY AND TRANSPARENCY WHEN INTRODUCING TECHNOLOGICAL INNOVATIONS INTO SOCIETY

Successful Launch of Swiss Biosafety Research Network

The safety of technological applications is a prerequisite to their introduction into society. Swiss policy makers have recognized the importance of a publicly funded program for carrying out biosafety research and technology assessment, and view this as a service to society. A special unit on Biosafety Research and Technology Development was created by the SPP BioTech to address the safety aspects of biotechnological applications. A national agency for biosafety research, the agency BATS, was also created to coordinate research projects on the biosafety of transgenic organisms, as well as hazardous, naturally occurring organisms (Table 5). The safety of transgenic plants has been an area of intense activity, for which there is considerable effort invested in safety research and the development of methodologies for the safety assessment of open biological systems (8–10).

Table 5. Biosafety Research Projects

Viral recombination related to virus-resistant transgenic plants
Vertical gene flow
Biological containment for transgenes in plants
Ecological effects of transgenic plants
Horizontal gene transfer between plants and microorganisms, in aquatic systems and in the environment.
Fate of microorganisms in the soil
Projects related to health watch: monitoring for recombinant or pathogenic microorganisms, retroviruses, prions in water, food, and composts

Biosafety Research and Technology Assessment

Though important, safety is not the only criterion that is considered in evaluating technology applications. Decision makers in Switzerland rely on technology assessment (TA) for understanding the interrelationship of a technology or a product and society or its environment. In Switzerland, TA is coordinated by a central TA unit, founded in 1991, which was a time when the Federal Council and the Parliament decided that the Swiss Science Council (SSC) should develop a Swiss model for the assessment of the effects of technologies. In the field of biotechnology, the TA unit of SSC has coordinated several studies and organized a Publiforum, as the consensus conferences are called in this country. They are presently organizing a second Publiforum on xenotransplantation. The agency BATS is one of the partner institution of this federal TA program.

One example of a TA is a study that was carried out on the effects of potential widespread use of transgenic crops in Switzerland. This TA focused on the impact assessment in the ecological, toxicological, and economical dimensions (11,12). The culture of transgenic crops was compared to other agricultural alternatives, such as organic farming and integrated production, for each of the three impact dimensions mentioned. Another example of a TA project is a study of alternative agricultural production strategies for sustainability, based on ecological and economic indicators. From the information generated by this interdisciplinary effort, the existing scientific knowledge could be presented in a format that is useful to decision makers who define policy options related to transgenic crops.

Making Scientific Knowledge Available to Society

Access to reliable information is fundamental to good decision making on the personal and governmental level.

Members of the public require an adequate understanding of the meaning of new discoveries in order to make personal choices. Officials on all administrative levels need easily accessible knowledge and resources for the preparation of new legislation or for regulatory oversight.

The agencies BATS and BICS, granted by the SPP BioTech to provide information on all aspects of biotechnology, are both non profit and non lobby organizations. BICS publishes the unique Swiss quarterly review on Biotechnology, *BioTeCH forum*, available in a bilingual French/German edition. Other publications of BICS include facts sheets and brochures which are also available online and are a source of useful information not found in the media. The home page of the agency BICS allows the visitor an easy access to a vast selection of links covering all aspects of this field (13). The objectives of the Internet site developed by the agencies BATS and BICS are to (*1*) offer value-added knowledge on biotechnology impacts, and (*2*) pool and organize digital information for easy access (14). Contributors of information are research institutions, government agencies, and nongovernmental institutions. At this site the visitor can find information on a range of biotechnological applications. The information for the site is carefully gathered and checked for the quality of the source. In addition the retrieval of information is facilitated through a full-text retrieval system. Links are also given to other relevant sites. These sites are constantly being improved in order to serve the public better. A new feature of the *bioweb* site provides an interactive forum. Visitors to the site can discuss issues related to biotechnology with other citizens and a panel of scientists knowledgeable in the field.

REGULATORY FRAMEWORK AND PUBLIC DEBATE IN BIOTECHNOLOGY

Swiss Citizens Against a Ban on Genetic Engineering

Like their fellow European citizens, the Swiss are ambivalent about modern biotechnology. They have nevertheless acquired the distinction of being the first in the world to call for a national referendum, based on the complex technical and emotional issues surrounding genetic engineering. In June 1998 the Swiss were asked to vote on a constitutional prohibition of genetic engineering, therefore banning the use and patenting of transgenic animals and the deliberate release of transgenic animals, plants, or microorganisms into the environment. The political campaign leading to the vote lasted two and a half years and provided a unique opportunity for public and private organizations to hold informational meetings and public debates on the issues of genetic engineering. There was extensive media coverage of public debates, which helped to improve the overall public understanding of science, and to a significant extent, also the public acceptance for certain applications of genetic engineering. In the heat of the public discussion preceding the referendum, the Swiss Parliament also committed itself, in a motion called Gen-Lex, to enact a strict regulatory framework, in place of an all-out ban. On the day of the referendum, the Swiss people rejected a general ban on genetic engineering by a margin of 2 to 1.

At the time of writing, the Swiss government is still in the process of drafting the legislative framework regarding applications in genetic engineering. Nine existing laws pertaining to the various aspects of the use of genetically modified organisms (GMOs) and modern biotechnology are in the process of revision. Swiss legislation, based on the European Directives 90/219 on contained use of GMOs and 90/220 on deliberate release, has also introduced pioneering regulatory measures (14).

Prior to the drafting of the new legislation, a Federal Commission of Experts on Biosafety was created to oversee the use of GMOs. This was in the early 1980s, and the Commission followed the U.S. National Institutes of Health (NIH) guidelines on the safe use of GMOs. As a result, even in the absence of legislation, there have not been any abuses of GMO in Switzerland. In January 1997 a Swiss Expert Committee for Biosafety was created, as foreseen in the law on the protection of the environment and on epidemics. The role of this Expert Committee is to advise the administration regarding authorizations of field releases of transgenic organisms and on the drafting of the new legislation to encompass the newest knowledge in this field.

On November 1, 1999, three ordinances were enacted to regulate (*1*) the contained use of GMOs and pathogenic organisms, (*2*) safety at the workplace, and (*3*) the deliberate release of GMOs, including experimental field release or commercialization (15). Other aspects of genetic engineering applications are still not covered and are being hotly debated. In human medicine, these are genetic testing and xenotransplantation. Other contentious issues are liability insurance, intellectual property rights, and international trade, which must respect the guidelines of international agreements on trade and safety. The Swiss regulation for the labeling of foods derived from GMOs is similar to that of Europe. Labeling must be clear and unambiguous in Switzerland, with mention of the GMO origin, for example, in the list of ingredients. Chemically defined substances that are purified from GMOs and are free from traces of modified DNA or proteins do not have to be labeled. Switzerland was the first country in Europe to introduce, on July 1, 1999 (15,16), a threshold value of 1 percent for compulsory declaration. This means that any unintentional inclusion of a GMO equivalent during processing or transport of a product does not have to be declared, as long as the threshold is not surpassed. The threshold value of 0.5 percent for animal feed has also been legally accepted.

Although not legally prohibited, food derived from GMOs are deliberately kept off the shelves of department stores and groceries. No GMO-labeled product is currently sold in Switzerland because of the massive rejection by consumers of GMO-containing foods. The major food producers now ensure a sufficient stock from suppliers guaranteeing GMO-free crops. In addition environmental organizations act as watchdogs for the systematic monitoring of products suspected of containing unintended traces of GMOs. When a product tests positive for the presence of GMOs, it is withdrawn immediately from the shelves.

Swiss law does not prohibit deliberate releases, although a de facto moratorium currently exists in

Switzerland on the deliberate release of transgenic animals, plants, or microorganisms into the environment. In the spring of 1999 two experimental field releases were denied authorization by the Federal Administration despite the recommendations of the Commission of Experts on Biosafety to authorize the release, based on a thorough examination of safety. The final decision of the Federal Administration not to grant authorization was swayed by public opinion, which was against any type of deliberate releases of transgenic plants. At the time of writing, there has not been any field releases in Switzerland, experimental or commercial. Independently of this federal decision, the concluding opinion of a Publiforum on gene technology and food (June 1999) called for an official moratorium on the commercial cultivation of transgenic plants but not for experimental field releases. The discussion on how Switzerland will proceed in the future is still ongoing within the Federal Administration.

Ethics of Nonhuman Applications of Biotechnology

In Switzerland, public concern for the dignity of human and nonhuman organisms is taken very seriously. Switzerland is the only country in the world with a Federal Commission of Experts on Bioethics for the Nonhuman Applications of Biotechnology. The members of this Commission represent the various schools of thought in philosophy and ethics rather than lobby groups. This Commission works in conjunction with other Expert Commissions for biosafety, animal experimentation, and human applications of biotechnology (at the time of writing, this commission does not exist yet but is planned). The role of the Federal Commission on Bioethics is to advise the Swiss authorities and to provide them with criteria for a comprehensive evaluation of the ethical dimension of genetic engineering applications on nonhuman organisms. The elaboration of evaluation criteria in this new field is a fascinating and pioneering aspect of the Commission's work and will contribute to the legal recognition of the intrinsic value of animals and the environment.

Regulatory Framework for Biotechnology Applications on Humans

At the time of writing, several biotechnology applications on humans are close to being regulated and are hotly debated. Preimplantation diagnostics will most probably be banned in Switzerland by a law that is presently in preparation. In addition freezing of additional embryos, cloning, or any type of research work with human embryos or embryonic cells will most probably be banned, to avoid any abuse. Genetic testing is currently being debated because of its social implications beyond medical diagnosis. A public hearing has been organized for citizens to assess the various aspects of genetic diagnostics; Swiss legislators are currently consulting the report of the hearing.

Opinion on xenotransplantation is divided in Switzerland, as the biosafety and ethical aspects of xenotransplantation are highly controversial. Some people would like to see a restrictive legal framework, while others prefer a contingent ban. A Publiforum has been planned for citizens to discuss this issue.

The liability concerning transgenic plants or products derived from GMOs is currently under discussion. It is possible that the liability period for adverse effects will be extended from 10 to 30 years for transgenic organisms.

Technology, Law, and Society

The Swiss legislation that is under preparation attempts to accommodate the needs of the commercial sector, as well as public expectations on safety, information, and dialogue. During the years 1995 to 1998, the threat of a ban on genetic engineering applications generated a feeling of uncertainty that hampered business decision making, particularly in small and medium-size companies. Some scientists and entrepreneurs considered leaving Switzerland. Therefore the outcome of the vote in June 1998 has given the development of Swiss biotechnology a definite boost.

In the aftermath of the referendum in 1998, media coverage on bio- and gene technology issues remains passionate in Switzerland and is sustained by public opinion coming from abroad, such as on the Pusztai report of the possible adverse effects on rats after eating transgenic potatoes containing lectins or the adverse effects of Bt treated corn on the monarch butterfly. The experience of the political campaign prior to the referendum in 1998 has also propelled the scientists into the public debate.

Gen-Lex, the proposed legislative framework for biotechnology applications, is now undergoing the final step of approval in Switzerland. Applications such as experimentation with human embryos, xenotransplantation, sociopolitical consequences of gene diagnostics, and liability insurance on agriculture products derived from GMOs are the topics currently being debated in the Parliament and by federal agencies and the public.

Switzerland's first-rate university research and strong position in modern biotechnology have already produced important results in the areas of health, nutrition, environmental protection, raw materials, and specialty chemicals. In building firmly on its existing strengths, Switzerland is in a good position to keep pace internationally with future rapid developments in biotechnology. This will take place, however, with a firm commitment to the needs of society for safety, information, and innovation. The structure is now in place, through government policies, for promoting public dialogue at all stages of technology development.

ACKNOWLEDGMENTS

Dr. Lillian Auberson is gratefully acknowledged for editing the manuscript. The authors are financially supported by the Swiss National Science Foundation.

BIBLIOGRAPHY

1. O. Ghisalba and H. Vogel, *Frueherkennungstudie zur Biotechnologie*, Swiss Science Council, Bern, 1990, pp. 97, 154.
2. Swiss Priority Programme Biotechnology, *Paving the Way for the Development of Swiss Biotechnology, Swiss National Science Foundation*, Bern, 1999, p. 12.
3. *Prisma 99, Statistics on the Priority Programmes in Switzerland, Swiss National Science Foundation*, Bern, 1999, p. 160.

4. Biotechnology Research Compendium Switzerland, *SPP BioTech*, vol. 7, Unitectra, Basel, 1997.
5. *Biotechnology Industry Guide Switzerland*, 2nd ed., Unitectra and SPP BioTech, vol. 5, Unitectra, Basel, 1999.
6. *Ernst & Young's European Life Sciences 99, Communicating Value*, Ernst & Young International, London, 1999.
7. *Evolution, Ernst & Young Seventh Annual European Life Sciences Report 2000*, Ernst & Young International, London, 2000.
8. O. Käppeli and L. Auberson, *Trends Biotechnol.* **15**, 342–349 (1997).
9. O. Käppeli and L. Auberson, *Chimia* **52**, 137–142 (1998).
10. O. Käppeli and L. Auberson, *Trends Plant Sci.* **3**, 276–281 (1998).
11. E. Schulte and O. Käppeli, eds., *Gentechnisch veränderte krankheits- und schädlingsresistente Nutzpflanzen. Eine Option für die Landwirtschaft?* vol. 1, Swiss National Science Foundation, Bern, 1996.
12. E. Schulte and O. Käppeli, eds., *Gentechnisch veränderte krankheits- und schädlingsresistente Nutzpflanzen. Eine Option für die Landwirtschaft?* vol. 2, Swiss National Science Foundation, Bern, 1997.
13. *http://bics.ch*
14. *http://www.bioweb.ch*
15. *http://www.admin.ch/buwal*
16. *http://www.admin.ch/ch/f/rs/81.html* in french or *http://www.admin.ch/ch/d/sr/81.html* in German.

See other INTERNATIONAL ASPECTS entries; INTERNATIONAL INTELLECTUAL PROPERTY ISSUES FOR BIOTECHNOLOGY.

INTERNATIONAL ASPECTS: NATIONAL PROFILES, UNITED KINGDOM

RICHARD E. ASHCROFT
Imperial College
London, England

BENJAMIN CAPPS
RICHARD HUXTABLE
University of Bristol
Bristol, England

OUTLINE

INTRODUCTION

The term "biotechnology" covers "any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use." This definition appears in Article 2 of the Convention on Biological Diversity (1). Biotechnology touches our lives in many ways. It is instrumental in health care, food, the environment and agriculture, as well as such diverse fields as waste disposal, biomediation, and the promotion of more energy efficient, less polluting, and cheaper production processes. The diversity of biotechnologies, and their relevance to so many different spheres of human concern, is reflected in the breadth of different kinds of regulatory instrument and the range different administrative responsibilities and structures in place in the United Kingdom of Great Britain and Northern Ireland (UK) at the present time. Similarly the relative novelty of many biotechnological methods, and the ethical uncertainty surrounding many them, is reflected in the mix of statute law, common law, advisory committees, and public policy in the UK at the present time. The British media and public have repeatedly indicated ambivalent concerns about much biotechnology, as is shown by ongoing debates over cloning, genetically modified (GM) foods, various reproductive medicine techniques, hunting and animal experimentation, to name but a few instances.

In order to understand the political and legislative structure of UK biotechnology regulation, we now briefly review UK law and policy making, and the relationship between this and European Community (EC) law and policy making.

UK Law

It is not strictly necessary to provide a detailed overview of the UK's system of law making, particularly as the UK regulatory scheme largely derives from two sources. For a general overview of the English legal system, see Ward (2). Those laws that possess relevance are not derived from the case law. Rather, in the context of biotechnology, various Acts of Parliament (or "statutes") and associated subordinated legislation govern the vast majority of the UK regulation. Essentially an Act of Parliament, once it has passed through Parliament (comprising the House of Commons and House of Lords) to become law, will bind those within the jurisdiction. Such Acts as apply to biotechnology often enable bodies other than Parliament to issue subordinate (or delegated) legislation. An example is the "statutory instrument," which enables a Minister to make a legally binding regulation. As we will witness

below, there is a proliferation of such regulations within the scheme for regulating biotechnology.

UK Law and European Law

Much of the impetus for the regulation in the UK has been provided by European legislation in the area. For a general overview of the European impact on UK law, consult Ward (2). Two sources of European legislation are of particular relevance. First, the UK has been a member state of the Council of Europe since its inception in 1949. The Council of Europe strives to establish Europe-wide standards on a range of issues. Bolstered by various authorities and institutional machinery (e.g., the European Court of Human Rights), the Council of Europe itself comprises the Committee of Ministers and the Parliamentary Assembly. The former is the decision-making body; it may lay down binding legislation such as Conventions, or adopt recommendations to governments. The latter is the deliberative body; it too may make recommendations or resolutions. The Council of Europe has issued a number of legislative documents relating to biotechnology and the most important of these will be noted.

Second, having joined the EC in 1972, EC law is of prime importance to the UK's regulation of biotechnology. Since the Treaty of Maastricht, the EC is but one of three "pillars" of the European Union (EU). For the most part, the following discussion relates to the EC (see Ref. 3). The EC treaties set out broad frameworks that must be fleshed out by more specific measures. This task is performed by the various EC institutions: the Court of Justice (ECJ), the Council, the Commission, and the Parliament. These bodies provide three types of secondary legislation: regulations, directives, and decisions.

Regulations are directly applicable in all member states, and are binding in their entirety. Directives differ because these may be implemented by those means chosen by the member states to which they apply. There are a number of such regulations and directives that relate to biotechnology, either directly or indirectly. Finally, decisions are addressed to a specific person or persons: for example, the decisions of the ECJ. This latter source of EC law does not warrant exhaustive study, as the numerous regulations and directives suffice to provide the EC perspective.

One other, albeit non-statutory, area of European standards-setting warrants mention. The European Committee for Standardisation (CEN) is responsible for the planning, drafting and adoption of European technical standards (with the exception of those pertaining to the electrotechnology and the telecommunications sectors). In Europe, CEN works in partnership with CENELEC — the European Committee for Electrotechnical Standardization and ETSI — the European Telecommunications Standards Institute.

Such technical specifications ensure compatibility between products; guarantee appropriate levels for their safety, quality, or efficiency; and provide the test methods necessary to establish conformity. Once the need for a European standard has been firmly established, and nonduplication of work verified (CEN may also use an international standard), an experts Technical Committee is established. CEN's remit is to promote voluntary technical harmonisation in Europe in conjunction with international bodies and its partners in Europe. This harmonization is designed to diminish trade barriers, promote safety, allow interoperability of products, systems and services, and to promote common technical understanding.

In essence, then, legislation in a particular area of biotechnology is governed through Acts of Parliament, Regulations and the relevant legislation from Europe. By way of example, [this example is discussed in greater detail later in this article (contained use of GMOs and deliberate release of GMOs).], Genetically Modified Organisms (GMOs) are controlled in the UK by a number of regulations, including the GMO (Contained Use) Regulations 1992 and the GMO (Deliberate Release) Regulations 1992. These three pieces of legislation were created under the Health and Safety Act 1974 (contained use) and the Environment Act 1992 (deliberate release). Within Europe there are two major Directives that relate to the use of GMOs and are implemented by the UK regulations: Directive 90/219 (contained use) and Directive 90/220 (deliberate release). To assist in the implementation of these Directives, CEN is drafting appropriate safety standards that will apply to Europe in the use and release of GMOs; these standards should further serve to harmonize work in this area.

UK REGULATION OF BIOTECHNOLOGY

Introduction

In this section we will outline the legislative and regulatory framework that exists at present in relation to biotechnology in the UK. The wide-ranging scope of biotechnology means that the regulatory bodies — which are both governmental and nongovernmental — often have blurred remits. Numerous committees, subcommittees, and groups advise these bodies. Because of the rapid advancement in this field, the existing bodies are often inadequately equipped to deal with specific issues. To compensate for this there is a practice of creating new statutory and ad hoc committees to fill in the gaps that develop in the framework. Additionally, the government is also advised through independent bodies, such as the Royal Society, which periodically produce reports on matters of interest.

The regulatory and advisory framework, according to the UK government, has two distinct functions: to consider whether to grant approvals for individual products or processes (based on ethical, legal, regulatory, and scientific criteria), and to set a strategic framework for development of the technology in the UK (4). The main measures to provide safeguards against any real or hypothetical risks in biotechnological products are rigorous pre-market assessment of safety, research to improve scientific understanding of the particular product, and health surveillance to provide reassurance against any unexpected adverse effects. There is a concentrated effort to ensure that the governmental review of, and

action on, issues relating to biotechnology are sufficiently transparent. Transparency is urged, in order to facilitate input from interested parties (such as the public and concerned industries) on matters of policy.

The regulation of biotechnology in the UK is divided into five areas of responsibility under a lead governmental departmental body. As observed, the scope of these areas is often blurred, with certain issues falling within the jurisdiction of one or more of the responsible organizations. The areas of legislative responsibility are divided among: the Department of Trade and Industry (DTI) (consumer safety, product liability, trading standards, and patents); the Health and Safety Executive (HSE) (health of biotechnology workers, the control of hazardous substances and the contained use of GMOs); the Department of the Environment, Transport and the Regions (DETR) (release and marketing of GMOs into the environment); the Ministry for Agriculture, Fisheries and Food (MAFF) (food safety and labeling, animal feed and veterinary medicines, and plant health and pesticides); and the Department of Health (DOH) (therapeutic medicinal products, medical devices, gene therapy, and medicines licensing). These same groups are responsible for specific guidance and advice, access to funding, and research and expert services. This additionally includes the input of research councils, universities and trade associations.

The various statutory and advisory committees that exist within these five areas regulate and provide advice on the safety and broader impact of biotechnology. In general, the issues fall within the remit of, on the one hand, food and agriculture and, on the other, medicines and therapeutics. Some bodies, however, such as the Advisory Committee on Genetic Modification and the Advisory Committee on Release into the Environment, involve themselves in each of the two areas. There are essentially two types of committee: those that are established (either statutorily or ad hoc) by the government to specifically address issues arising from developments in biotechnology, and those that are not biotechnology specific, but nevertheless undertake significant amounts of biotechnology based work. There are other groups, such as English Nature, that only occasionally touch on biotechnology as part of a much wider area of interest.

Eight of the existing committees have a statutory function to advise Ministers on the exercise of their powers under specific pieces of legislation. These committees are the Animal Procedures Committee (APC), the Veterinary Products Committee (VPC), the Human Fertilisation and Embryology Committee (HFEA), the Advisory Committee on Release into the Environment (ACRE), the Advisory Committee on Pesticides (ACP), the Food Advisory Committee (FAC), the Committee on Safety of Medicines (CSM) (which also advises the Medicines Control Agency, (MCA)), and English Nature (along with its equivalents in Scotland, Wales, and Northern Ireland). The relevant Ministers, and MCA in the case of CSM, are statutorily required to take into account the advice of three of these committees when taking decisions. Thus the relevant Secretary of State consults ACRE; the recommendations of ACP must be taken into account by Ministers at the Ministry of Agriculture, Fisheries, and Food, the Department of the Environment, Transport, and the Regions, the Department of Health, and the Scottish and Welsh Offices; and MCA must consult the CSM.

In addition there are a number of nonstatutory committees. These offer advice on specific matters of interest or concern. These committees are the Advisory Committee on Novel Foods and Processes (ACNFP), the Farm Animal Welfare Council (FAWC), the Advisory Committee on Genetic Modification (ACGM), the Human Genetics Advisory Commission (HGAC), the UK Xenotransplantation Interim Regulatory Authority (UKXIRA), the Genetics and Insurance Committee (GAIC), the Gene Therapy Advisory Committee (GTAC), the Advisory Group on Scientific Advances in Genetics (AGSAG), and the Advisory Committee on Genetic Testing (ACGT).

The Department of Health

DOH is concerned with the health of humans, and it plays an active role in relation to biotechnology and the relevant UK legislation that governs the medical field and public health. DOH's main responsibilities concern the protection of public health from any possible hazards arising from the application of biotechnology. With this overarching objective in mind, DOH acts as a focal point in developing and coordinating policies, both national and international. The Department has legislative responsibility for developments in biotechnology in therapeutic medicinal products, medical devices and gene therapy also plays a role in the active encouragement of inward investment and sponsors the UK's biopharmaceutical industry. Research plays a large part within DOH and is divided into three main areas: the policy research program, NHS research and development strategy, and the research of nondepartmental public bodies. It is also involved in the European Community's Biomedicine and Health Research Programme.

DOH and the Office of Science and Technology (OST) jointly form the secretariat of the Human Genetics Advisory Commission (HGAC) which reports to and advises DOH and OST. HGAC remit is to "keep under review scientific progress at the frontiers of human genetics and related fields; to report on issues arising from new developments in human genetics that can be expected to have wider social, ethical and/or economic consequences, for example in relation to public health, insurance, patents and employment; and to advise on ways to build public confidence in, and understanding of, the new genetics" (5). It works alongside other committees that have an interest in human genetics. DOH also jointly forms the secretariat of the Advisory Committee on Novel Foods and Processes with the Ministry of Agriculture, Fisheries, and Food and provides members for the Inter-department Group on New Food Developments which covers feed intended for animal consumption.

The following discussion examines those various forms of biotechnology regulation which effectively operate under the auspices of DOH. Thus we address research, medicinal products, medical devices, and a variety of measures designed to survey the general field of human genetics. Before commencing the overview, two caveats

must be noted. First, although efforts have been made to distinguish and categorize the areas of interest, in an effort to avoid unnecessary duplication, naturally, such duplication cannot be entirely avoided. Second, and related to this first point, we should note one issue that will not be rehearsed: contained use and deliberate release. Here DOH does not have a statutory responsibility, although it is closely involved with the relevant independent expert advisory committees (ACGM and ACRE) (6).

Research. Research in DOH is divided among the Public Health Laboratory Service (which develops and implements epidemiological typing methods and laboratory diagnostics), the National Institute for Biological Standards, the Centre for Microbiology Research, and the Edward Jenner Institute for Vaccine Research. The Chief Scientist Office in the Scottish Office Department of Health and the Welsh Office Health Department are the overseers of research in Scotland and Wales, respectively.

From the point of view of medical research on humans, the most important bodies are the local and multi-center research ethics committees. The first Local Research Ethics Committees (LRECs) began to be set up in the mid-1960s, in emulation of the U.S. Institutional Review Boards. Initially these LRECs were locally originated, and without official status, although their pattern of constitution and process was defined by the Royal College of Physicians guidance on the Research Ethics Committees in 1967. Gradually the number of committees grew and acquired Department of Health recognition. There are now over 230 LRECs in the UK, regulated since 1991 by official Department of Health guidelines, and responsible to the area health authorities. Since 1997 any research protocol involving five or more centers has been reviewed in the first instance by a Multi-centre Research Ethics Committee (MREC). There are 10 MRECs, one each for Scotland and Wales, and one for each of the eight English health Regions. The MRECs are responsible to the Regional Health Authorities, and to the Department of Health centrally. For information and documentation on RECs in the UK, see Smith (7).

LRECs and MRECs are responsible for any research on (National Health Service) NHS patients, the recently dead in NHS premises, fetal tissue research on NHS premises, access to the medical records of NHS patients, and any other research on human beings that takes place on NHS premises. Their remit is to protect the subjects of such research, and to facilitate useful research; when these two aims conflict, the presumption is supposed to be that protection of individual patients takes priority (as required by the Declaration of Helsinki).

The constitutions of LRECs and MRECs are similar. Such committees must have a chair and a vice-chair, one of whom must be a lay person, together with at least eight members drawn from a range of professional backgrounds and including at least one other lay person. The committees are consensus-forming committees, and are not intended to be voting committees or "representative" in any but the broadest political or social sense.

Research approved by the MREC must also be reviewed by the relevant LRECs, but the LRECs in this situation can only consider four factors: the suitability of the site for the research, the suitability of the local investigator to do this research, the suitability of the local population to take part in this research, and the usability and comprehensibility of the patient information sheet to patients in this locality.

LRECs and MRECs are expected to take their decisions in the light of a fair process of discussion, and in the light of the best available written guidance (including that from government advisory committees such as the ACGT, the medical Royal Colleges, and international agencies such as the World Medical Association, the Council of the International Organisations of Medical Sciences, and the International Committee on Harmonisation of Good Clinical Practice).

Medicinal Products. Therapeutic medicinal products are primarily controlled by the Medicines Control Agency (MCA). MCA acts on behalf of Health Ministers and the UK licensing authority to issue marketing authorizations for medicinal products for human use and other licenses governing manufacture, clinical trials, wholesale dealing and parallel imports. This is based on the product reaching required levels of safety, efficiency, and quality. Independent advisory committees support the MCA in its tasks: of chief importance are the Committee on the Safety of Medicines and the Medicines Commission. The MCA is responsible for enforcing compliance with those authorisation provisions issued under the *Medicines Act 1968* (and its associated legislation, such as the Medicines for Human Use (Marketing Authorisations etc.) Regulations 1994) (8), and is assisted in this by the Royal Society of Great Britain, the DOH and the Social Services in Northern Ireland.

The Committee on Safety of Medicines (CSM) is a statutory body that advises the UK Licensing Authority (part of DOH) and MCA on the quality, efficiency, and safety of medicines in order to ensure that appropriate public health standards are met and maintained. A final important feature of the regulation in this context operates at the international level. The National Institute for Biological Standards and Control (NIBSC) oversees international standards in medicinal products. This is the executive arm of the National Biological Standards Board, which essentially exercises certain controls on the purity and potency of biological substances. Specialist medicines inspectors are responsible for the inspection of biological — including biotechnological — manufacturing sites.

The *Medicines Act 1968* and its associated legislation gives effect to the European medicines legislation, as initially laid down in *Directive 65/65* (9). A subsequent Directive in 1975 began to flesh out this basic framework in, among other matters, providing the data requirements for testing (10). Another Directive from 1975 established the Committee for Proprietary Medicinal Products (CPMP) (11), which now forms part of the European Agency for the Evaluation of Medicinal Products (EMEA). This latter agency controls the European authorisation of UK medicinal products. *Council Regulation 2309/93* set up the EMEA and a centralized procedure for biotechnological, and other "high-tech," medicines (12). This allows an authorization, through the EMEA, which is valid

throughout the Community. There is also a decentralized procedure for member states to mutually recognise each other's authorizations, which is also controlled through the EMEA.

Other Directives provide further amendment to, and elaboration on, the European legislative scheme relating to medicinal products. Those warranting particular mention are the laws governing specific products such as immunological products (13), radiopharmaceutical products (14) and products derived from human blood or human plasma (15), alongside more general measures dealing with good manufacturing practice (16), wholesale distribution (17), the classification of medicinal products (18), labels and leaflets (19), and advertising (20).

Medical Devices. Medical devices "are those diverse and extensive products, other than medicines, which are used in the healthcare field for the prevention, diagnosis, monitoring and treatment of disease and injury" (21). These are controlled by the Medical Devices Agency (MDA). This body is responsible for ensuring that medical devices and equipment for sale or use in the UK meet acceptable standards of safety, quality, and effectiveness and that these standards comply with the relevant EC Directives. At present, there are three directives that regulate the safety and marketing of medical devices throughout the EU. These are Directive 90/385 (governing active implantable medical devices) (22), Directive 93/42 (which covers all medical devices except those covered by Directive 90/385 and medical devices for in vitro diagnostics) (23), and Directive 98/79 (governing in vitro diagnostic medical devices) (24). The first of these Directives, which effectively covers such devices as heart pacemakers and cochlear implants, finds expression in UK law in the Active Implantable Medical Devices Regulations 1992 (25), as amended by the Active Implantable Medical Devices (Amendment and Transitional Provisions) Regulations 1995 (26). The second Directive, which has a broader scope, is implemented by the *Medical Devices Regulations* 1994 (27). It appears that the third of these Directives has yet to be implemented.

A number of aims and themes can be identified in this legislation. Emphasis is placed on the requirements that devices must not compromise the health and safety of the patient, user or any other person, and that the risks associated with the device must remain compatible with the patients health and protection. In order to achieve these aims, a number of specific requirements must be met. Thus clinical investigations are to be carried out, adverse incidents must be reported, devices must be classified and controlled according to the degree of risk inherent in their application, and monitoring must occur in order to ensure compliance with the requirements.

Gene Therapy. The UK's approach to the control of gene therapy has been informed by the 1992 *Clothier Report* (28). The Clothier Committee observed that gene therapy should be regarded as research involving human subjects. It concluded, inter alia, that research in the area should be restricted to disorders that are life threatening or cause serious handicap and for which treatment is either unavailable or unsatisfactory. The Committee decided that, for the time being, no attempt should be made to intervene in germ line cells. Finally, the recommendation a national supervisory body be established to consider and advise on the acceptability of gene therapy protocols resulted in the establishment of the Gene Therapy Advisory Committee (GTAC). GTAC is responsible for both the case-by-case review of individual protocols and an assessment of more general issues relating to such therapy. In addition GTAC provides advice to Health Ministers on developments in this field and on their implications. As well as working closely with LRECs, GTAC works closely with the MCA. As previously noted under the *Medicines Act 1968* and *Directive 65/65* (as modified), the MCA has a responsibility for regulating the quality, safety, and efficiency of medical products and applications for their clinical trials. This possesses relevance for gene therapy. HSE and DETR are similarly involved when the scope of the gene therapy falls within their remit.

Genetic Testing. The Advisory Committee on Genetic Testing (ACGT) is a nonstatutory body that advises UK Health Ministers on developments on genetic testing, taking account of ethical, social, and scientific aspects. Established in 1996, its remit is to provide advice on developments in testing for genetic disorders and to establish requirements, especially in respect of efficiency and product information, to be met by manufactures and suppliers of genetic tests.

Xenotransplantation. The UK Xenotransplantation Interim Regulatory Authority (UKXIRA) is a nonstatutory body that provides the voluntary regulatory framework for biotechnology in the area of human genetics and xenotransplantation. At present there is no domestic legislation, although the need for primary legislation has been realised. Set up through the Advisory Group on the Ethics of Xenotransplantation, UKXIRA, with the assistance of the Committee on Dangerous Pathogens, reviews and assesses the safety and efficiency of xenotransplantation procedures. All treatments for patients in this field have to be approved by the Secretary of State (29). As for specific legislation, some aspects of xenotransplantation are covered by the Human Fertilisation and Embryology Act 1990 (including cell therapies and gene therapies involving viable tissue) (see also the section "Home Office").

Genetically modified (GM) animals created in the course of xenotransplantation research are disposed of (although some GM animals can be used as food) under the assistance of the ACNFP, FAC (for labeling issues) and ACRE. Any live GM animals used in containment will be subject to the GMO (Contained Use) Regulations 1992 (as amended 1996) and the GMO (Risk Assessments) (Records and Exemptions) Regulations 1996 with respect to environment risk assessment (see below). The control of animals under these regulations will be the responsibility of HSE and DETR, as advised by ACGM. Xenotransplantation protocols that involve animals are covered by the Animals (Scientific Procedures) Act 1986,

and UKXIRA works closely with the Home Office in this matter. It is widely accepted that primates should not be used in such procedures, although the possibility has not been entirely ruled out (30). As of yet, no xenotranplantations from animals to human beings have taken place in the UK, although, while being cautious in its policy, the government has by no means excluded the possibility (31).

Infertility Treatment. The Human Fertilisation and Embryology Act 1990 established the Human Fertilisation and Embryology Authority (HFEA). Another important, albeit nonbiotechnology specific, piece of legislation in the field of infertility is the *Surrogacy Arrangements Act 1985* (which essentially prohibits commercial surrogacy arrangements). Its primary function is to licence and monitor centres providing treatment, research and care in this field. The Act and the Authority are therefore concerned with the use of gametes and embryos; as for HFEA's more general functions, advice will be disseminated on issues arising from developments in assisted conception and associated research. Key concerns include the need for safe, efficient, and ethical advances in the field.

The Health and Safety Executive

The Health and Safety Executive (HSE) is responsible for the health and safety of workers (and others) engaged in biotechnology in Great Britain. In Northern Ireland, the Health and Safety Inspectorate of the Department of Economic Development is responsible.

In Great Britain, the Health and Safety Commission (HSC) aids the HSE in its duties. HSC, whose members are appointed by the Secretary of State for the Environment, Transport, and the Regions, therefore considers, and also develops, health and safety policy. HSE advises the HSC on the shaping of policy and is responsible for implementing it. HSE specialist inspectors provide advice on the areas in question, specifically on risk assessment and containment. In relation to biotechnology, HSE's main responsibilities concern the regulation of the contained use of GMOs and the implementation of general health and safety legislation. (HSE works with the DETR with regard to the related issue of deliberate release when the release has implications on the health and safety of individuals. The COSHH regulations deal with certain deliberate releases of GMOs.)

Contained Use of GMOs. The former responsibility derives from those requirements laid down in the Genetic Manipulation Regulations 1989 (32), and the GMO (Contained Use) Regulations 1992 (as amended) (33). These regulations were made under the *Health and Safety at Work Act 1974*. The regulations revoke the regulations from 1989 (32) and replace them insofar as they relate to contained use. The latter contained use legislation assesses risks to humans and the environment, and essentially derives from EC Directive 90/219 (34). The 1996 amendments maintain this dual aim, as well as including various other requirements. These include the need to keep records; the need to establish a local GM safety committee, the need to classify all activities and organisms used; the need to notify the HSE of first an intention to use premises for GM for the first time and second, an intention to engage in certain subsequent individual activities (and, in some cases, work must not begin without the HSE's prior consent); and, finally, the need to adopt adequate controls, including suitable containment measures. Since the amendment, DETR also plays a role — alongside the HSE — in the regulation of this area.

The contained use regulations are thus concerned with the health and safety of both workers involved in the contained use of genetic engineering and those members of the public who may come into contact with such work. HSE and HSC rely on advice from the Advisory Committee on Genetic Modification (ACGM). ACGM advises all relevant government departments on human health and the environmental aspects of the contained use of GMOs, including laboratory and industrial installations. It is not involved in product approval. ACGM is accordingly advised by a technical subcommittee formed to provide specialised technical advice on all aspects of the human and environmental safety of the contained use of GMOs. The Advisory Committee on Dangerous Pathogens (ACDP) and the Department of Health's Health Promotion Division Select Committee on Science and Technology (SCST) also have a role in the control of the contained use of GMOs. SCST is divided into three subcommittees: the Human Genetics Advisory Commission (nonstatutory advisory body that also advises OST), the National Screening Committee, and the Advisory Committee on Genetic Testing (nonstatutory advisory body). Other committees may or may not advise on specific issues of GMO contained use depending on their remit. Finally in relation to contained use, ACRE advises the HSE/HSC and any other bodies appropriate on the possible human consequences of releases into the environment.

General Health and Safety. The legislation focusing upon general health and safety includes the *Health and Safety at Work Act 1974* and the *Control of Substances Hazardous to Health (COSHH) Regulations 1999* (35). Here the risks to be assessed are those risks to humans. The *Health and Safety at Work Act 1974* applies to all persons at work in Great Britain, whether employees or self-employed. Its requirements cover biotechnology, including the application of genetic modification techniques. Under this Act, employers have a duty to ensure the health and safety of the employees and to ensure that the general public is not put at risk by the work. The COSHH regulations apply to biological agents, including those which have been genetically modified, that may cause an infection, allergy, toxicity, or otherwise cause a hazard to human health. The COSHH regulations implement those EC directives relating to the protection of workers from risks associated with biological agents (36). Employers must therefore assess the risk of working with certain biological agents, to adopt appropriate control measures, and to notify the HSE of work involving certain biological agents.

The Department of the Environment, Transport and Regions

The Department of Environment, Transport and the Regions (DETR) is responsible for the regulation of the deliberate release and marketing of GMOs in Great Britain, and in furtherance of its aims, it promotes an extensive research program into associated risks. Before analyzing the various bodies that work with DETR, the (distinct) position in Northern Ireland deserves mention. There the position is virtually identical to that in Great Britain, although the requisite notification must be made to the Department of Economic Development; it is enforced by its Health and Safety Inspectorate. The Department of the Environment for Northern Ireland controls legislation governing the release and marketing of GMOs.

Other bodies, however, also have a role to play in this context. In addition to DETR, the Ministry of Agriculture, Fisheries, and Food, the Scottish Office Agriculture, Environment and Fisheries Department, the Department of the Environment for Northern Ireland, and the Welsh Office Agriculture Department are implicated in regulating specific areas. In addition the Department of Health addresses those releases of GMOs that have an impact on human health. The Ministers are advised primarily by the Advisory Committee on Release to the Environment (ACRE) and the Advisory Committee on Genetic Modification (ACGM). With regard to food and the marketing of GMOs, the Food Advisory Committee (FAC) and the Advisory Committee on Novel Foods and Processes (ACNFP) have particular relevance. FAC advises Ministers on the exercise of powers in the *Food Safety Act* relating to the labeling, composition, and chemical safety of food. It also advises on general matters relating to food safety.

Turning to the particular role of DETR, the regulation of the release of GMOs is primarily covered by the Genetically Modified Organisms (Deliberate Release) Regulations 1992 (37), issued in accordance with the Environmental Protection Act 1990 (Part IV). Other domestic legislation that impacts upon biotechnology and the environment must also be noted (38). The latter Act sets out the offences and penalties which apply in the event of a breach of its requirements. Any release of GMOs, with a few specialized exceptions, into the environment must be approved by the Secretary of State for the Environment, Transport, and Regions (acting jointly with other appropriate Ministers). Its purpose is to minimize any damage to the environment or the public that might arise from the deliberate release or escape of GMOs. The Secretary of State is therefore empowered to revoke a consent and to take enforcement action. The *GMO (Deliberate Release) Regulations* are enforced jointly by DETR and HSC. The HSC, in turn, can direct the HSE inspectors to perform the delegated enforcement functions.

Both of the relevant pieces of legislation (i.e., the Environmental Protection Act 1990 and the GMO (Deliberate Release) Regulations 1992; note that parallel legislation exists in Northern Ireland) implement the EC *Directive 90/220* (39), which specifically addresses the deliberate release of GMOs. This Directive has been amended in the light of progress in relation to the (new, simplified) procedure for applications to release GM crop plants (40), and the technical progress made, for example, regarding the information requirements of GM higher plants (41). These amendments were implemented by the GMO (Deliberate Release) Regulations 1995 (Section 7.40). The EC has passed other legislation dealing with deliberate release, but this may best be dealt with in other contexts. It should also be noted that the Council of Europe is similarly committed to the safety of both humans and the environment, as evinced in a number of legislative documents (42).

Ministry of Agriculture, Fisheries and Food

The policy of the Ministry of Agriculture Fisheries and Food (MAFF) is to support the development and exploitation of biotechnology within the food and agriculture industries, while protecting people, livestock, crops, and the natural environment. Where applicable, MAFF is jointly responsible for regulations governing pesticides, plant health, veterinary medicines, food products and imports. The approach of MAFF is coordinated with other government departments. Thus, for example, it consults with ACNFP to provide guidance on, and to regulate the use of, GM food. ACNFP is also responsible for assessing all applications made under the EC regulations relating to novel food and novel food ingredients. HSE, DETR, and the DOH also consult MAFF on the contained use and deliberate release of GMOs and act jointly, where appropriate. Finally, in Northern Ireland, the Department of Agriculture for Northern Ireland coordinates its work with MAFF with regard to biotechnology as it applies to its jurisdiction.

Novel Foods. In the UK assessment of GM and other novel foods, Ministers are advised by the independent Advisory Committee on Novel Foods and Processes (ACNFP). This committee carries out safety assessments of individual novel foods as part of the pre-market approval scheme controlled by the EC. In carrying out such assessments, the ACNFP is assisted by other governmental advisory committees, such as the Committee on Medical Aspects of Food and Nutrition Policy, the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment, and the Food Advisory Committee on the labeling of GM foods.

MAFF operates under the Food Safety Act 1990 (which applies to England, Scotland, and Wales and is parallel to the Food Safety Act (NI) 1991), which controls food consumption in Northern Ireland. The legislation makes it an offence to render any food injurious to health by adding or using any article or substance, abstracting any constituent from the food, or by subjecting the food to any other processes or treatment. In the main, local authorities enforce those portions of the Act that relate to hygiene and health. Environmental Health Officers and Trading Standards Officers enforce the requirements governing labeling and composition. In the capital, the London borough and Metropolitan authorities carry out both of these sets of enforcement duties.

Evidently food safety is a key concern. This concern is also detectable at the more general European level, although competing interests have been cited. For

example, Reports from the Council of Europe concede that risks should be assessed and minimised, but observe that advances — specifically in biotechnology — might increase yields and therefore prosperity (43).

Nevertheless, in the specific field of biotechnology, food safety is the pervasive theme. The primary piece of European legislation relating to novel foods and novel food ingredients is EC Regulation 258/97 (44). The Commission has taken many decisions in this context, too numerous to mention. These decisions concern such plants and vegetables as swede, maize, and soya, and an overriding concern is that the products will not adversely affect health. The UK has provided for its enforcement in The Novel Foods and Novel Food Ingredients Regulations 1997 (45). The EC regulation created a comprehensive EU-wide regulatory framework controlling all aspects of GM crops in Europe, from seed to final product. Accordingly, member states cannot introduce their own requirements in this area without the agreement of the other countries of the Commission, who are advised by the EC Scientific Committee for Food (SCF). (The SCF also re-evaluates any additives if they are prepared significantly differently from the original evaluation.) The regulation introduced a statutory pre-market clearance system for all novel foods, including those produced by genetic modification, and it is binding on all member states. Under this regulation the safety of individual GM foods is assessed by all member states, and any differences of scientific opinion are resolved by reference to a number of scientific committees within the EC.

The primary regulation has since been the subject of a recommendation concerning the scientific aspects and the presentation of information relevant to a safety assessment (46). Materials not originally covered by the regulation have also been brought within its procedures. Accordingly Regulation (EC) No 1813/97 (47), which generally concerned labeling requirements, dealt with genetically modified soya and maize, which was originally approved for food use under Directive 90/220 (supra), prior to the novel foods regulation. Detailed rules relating to the labeling stipulations contained in the later regulation have since been laid down in *Council Regulation (EC) No 1139/98* (48). These regulations governing the labeling of GM foods enter UK law via the Food Labelling (Amendment) Regulations 1999 (49).

Animal Feedingstuffs. The Agriculture Act 1970 (as amended) governs the marketing of animal feed in the UK. The Act makes it an offence to sell any material for use as feed that contains any ingredient that is deleterious to animals and, secondly, to human beings, who consume the products of an animal fed with the material. The Feeding Stuffs Regulations 1995 (as amended) (50) implement those EC Directives. The regulations cover a permitted list of single-cell proteins in feedingstuffs (51) (this Directive may be extended to encompass novel feed material; at present, a voluntary scheme for the approval of new feed material is in operation in the UK); assessment of "certain products" used in animal nutrition (52); and assessment of additives used in animal feedingstuffs (53). These Directives set out permitted additives that are allowed to be used or present in animal feed. The Directives also laid down requirements governing the information that must be provided (54).

Veterinary Medicines. The manufacture, distribution, marketing, and administration of veterinary medicines are controlled by the Medicines Act 1968, in conjunction with the secondary legislation issued under it and other UK legislation implementing the apposite EC legislation. Veterinary medicines are also controlled by the Marketing Authorisations for Veterinary Medicinal Products Regulations 1994 (55). These, in implementing parts of Directive 81/851/EEC (as amended) (56) state that only veterinary medicinal products subject to a marketing authorisation valid in the UK may by placed on the UK market. The Agriculture Departments and the Department of Health and Social Security (Northern Ireland) enforce the provisions governing veterinary procedures, acting on behalf of the health and agriculture ministers. The Royal Pharmaceutical Society of Great Britain also undertakes responsibility for the enforcement of provisions relating to certain retail sales. The Veterinary Medicines Directorate (VMD) administers the control of veterinary medicines, on behalf of DOH and MAFF. These bodies are advised by the independent Veterinary Products Committee (VPC) (created under the Medicines Act 1968) on the safety, quality, and efficiency of veterinary medicines covered by the 1968 Act. The VMD monitor and regulate all veterinary products on the market, including postauthorization monitoring.

The Council's Regulation (EEC) 2309/93 establishes a European centralized authorization procedure for high technology products in veterinary medicine (57). Compliance with this procedure is obligatory for certain biotechnological products and for novel growth promoters. Under this regulation initial applications are made to the European Medicines Evaluation Agency. The application is then submitted to the EC Committee for Veterinary Medicinal Products, comprising representatives from all member states, which assesses the application for authorization throughout the Community.

Pesticides and Plant Health. The Control of Pesticides Regulations 1986 (COPR) (as amended) addresses the majority of pesticides (58). The responsibility for regulation in this area is divided between MAFF and HSE. MAFF deals with the approval of products for use in agriculture and horticulture and in food storage practice. HSE is concerned with products for use with regard to mainly nonagricultural and nonfood uses.

COPR, as it relates to the MAFF remit, is progressively being superseded by the body of legislation which implements EC law. MAFF is responsible for the domestic legislation that implements Directive 91/414/EC (59), which covers the placing on the market of plant protection products (broadly, agricultural pesticides) and the import of pesticides. Included within this legislation are the Plant Protection Products Regulations 1995 (as amended in 1996 and 1997) (60), the Plant Protection Products (Fees) Regulations 1995 (as amended in 1997) (61), and the Plant Protection Products (Basic Conditions) Regulations

1997 (62). The legislation also contains powers to control the import of pesticides.

Under this legislation, manufacturers seeking to gain approval for pesticide products must apply to the independent Advisory Committee on Pesticides (ACP), according to the Food and Environment Act 1985. It is expected that this legislation will be superseded by the recent EC Directive on biocidal products (63), which will have a larger scope than that presently under the control of the COPR.

The legislation is enforced by the HSE, local authorities, and agriculture departments. Local authorities are concerned with consumer aspects of the legislation (as overseen by Trading Standards Officers) and issues, including storage, which are not covered by HSE (as overseen by Environmental Health Officers). The agriculture departments enforce provisions relating to wildlife, including the impact of pesticides and of pesticide residues in the environment.

Finally with regard to plant protection measures, following European legislation (64), the Plant Health (Great Britain) Order 1993 (as amended) places restrictions on the import and movement within the EC, and keeping in Great Britain, of particular plant pests, including GM plant pests, plants and products (65). The Order further provides that no unauthorized person may engage in any activity that involves genetic modification of a plant pest without proper authorization. Licences to undertake such work are supplied by the Plant Health Division of MAFF and, in Scotland, by the Scottish Office Agriculture Environment and Fisheries Department.

Plant Breeders and Plant Varieties. Applications for plant breeders' rights and the National Listing of Varieties are governed by the Plant Varieties and Seeds Act 1964 (as amended), the Plant Breeders' Rights Regulations 1998 (66), and the Seeds (National Lists of Varieties) Regulations 1982 (as amended) (67).

Discussions are persisting on the most appropriate manner in which to embody the EC provisions in this context. Although a lengthy discussion of the European law in this context is unnecessary, a few points do warrant mention. Accordingly there are provisions relating to the marketing of GM material that require domestic adoption. The new EC Plant Varieties system (introduced on April 27, 1995) makes specific licensing provisions for essentially derived varieties (i.e., those produced from existing varieties using biotechnological techniques). A recent important Directive allows varieties to be marketed. Thus member states may authorize producers in their own territory to place GM materials on the market (68). Such authorization may be granted only if all appropriate measures have been taken to avoid adverse effects on human health and the environment (as determined in accordance with other Directives) (69). Other stipulations relate to such matters as the need for labels to identify GM products and the need to protect varieties threatened with genetic erosion. (The novel foods legislation is also taken into account.)

Animal Welfare. The Farm Animal Welfare Council (FAWC) advises MAFF Ministers on the welfare of farm animals on agricultural land, at market, in transit, and at place of slaughter. It can freely investigate, advise, and communicate with any outside body, including the European Commission and the public, on any legislative or any other changes that may be necessary in this context.

As for existing legislation in this context, the Animal Health Act 1981 provides for Ministers to control the spread of disease, and the Specified Animal Pathogens Order 1993 (SAPO) (70), prohibits the import of animal pathogens and carriers of pathogens except under licence. It appears that the prohibition essentially relates to Third World countries. Under the latter Order, GMOs require a licence regardless of their origin (i.e., whether or not they are from the Third World). The licence is granted by Agriculture Departments, as advised by the state veterinary service. With specific regard to fish, broadly similar requirements apply, as provided for in the Diseases of Fish Act 1983. (The approaches of the EC and the Council of Europe to animals (and specifically animal welfare), are discussed below.)

Department of Trade and Industry

The Department of Trade and Industry (DTI) is legislatively responsible for product liability, trading standards, and the Patent Office. DTI is also the lead sponsor department for biotechnology. DTI's Chemicals and Biotechnology Directorate works within the DTI and with other government departments on the regulation and general appraisal of issues surrounding competitiveness in biotechnology. DTI thus has the ultimate responsibility for championing the biotechnology industry in all aspects of governmental, European, and international policies that affect its competitiveness. For this reason DTI has strong links with both the industry and the regulatory bodies of the UK government. The Office of Science and Technology, part of DTI, is responsible for managing the science budget and coordinating government policies on science and technology.

Product Liability. The DTI regulates consumer safety through a number of specific Acts and Regulations, depending on the product. Among these are the General Product Safety Regulations 1994 (71), and those other measures that implement the relevant EC Directives. The 1994 regulations impose a general requirement for safety in all consumer products which have not already been comprehensively covered by extant, specific European product Directives, and British and European standards. Local Authority Trading Standards Officers enforce the regulations. There is also the Consumer Protection Act 1987, which implements the EC Directive on product liability.

Patenting. The central piece of legislation in the UK is the *Patents Act 1977*, which succeeded the prior European Patent Convention (72). In the context of biotechnology, it must be emphasized that animal or plant varieties and biological processes for the production of animals or plants not involving significant technical intervention *cannot* be patented. (However, certain interested parties are not left unprotected: for example,

plant breeders; see the discussion below.) Nevertheless, as certain biotechnological innovations do remain subject to patenting, a basic overview of the UK system must be provided.

As with other inventions, the granting of a UK patent for a biotechnological invention depends on satisfying certain criteria, including those of novelty, inventiveness, and industrial applicability. In order to obtain a patent for any invention, an application that clearly and fully discloses the invention must be filed. It has been recognized that in some cases it may not be possible to describe a microorganism in words. In such cases a culture of the microorganism must be deposited in a culture collection not later than the date of the filing of the application. Indeed, the National Culture Collections offer a number of services, including the supply, identification, and safe deposit of cultures for patent purposes. Once the applicant has satisfied certain procedural requirements, a patent may be granted.

In the UK an applicant may deal with the Patent Office or the European Patent Office. The systems are broadly similar, although differences do exist. It should be noted that a UK grant of a patent is effective only in the UK. Patents for other countries generally have to be obtained locally. It is possible under the Patent Co-operation Treaty to obtain patents in a number of countries through a single initial application. It is also possible to obtain patents in up to 18 European countries, including the UK, by a single application to the European Patent Office.

We do not need to further probe the UK system governing patents because, with regard to Europe, the position on patenting is due to undergo some revision. It has been recognized that the biotechnology industries require a secure and effective intellectual property regime. The EC has answered the calls for such a regime in a 1998 Directive (73). This constitutes a significant piece of binding legislation, particularly for biotechnology. Member states will have until July 2000 in which to ensure that their national patent law is consistent with the requirements laid down by the Directive. The Directive maintains the general position with regard to patenting and biotechnology. Notable additional requirements include the denial of patents to inventions whose commercial use would be "immoral" (e.g., where suffering may be caused to a GM animal without particular gains for humans or animals). The Parliament and Council have also issued a regulation that impacts upon patenting biotechnology (74). This regulation enables certification, and strives to overcome some of the difficulties surrounding the granting of patents. Nevertheless, the 1998 Directive is undoubtedly the most important development in the area. Finally, and in contrast to the EC, the Council of Europe has appeared less willing to perceive patents as the panacea for the troublesome issue of rights in biotechnological advances. In that context, intellectual property rights are still being debated (75).

Home Office

In relation to the Home Office, a particular area of interest is the use of animals in scientific procedures (see the discussion below). The Animal Procedures Committee (APC) advises the Home Secretary on this issue, under the Animal (Scientific Procedures) Act 1986. The Committee is bound to have regard for both the legitimate requirements of science and industry and the protection of animals from avoidable suffering and unnecessary use in scientific procedures.

Similar themes are detectable in the relevant European legislative documents (for obvious reasons, not every document will be cited). The Council of Europe has had much to say on the status and treatment of animals (76). A key convention from 1986 regarding animals used in experimentation was, to a large extent inspired, by Directive 86/609 from the EC (77). It is notable that the contents of the two documents are broadly similar in their concern with ethical issues, such as the welfare of animals. Nevertheless, there are some distinctive differences. Thus, for example, whereas the Directive is primarily interested in the harmonization of national laws, in order to avoid any distortions in the internal market, provides better explanations of "alternative methods," and is supported by an Advisory Committee, the convention, by way of contrast, directs increased attention to the ethical issues, such as animal rights and humankind's needs.

In relation to biotechnology, two issues seem to have particular relevance to animals. These are cloning and genetic modification. As to the former issue, the EC has called for strict controls, with a particular view to ensuring that harm — to humans, animals and the environment — is minimized (78). As for the latter issue, both the EC and the Council of Europe have devoted some considerable efforts to assessing the permissibility of transgenesis; it is evident that each the ethical — primarily welfare-related — issues (79).

LOOKING TO THE FUTURE OF BIOTECHNOLOGY REGULATION IN THE UK

It is clear from the developments surveyed below that the regulatory philosophy of the UK government is unlikely to change much in the near future. It is unlikely, for instance, that a single National Bioethics Commission will be set up, partly because of the complexity of existing relationships within the administrative structure of the state (as shown), partly because of some scepticism regarding the merits of "bioethics" as an academic or policy discipline within the UK, partly because the experience of international debates with countries that have such a commission do not lead UK commentators to hope for much from such a commission, partly because of such a commission's enormous work load, and partly because there is no pressure for change at present. Bioethics advice is normally seen as the province of scientific experts, together with individual profession-based insights from key members of religious confessions (especially the established Church of England), the legal professions, and some social scientists and philosophers.

The credibility of some of the recent advice (e.g., on GM foods) has come under attack of late, in the main because many expert advisors have become seen to be out of step with the public mood, and in part also because some advisors have been seen as *parti pris*, in virtue of

the intellectual or commercial links with the activities they regulate. This is a problem in particular for scientific advisors, as has been seen in the GM foods debate.

It is true that the present Blair administration, and the previous Major administration, public commitments have been made to "open government" and to accountability of governmental and quasi-governmental committees. Appointments to all advisory committees must now be opened to applications for membership, and stringent appointment procedures now apply. But many important groups (e.g., commissions of inquiry) remain committees appointed by the minister on advice of his or her civil servants. Accountability is secured by accountability to parliament, and to the press, rather than any more direct method (e.g., public meeting).

Moving on from processual considerations of how committees are set up, structured, and members appointed, it has been noted by some commentators that most UK committees apply some very specific kinds of ethical and regulatory considerations to the legitimation of research and technology. Most committees apply some sort of risk–benefit calculus, modified in the light of broad social or cultural concerns. The initial Warnock report on Human Fertilisation and Embryology was much criticized by philosophers and by members of conservative pressure groups for the way it tried to synthesize analytic philosophical argument with more intuitive ideas about right and wrong. The result was felt to be unsatisfactory both from a consequentialist point of view (which underlies the risk–benefit arguments) and from the mainstream religious viewpoints (because the risk–benefit arguments were felt to miss the point). Nonetheless, the Warnock committees recommendations were broadly taken up in the form of the Human Fertilisation and Embryology Act. The general methodology of Warnock is similar to that of all the chief regulatory committees; what differentiates them is the way they assess risk, the significance attached to particular risks, and the way in which socially expressed concerns are permitted to influence these considerations (80). This point of method has significant consequences for the politics of such decision making (Risk to whom? How much is tolerable and by whom? Which risks are considered?), not least in the essential informality of the way judgments about risk must be made. At least one chair of a major committee has indicated great dissatisfaction with this state of affairs, but it is unlikely that the philosophy of such committee decision making will change significantly unless a different but similarly mechanical means for taking decisions can be found (81).

One major change over the next few years will be the impact of research programs in bioethics and biolaw in the UK; recent funding initiatives by the Wellcome Trust and by the European Commission, among others, are likely to result in major research projects into both the sociology of public discourse on biotechnology, and the incorporation of public attitudes into bioethical decision making. It will be interesting to see how (and if) such research affects policy making in the more formal arena of central government. However, it is likely that the main source of advice will continue to be the scientific and medical communities, whether or not the public credibility of these sources of advice, qua impartial advice, rises or falls.

In the light of these considerations, then, the main structural changes to biotechnology regulation will be constitutional changes in the state, rather than in the organization of regulation, its deliberative processes, or its underlying philosophy. The two main changes in the next few years are increased devolution of power away from the UK government in London "down" toward the national and regional authorities in Scotland, Wales, Northern Ireland, London, and the English regions, and "up" toward the European Commission and the juristic functions of the European Courts.

Recent Changes: The Implications of Devolution

With the devolution to the Scottish, Welsh, and Northern Ireland Parliaments, certain changes will be made to the regulation of biotechnology in the UK and Northern Ireland. Assisted conception, genetics, and health and safety legislation will be reserved to the Westminster Parliament, and therefore, the makeup of the primary and secondary legislation-making process and the advisory structure will generally remain the same. All other biotechnology related fields will become the responsibility of the devolved parliaments. The Northern Ireland and Scotland Parliaments will gain the ability to make primary legislation under the new structure, while the Welsh Assembly will acquire a similar range of functions, including powers to make secondary legislation.

All the devolved administrations will have access to the existing committees, irrespective of whether legislation is devolved or not. Where legislation is devolved, the new administrations, once in place, will be required to decide whether they wish to use the existing committees or create an alternative. They will also possess powers with regard to appointments to existing committees that are effected by devolution. The FSA and the proposed HGC and AEBC will operate on a UK-wide basis, with the devolved administrations having a say in appointments.

Of the statutory committees, those that will remain reserved to Westminster will be HFEA, CSM (which primarily advises the MCA, which will itself have some of its remit devolved; the CSM itself will become a UK-wide committee), ACGM (the environmental aspect of this committee will become devolved), VPC, and APC. The nonstatutory committees will all remain reserved, with no changes to the ACGT, HGAC, AGSAG, and GTAC. The remaining four committees, GAIC, UKXIRA, FAWC, and ACAF, will be altered so as to report to all the UK administrations within a UK-wide remit. Four of the remaining five statutory groups, although operating under devolved legislation, will continue to advise Westminster, and the devolved administration, if requested to do so. English Nature will be replaced in its UK advisory role by the Joint Nature Conservation Committee.

Recent Changes to UK Biotechnology Regulation

In May 1999 a government report was published that reviewed the UK's regulatory and advisory structure in biotechnology (4). In the report the government confirmed its existing policies, which are to protect the health of the public and to protect the environment while ensuring that

the potential benefits of this technology are not denied to the British people. Through consultation with the public, industry and experts, it was decided that there was a need to introduce a new regulatory framework. The new framework is designed to be less complex and more transparent, reflect the broader ethical and environmental questions and views of potential stakeholders, and is sufficiently forward-looking to encompass the rapid developments in this field.

The new comprehensive strategic advisory structure will be headed by the Food Standards Agency (FSA), and two new nondepartmental public bodies, the Human Genetics Commission (HGC) and the Agriculture and Environment Biotechnology Commission (AEBC). The remit of these three bodies is to advise relevant ministers on issues, receptively, on food safety, genetic technologies, and their impact on humans, as well as all other aspects of biotechnology without a direct impact on humans and food safety or food standards. Many of the present advisory committees' work has been taken on by the commissions, while other committees and technical bodies will be involved in cross-boundary issues between the three commissions. The commissions are not be involved in case-by-case review. This is still the responsibility of the remaining specialist technical bodies (84).

The ultimate responsibility in this area will lie with the appropriate Ministers. They will be responsible for giving guidelines to regulatory and technical bodies, changing the regulatory and advisory structure, and granting individual consent decisions. To aid the Ministers, the framework of independent expert regulatory and advisory committees will still exist under the remit of three nondepartmental public commissions. Direct advice to Ministers will come from these three strategic advisory commissions. The commissions will lease with the regulatory and technical bodies to give Ministers an overall picture of issues in biotechnology. In addition they advise on future strategy, changes in the guidelines for advisory/regulatory bodies, broader issues including ethics, and gaps in the framework.

Proposed Changes in to the European Legislation

There are a number of changes being proposed at the European level. Some of these are only entering the earliest debate and report stages, while others have progressed to encompass a draft Directive, which needs only to be debated. Due to the usual constraints, this discussion of potential innovations will have to be restricted. One of the more important proposals relates to those EC Directives governing deliberate release. In 1998 a draft Directive was drawn up (82), which is proposed to amend the central Directive 90/220 in a number of significant ways. These include strengthening the provisions on environmental risk assessment and monitoring, streamlining some procedures, ensuring greater transparency and improved labeling requirements, and providing for the consultation of EU ethics and scientific committees. Advances have also been proposed in the specific of novel foods, specifically in relation to additives (83).

CONCLUSION

The main advantages of the UK regulatory system are that it is flexible and responsive to the rapid changes in the biotechnology world; it is accountable readily to parliament and the courts; and some sort of workable balance is maintained, for the most part, between the politicians, scientists, technologists and industrialists, industrial consumers, and public consumers of biotechnological goods. For the most part the existing system has maintained public credibility, despite such recent disasters as the BSE/CJD affair and the GM foods affair. To some extent, at least, these crises are seen as failures of government rather than failures of advice. But underlying this assignment of responsibility is a growing sense of the politicization of science and the interested nature of scientific advice, and a response to this politicization that has not been seen since the nuclear power debates of the 1960s and 1970s. The weaknesses of the system of regulation are then (perceived to be) vulnerability to regulatory capture and lack of responsiveness to public debate. There is a sense abroad that "public debate" is perceived by government and the scientific community as necessarily ill-informed and misled, and therefore to be ignored or "educated." At the same time, certain kinds of public concern are taken note of (as in the cloning debate). It is not clear which sorts of public concern get to dominate scientific concerns and which are dominated by scientific and policy concerns. What is clear is that the political agenda, if not regulatory practice, will be increasingly shaped by the new single-issue agendas prompted by biotechnological change.

BIBLIOGRAPHY

1. Convention on Biological Diversity, *Off. J. Eur. Commun.* **L 309**, 3–20 (1993).
2. R. Ward, *Walker and Walker's English Legal System*, 8th ed., Butterworth, London, 1998.
3. P. Craig and G. de Burca, *EU Law: Text, Cases, and Materials*, 2nd ed., Oxford University Press, Oxford, UK, 1998, p. 3.
4. Cabinet Office: Office of Science and Technology, *The Advisory and Regulatory Framework for Biotechnology: Report from the Government's Review*, London, Cabinet Office, 1999, p. 4.
5. See HGAC at http://www.dti.gov.uk/hgac/
6. *GMOs (Contained Use) Regulations 1992* (see HSE, section 2.3.1) and the *GMOs (Deliberate Release) Regulations 1992, 1995* (see DETR, section 2.4).
7. T. Smith, *Ethics in Medical Research*, Cambridge University Press, Cambridge, UK, 1999.
8. SI 1994 No. 3144.
9. Council *Directive 65/65/EEC* of 26 January 1965 on the approximation of provisions laid down by Law, Regulation or Administrative Action relating to proprietary medicinal products [*Off. J. Eur. Commun.* **022**, 0369–0373 (1965)].
10. Council *Directive 75/318/EEC* of 20 May 1975 on the approximation of the laws of Member States relating to analytical, pharmaco-toxicological and clinical standards and protocols in respect of the testing of proprietary medicinal products [*Off. J. Eur. Commun.* **L 147**, 0001–0012 (1975)].
11. Second Council *Directive 75/319/EEC* of 20 May 1975 on the approximation of provisions laid down by Law, Regulation

or Administrative Action relating to proprietary medicinal products [*Off. J. Eur. Commun.* **L 147**, 0013–0022 (1975)].

12. Council *Regulation (EEC) No 2309/93* of 22 July 1993 laying down Community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Agency for the Evaluation of Medicinal Products [*Off. J. Eur. Commun.* **L 214**, 0001–0021 (1993)].
13. Council *Directive 89/342/EEC* of 3 May 1989 extending the scope of Directives 65/65/EEC and 75/319/EEC and laying down additional provisions for immunological medicinal products consisting of vaccines, toxins or serums and allergens [*Off. J. Eur. Commun.* **L 142**, 0014–0015 (1989)].
14. Council *Directive 89/343/EEC* of 3 May 1989 extending the scope of Directives 65/65/EEC and 75/319/EEC and laying down additional provisions for radiopharmaceuticals [*Off. J. Eur. Commun.* **L 142**, 0016–0018 (1989)].
15. Council *Directive 89/381/EEC* of 14 June 1989 extending the scope of Directives 65/65/EEC and 75/319/EEC on the approximation of provisions laid down by Law, Regulation or Administrative Action relating to proprietary medicinal products and laying down special provisions for medicinal products derived from human blood or human plasma [*Off. J. Eur. Commun.* **L 181**, 0044–0046 (1989)].
16. Commission *Directive 91/356/EEC* of 13 June 1991 laying down the principles and guidelines of good manufacturing practice for medicinal products for human use [*Off. J. Eur. Commun.* **L 193**, 0030–0033 (1991)].
17. Council *Directive 92/25/EEC* of 31 March 1992 on the wholesale distribution of medicinal products for human use [*Off. J. Eur. Commun.* **L 113**, 0001–0004 (1992)].
18. Council *Directive 92/26/EEC* of 31 March 1992 concerning the classification for the supply of medicinal products for human use [*Off. J. Eur. Commun.* **L 113**, 0005–0007 (1992)].
19. Council *Directive 92/27/EEC* of 31 March 1992 on the labelling of medicinal products for human use and on package leaflets [*Off. J. Eur. Commun.* **L 113**, 0008–0012 (1992)].
20. Council *Directive 92/28/EEC* of 31 March 1992 on the advertising of medicinal products for human use [*Off. J. Eur. Commun.* **L 113**, 0013–0018 (1992)].
21. D. Longley, *Med. Law Inter.* **3**(4), 319–345 (1998).
22. Council *Directive 90/385/EEC* of 20 June 1990 on the approximation of the laws of the Member States relating to active implantable medical devices [*Off. J. Eur. Commun.* **L 189**, 0017–0036 (1990)].
23. Council *Directive 93/42/EEC* of 14 June 1993 concerning medical devices [*Off. J. Eur. Commun.* **L 169**, 0001–0043 (1993)].
24. *Directive 98/79/EC* of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices [*Off. J. Eur. Commun.* **L 331**, 0001–0037 (1998)].
25. SI 1992 No. 3146.
26. SI 1995 No. 1671.
27. SI 1994 No. 3017. (Note that these Regulations are pursuant to the *Consumer Protection Act 1987*.)
28. *Report of the Committee on the Ethics of Gene Therapy*, H.M. Stationary Office, London, 1992, Cm 1788.
29. See Health Service Circular HSC 1998/126 (*Clinical Procedures involving Xenotransplantation*), which utilizes powers within the *National Health Service Act 1977* and the *Community Care Act 1990*.
30. J.K. Mason and R.A. McCall Smith, *Law and Medical Ethics*, Butterworth, London, 1999, p. 340.
31. *The Government's Response to Animal Tissue into Humans: The Report of the Advisory Group on the Ethics of Xenotransplantation*, London, Advisory Group, 1997.
32. SI 1989 No. 1810.
33. SI 1992 No. 3217, SI 1995 No. 2626, SI 1996 No. 967, and SI 1998 No. 1548.
34. Council *Directive 90/219/EEC* of 23 April 1990 on the contained use of genetically modified micro-organisms [*Off. J. Eur. Commun.* **L 117**, 0001–0004 (1990)].
35. SI 1999 No. 437 (revoking SI 1994 No. 3246), which implements a number of EC Directives, including some of those noted below (n. 39). In Northern Ireland, see *Control of Substances Hazardous to Health (COSHH) Regulations*, Northern Ireland, 1994.
36. The primary piece of European legislation is Council *Directive 90/679/EEC* of November 26, 1990 on the protection of workers from risks related to exposure to biological agents at work (seventh individual Directive within the meaning of Article 16 (1) of *Directive 89/391/EEC*) [*Off. J. Eur. Commun.* **L 374**, 0001–0012 (1990)]. This Directive has been amended since. See Council *Directive 93/88/EEC* [*Off. J. Eur. Commun.* **L 268**, 0071–0082 (1993)], Commission *Directive 95/30/EC* [*Off. J. Eur. Commun.* **L 155**, 0041–0042 (1995)], Commission *Directive 97/59/EC* [*Off. J. Eur. Commun.* **L 282**, 0033–0035 (1997)], and Commission *Directive 97/65/EC* [*Off. J. Eur. Commun.* **L 335**, 0017–0018 (1997)].
37. SI 1992 No. 3217. This has since been amended by SI 1993 No. 152 and SI 1995 No. 304. See also *Genetically Modified Organisms (Deliberate Release and Risk Assessment–Amendments) Regulations 1997* (SI 1997 No. 1900).
38. See the *Health and Safety at Work Act 1994, Wildlife and Countryside Act 1981* (s. 14) (as amended in 1985), the *Nature Conservancy Council Act 1973*, the *Countryside Act 1968*, the *Food and Environment Protection Act 1985*, and the *Food Safety Act 1990*. More specifically, see the *GMO (Risk Assessment/Records and Exemptions) Regulations 1996* (SI 1996 No. 1106, as amended by SI 1997 No. 1900) (restricting import and acquisition of GMOs) and the *Environmental Information Regulations 1992* (SI 1992 No. 3240, implementing Council *Directive 90/313/EEC* of 7 June 1990 on the freedom of access to information on the environment [*Off. J. Eur. Commun.* **L 158**, 0056–0058 (1990)].
39. Council *Directive 90/220/EEC* of 23 April 1990 on the deliberate release into the environment of genetically modified organisms [*Off. J. Eur. Commun.* **L 117**, 0002–0024 (1990)].
40. Commission *Decision 94/730/EC* of 4 November 1994 establishing simplified procedures concerning the deliberate release into the environment of genetically modified plants pursuant to Article 6.5 of Council Directive 90/220/EEC (*Official Journal of the European Communities* L 292 of 12/11/1994 pp. 0031–0034). See also the Written Questions Nos. 2394/96 to 2405/96 by Hiltrud Breyer to the Commission, on establishing simplified procedures concerning the deliberate release into the environment of genetically modified plants [*Off. J. Eur. Commun.* **C 96**, 0005–0009 (1997)].
41. See, first, Commission *Directive 94/15/EC* of 15 April 1994 adapting to technical progress for the first time Council *Directive 90/220/EEC* on the deliberate release into the environment of genetically modified organisms [*Off. J. Eur. Commun.* **L 103**, 0020–0027 (1994)]. See also Commission *Directive 97/35/EC* of 18 June 1997 adapting to technical progress for the second time Council *Directive 90/220/EEC* on the deliberate release into the environment of genetically

modified organisms [*Off. J. Eur. Commun.* **L 169**, 0072–0073 (1997)].

42. See *Convention on Civil Liability for Damage Resulting from Activities Dangerous to the Environment* (ETS No. 150) of 21 June 1993, *Recommendation R (92) 9 of the Committee of Ministers to Member States* on the potential ecological impact of the contained use and deliberate release of genetically modified organisms, and *Resolution 870 (1986)* on the biogenetic revolution in agriculture — a blessing or a curse? (as essentially adopted in Recommendation 1213 (1993) on developments in biotechnology and the consequences for agriculture).
43. See *Report (Doc. 7943)* of 7 October 1997 on food supply in the world and *Report (Doc. 8194)* of 14 September 1998 of the Committee on Agriculture and Rural Development on consumer safety and food quality.
44. European Parliament and the Council of 27 January 1997 [*Off. J. Eur. Commun.* **L 43**, 0001–0007 (1999)].
45. SI 1997 Nos. 1335 and 1336.
46. Commission *Recommendation 97/618/EC* of 29 July 1997 concerning the scientific aspects and the presentation of information necessary to support applications for the placing on the market of novel foods and novel food ingredients and the preparation of initial assessment reports under *Regulation (EC) No 258/97* of the European Parliament and of the Council [*Off. J. Eur. Commun.* **L 253**, 0001–0036 (1997)].
47. Commission *Regulation (EC) No. 1813/97* of 19 September 1997 concerning the compulsory indication on the labelling of certain foodstuffs produced from genetically modified organisms of particulars other than those called for in Directive 79/112/EC [*Off. J. Eur. Commun.* **L 257**, 0007–0008 (1997)].
48. Council *Regulation (EC) No. 1139/98* of 26 May 1998 concerning the compulsory indication of the labelling of certain foodstuffs produced from genetically modified organisms of particulars other than those provided for in *Directive 79/112/EEC* [*Off. J. Eur. Commun.* **L 159**, 0004–0007 (1998)].
49. SI 1999 No. 747.
50. SI 1995 No. 1412, as amended by SI 1996 No. 1260, SI 1998 No. 104, SI 1998 No. 2072, and SI 1999 No. 1528.
51. Council *Directive 82/471/EEC* of 30 June 1982 concerning certain products used in animal nutrition [*Off. J. Eur. Commun.* **L 213**, 0008–0014 (1982)].
52. Council *Directive 83/228/EEC* of 18 April 1983 on the fixing of guidelines for the assessment of certain products used in animal nutrition [*Off. J. Eur. Commun.* **L 126**, 0023–0027 (1983)].
53. Council *Directive 87/153/EEC* of 16 February 1987 fixing guidelines for the assessment of additives in animal nutrition [*Off. J. Eur. Commun.* **L 064**, 0019–0028 (1987)], as revised by Commission *Directive 94/40/EC* of 22 July 1994 amending Council *Directive 87/153/EEC* fixing guidelines for the assessment of additives in animal nutrition [*Off. J. Eur. Commun.* **L 208**, 0015–0019 (1994)] and Commission *Directive 95/11/EC* of 4 May 1995 amending Council Directive 87/153/EEC fixing guidelines for the assessment of additives in animal nutrition [*Off. J. Eur. Commun.* **L 106**, 0023–0024 (1995)].
54. See also Council *Directive 93/113/EC* of 14 December 1993 concerning the use and marketing of enzymes, micro-organisms and their preparations in animal nutrition [*Off. J. Eur. Commun.* **L 334**, 0017–0023 (1993)].
55. SI 1994 No. 3142.
56. Council *Directive 81/851/EEC* of 28 September 1981 on the approximation of the laws of the Member States relating to veterinary medicinal products [*Off. J. Eur. Commun.* **L 317**, 0001–0015 (1981)].
57. Council *Regulation (EEC) No. 2309/93* of 22 July 1993 laying down Community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Agency for the Evaluation of Medicinal Products [*Off. J. Eur. Commun.* **L 214**, 0001–0021 (1993)].
58. SI 1986 No. 1510. This was made under the *Food and Environment Act 1985*. It has since been amended (see SI 1990 No. 2487, SI 1994 No. 3142, and SI 1997 No. 188). This regulatory scheme has been partially superseded by the *Plant Protection Regulations*, which provide that the *Control of Pesticides Regulations* do not apply to any "plant protection product" (as defined in SI 1995 No. 887).
59. Council *Directive 91/414/EEC* of 15 July 1991 concerning the placing of plant protection products on the market [*Off. J. Eur. Commun.* **L 230**, 0001–0032 (1991)].
60. SI 1995 No. 887, as amended by SI 1997 No. 7, SI 1997 No. 2499, and SI 1999 No. 1228. The latter of these revoked the amendments in SI 1996 No. 1940 and SI 1998 No. 2760).
61. SI 1995 No. 888.
62. SI 1997 No. 189.
63. *Directive 98/8/EC* of the European Parliament and of the Council of 16 February 1998 concerning the placing of biocidal products on the market [*Off. J. Eur. Commun.* **L 123**, 0001–0063 (1998)].
64. See Council *Directive 93/77/EEC* of 21 September 1993 relating to fruit juices and certain similar products [*Off. J. Eur. Commun.* **L 244**, 0023–0031 (1993)]; Commission *Directive 92/76/EEC* of 6 October 1992 recognizing protected zones exposed to particular plant health risks in the Community [*Off. J. Eur. Commun.* **L 305**, 0012–0015 (1998)]; Commission *Directive 95/44/EC* of 26 July 1995 establishing the conditions under which certain harmful organisms, plants, plant products and other objects listed in Annexes I to V to Council *Directive 77/93/EEC* may be introduced into or moved within the Community or certain protected zones thereof, for trial or scientific purposes and for work on varietal selections [*Off. J. Eur. Commun.* **L 184**, 0034–0046 (1995)]; Commission *Directive 98/22/EC* of 15 April 1998 laying down the minimum conditions for carrying out plant health checks in the Community, at inspection posts other than those at the place of destination, of plants, plant products or other objects coming from third countries [*Off. J. Eur. Commun.* **L 126**, 0026–0028 (1998)]; and Commission *Decision 98/109/EC* of 2 February 1998 authorising Member States temporarily to take emergency measures against the dissemination of Thrips palmi Karny as regards Thailand [*Off. J. Eur. Commun.* **L 027**, 0047–0048 (1998)].
65. SI 1993 No. 1320, made under the *Plant Health Act 1967* and *Agriculture (Miscellaneous Provisions) Act 1972*. The Order replaced the *Plant Health Order (Great Britain) 1987* (SI 1987 No. 1758). The Order has been amended by: SI 1993 No. 3213, SI 1995 No. 1358, SI 1995 No. 2929, SI 1996 No. 25, SI 1996 No. 1165, SI 1996 No. 3242, SI 1996 No. 1145, SI 1997 No. 2907, SI 1998 No. 349, SI 1998 No. 1121, and SI 1998 No. 2245.
66. SI 1998 No. 1027 (revoking SI 1978 No. 294).
67. SI 1982 No. 844, as amended by SI 1985 No. 1529, SI 1989 No. 1314, SI 1990 No. 1353, SI 1992 No. 1615, and SI 1998 No. 2726.

68. Council *Directive 98/95/EC* of 14 December 1998 amending, in respect of the consolidation of the internal market, genetically modified plant varieties and plant genetic resources, Directives 66/400/EEC, 66/401/EEC, 66/402/EEC, 66/403/EEC, 69/208/EEC, 70/457/EEC and 70/458/EEC on the marketing of beet seed, fodder plant seed, cereal seed, seed potatoes, seed of oil and fibre plants and vegetable seed and on the common catalogue of varieties of agricultural plant species [*Off. J. Eur. Commun.* **L 025**, 0001–0026 (1999)].
69. Such as *Directive 90/220* (see n. 49).
70. SI 1993 No. 3250, as amended by SI 1994 No. 3142 and SI 1994 No. 3144.
71. SI 1994 No. 2328, as amended by SI 1994 No. 3142 and SI 1994 No. 3144.
72. Convention on the Grant of European Patents (Munich, 5 October 1973; TS 20 (1983); Cmnd 7090).
73. *Directive 98/44/EC* of the European Parliament and of the Council of 6 July 1998 on the legal protection of biotechnological inventions [*Off. J. Eur. Commun.* **L 213**, 0013–0021 (1998)].
74. *Regulation (EC) No. 1610/96* of the European Parliament and of the Council of 23 July 1996 concerning the creation of a supplementary protection certificate for plant protection products [*Off. J. Eur. Commun.* **L 198**, 0030–0035 (1996)].
75. See, particularly, Report (Doc. 8459) of the Committee on Agriculture and Rural Development of 9 July 1999 on biotechnology and intellectual property.
76. See ETS Nos. 65, 87, 102, 125, 145 and Recommendations Nos. 1213 and 1289.
77. European Convention for the Protection of Vertebrates used for Experimental or Other Scientific Purposes (ETS No. 123) of 18 March 1986; Council *Directive 86/609/EEC* of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes [*Off. J. Eur. Commun.* **L 358**, 0001–0027 (1986)].
78. See, amongst other documents, the Resolution on cloning [*Off. J. Eur. Commun.* **C 115**, 0092 (1997)], which followed the reports of the successful cloning of a sheep ("Dolly").
79. See, for example, the Council of Europe documents (ETS Nos. 87 and 145). With regard to the EC, see Council *Decision 1999/167/EC* of 25 January 1999 adopting a specific programme for research, technological development and demonstration on quality of life and management of living resources (1998 to 2002) [*Off. J. Eur. Commun.* **L 064**, 0001–0019 (1999)]; Council *Directive 93/35/EEC* of 14 June 1993 amending for the sixth time *Directive 76/768/EEC* on the approximation of the laws of the Member States relating to cosmetic products [*Off. J. Eur. Commun.* **L 151**, 0032–0037 (1993)]; and *Study Concerning the Ethical Implications of Transgenic Animals (Legal Aspects)* SEC/8581/95: A report to the European Commission Group of Advisors on the Ethical Implications of Biotechnology, May 1995.
80. A. Maclean, *The Elimination of Morality*, Routledge, London, 1992.
81. M. Banner, *Christian Ethics and Contemporary Moral Problems*, Cambridge University Press, Cambridge, UK, 1999, pp. 204–224.
82. Proposal for a European Parliament and Council Directive amending *Directive 90/220/EEC* on the deliberate release into the environment of genetically modified organisms [COM (1998) 85 final of 23/02/1998].
83. Draft Directive III/5574/98.
84. http://www.hgc.gov.uk

See other INTERNATIONAL ASPECTS entries; INTERNATIONAL INTELLECTUAL PROPERTY ISSUES FOR BIOTECHNOLOGY.

INTERNATIONAL INTELLECTUAL PROPERTY ISSUES FOR BIOTECHNOLOGY

CYNTHIA M. HO
Loyola University of Chicago School of Law
Chicago, Illinois

OUTLINE

OVERVIEW OF INTERNATIONAL INTELLECTUAL PROPERTY

Introduction

International intellectual property issues are becoming increasingly important as biotechnology is used and sold worldwide. This is because intellectual property rights are inextricably linked to the right to exclude others from use. The right to exclude can provide a competitive advantage or a barrier to entry into a commercial market. Several types of intellectual property provide some right to exclude others from an invention in the area of biotechnology — trade secrets, patents, and plant variety rights. Trade secrets enable their holder to prevent others from wrongfully appropriating valuable information for a potentially infinite time period. But they do not protect against independent invention, and they terminate once the information becomes public. A patent provides a right to exclude others, including independent inventors, from using the patented invention without consent, and only for a limited time. Plant variety rights function similarly to patents with respect to the ability to exclude others, but they are only available for plant "varieties."

Types of Protection

Trade Secrets. A trade secret typically consists of any information that is not generally known to others in the same business; it provides a competitive advantage to its owner. This is the easiest type of intellectual property

to acquire, but it also provides the weakest protection. Unlike other types of intellectual property protection, formal procedural requirements are usually unnecessary to "obtain" a trade secret. Rather, the inventor of a trade secret merely needs to keep the information reasonably secret. The trade secret lasts as long as the information remains secret. However, once the information becomes publicly known, the trade secret ceases to exist. One way this can happen is if the information is independently developed and patented by another. Under this scenario, a trade secret would be exterminated because a patent on the same information would reveal the information to the public (since patents are public documents). Moreover, under this scenario, the patent owner could preclude the former trade secret owner from using the now patented invention because patent owners generally have rights to exclude all others from the patented invention. Although some countries soften this approach by providing those who used an invention prior to its patenting by another (prior users) with the right of continued use, there is no uniformity in such protection; for example, while some European countries allow a limited right, no analogous protection exists for biotechnology trade secrets in the United States.

Even without the complications of a superceding patent, trade secret protection offers minimal protection. To begin with trade secret protection does not allow exclusion of those who independently invent the identical "trade secret" therefore, two or more individuals or corporations could theoretically be using the same trade secret without infringing on each other's rights if they all did so independently. A trade secret does not confer any affirmative rights except as to those who misappropriate the information (e.g., an employee who leaves with confidential information). Even then, the "protection" provided is usually inadequate because monetary compensation for the trade secret misappropriation cannot restore information to trade secrecy status if it has been disclosed to the public.

Patents. Patents provide inventors a reward or incentive for publicly disclosing an invention, by providing the inventor with the right to exclude others from the patented invention for a limited term. The exclusivity provided by a patent is considered critical to stimulate ideas and lead to further advances. The requirements established for patentability are aimed at securing the goal of promoting innovation. Although the requirements vary somewhat between various countries, the typical requirements for obtaining a patent are that (*1*) the invention constitute patentable subject matter, namely constitute the type of subject matter that country wants to encourage innovation in, and (*2*) that the invention, as disclosed in a patent application, satisfies technical patentability requirements. The scope of patentable subject matter may include both products and processes in all areas of technology. Technical patentability requirements typically require that the invention be at least "new," "useful" (or have "industrial application"), "nonobvious" (or have an "inventive step"), and fully disclosed in a written document such that someone who was similarly technically competent could reproduce the invention. These requirements are intended to define inventive activity deserving of a patent. A national patent office typically examines patent applications to determine whether the patentability requirements are met. Additionally, in some countries, third parties are allowed to oppose the issuance of a patent or petition to revoke an existing patent for failing to meet the technical requirements.

Patents are generally considered the preferable type of intellectual property to protect biotechnology because they provide the most exclusive rights. Unlike trade secrets, patents protect against independent invention because the owner of a patent can exclude all others from using the patented invention. In addition, because a patent can entitle its owner to exclude others, including competitors, a patent or even a potential patent can justify the often high cost of research and development involved in biotechnology. The potential to exclude all others through patent protection is considered more valuable than attempting to maintain a trade secret indefinitely with no potential to affirmatively exclude others. Thus patents are the principal type of protection that is sought for biotechnology even though disclosure of the invention is required and patent protection is not a certainty.

Plant Variety Rights. A plant variety right also confers some exclusive rights. A plant variety right, which is also referred to as a "breeder's right" (because the right is provided to the breeder of a plant variety), functions analogously to patent rights — a relatively exclusive right is provided to breeders of new plant varieties to further the development of agriculture. As with patents, plant variety rights are not automatic; rather, they must be applied for and examined to determine whether they meet the requisite technical requirements. The requirements for plant variety rights are intended to function similarly to those for patent rights in that both are intended to provide protection to subject matter that is truly innovative. Some of the technical requirements for plant breeder rights parallel those for patent rights — a variety must be "new" (although the definition differs from that for patents, as will be discussed later), "distinct," "uniform," and "stable." These requirements are generally less onerous than those needed to meet patentability requirements. In addition, unlike patents, disclosure of the invention, or at least the method of making the invention, is not always required, which could be seen as an advantage. Although plant variety rights will be addressed in this article, the focus is on patent protection because patents provide coverage for more types of biotechnology and broadest protection.

Examining International Issues

Intellectual property rights are national rights provided by individual countries such that examining international protection requires examining the laws of individual countries. The need to examine what protection is available for biotechnology on an international scale is of obvious importance. However, national laws are currently evolving, particularly in the area of biotechnology. One way in which an international perspective of important issues can be obtained is by examining key international agreements, as well as illustrative national

laws. Although no international agreement creates an international right, or even identical national rights, some agreements mandate minimum levels of protection by member countries, such that examining these agreements establishes what minimum levels of protection are globally available for biotechnology. The most important international agreement for this discussion is the Trade-Related Agreement on Intellectual Property (TRIPS) (1), which mandates minimum levels of protection for patents concerning all technology, as well as other types of intellectual property. In addition the International Convention for the Protection of New Varieties of Plants (UPOV) requires that its members provide minimum levels of protection for plant "varieties" (2,3). Both TRIPS and UPOV provide a useful framework for analyzing international protection as they establish a foundation of protection for all member countries. However, TRIPS and UPOV only provide a framework of minimum protection, rather than binding law. To determine global protection of intellectual property, national and regional laws will be discussed as illustrative of current approaches, as well as anticipated approaches in the near future. The laws applicable to Europe and Japan will be addressed as representative of current law impacting international protection of biotechnology (relevant U.S. law is addressed in a separate article of this encyclopedia). In addition to representing areas that are generally considered important for biotechnology, the laws of Europe and Japan highlight the application of both older and newer laws to biotechnology. Europe will be addressed as a single entity for this discussion because many European countries have adopted similar laws pursuant to international agreements other than TRIPS. The applicable agreements are the European Patent Convention (EPC), which applies to any European nation that signs the agreement and the European Union's Directive on the Legal Protection of Biotechnological Inventions (EU Directive), which applies to any country that is or becomes a member of the European Union (EU) (4,5).

Table 1 highlights membership in relevant international conventions based upon information obtained from the official website for each convention. However, it should be noted that this table only presents some members of these conventions — there are presently over 130 members of TRIPS and over 40 members of the UPOV and Budapest Conventions. For complete and membership regarding all conventions summarized in Table 1 (TRIPS, UPOV, and the Budapest Convention), the Web sites of those specific conventions should be consulted (6–11). In addition, for current membership to any of these agreements, official Web sites should be examined; this is particularly true for both TRIPS and the EU, as the scope of membership is expected to increase. For example, China is applying for membership to the WTO and if accepted, would be bound to comply with TRIPS. In addition countries that may be included within the EU include Hungary, Poland, and the Czech Republic (12).

Table 1 includes all members to EPC, as well as the EU (6–7). This table also includes important areas outside of Europe for which biotechnology protection is often considered important such as Australia, Canada, Japan, and New Zealand. Some areas where biotechnology protection is beginning to develop are also included for comparison such as South Korea and South Africa. For all these countries, Table 1 shows which are bound to the requirements of TRIPS and which provision of UPOV, if any, they comply with (as there are currently two versions of UPOV which govern) (8,9). Also countries that subscribe to the Budapest Convention, which governs deposits of biological material in relation to obtaining patent protection, are noted (10–11).

Table 1. Membership in International Conventions Impacting Biotechnology

Country	EPC	EU	TRIPS	UPOV 1991	UPOV 1978	Budapest Convention
Australia			Y	Y		Y
Austria	Y	Y	Y		Y	Y
Belgium	Y	Y	Y		Y	Y
Canada			Y		Y	Y
Denmark	Y	Y	Y	Y		Y
Finland	Y	Y	Y		Y	Y
France	Y	Y	Y		Y	Y
Germany	Y	Y	Y	Y		Y
Greece		Y	Y	—	—	Y
Ireland	Y	Y	Y		Y	Y
Israel			Y	Y		Y
Italy	Y	Y	Y		Y	Y
Japan			Y	Y		Y
Korea			Y			Y
Liechtenstein	Y		Y	—	—	Y
Luxembourg	Y	Y	Y	—	—	
Monaco	Y			—	—	Y
Netherlands	Y	Y	Y	Y		Y
New Zealand			Y		Y	
Norway			Y		Y	Y
Portugal	Y	Y	Y		Y	Y
South Africa			Y		Y	Y
Spain	Y	Y	Y		Y	Y
Sweden	Y	Y	Y	Y		Y
Switzerland	Y		Y		Y	Y
United Kingdoms	Y	Y	Y	Y		Y

DEFINING THE PARAMETERS OF PROTECTION

Patents Pursuant to TRIPS

It is important to note that although TRIPS does not create a uniform worldwide patent law, it does establish a minimum floor below which no member may go without subjecting itself to potential retaliation by other countries in the form of trade sanctions. All countries that are members of the World Trade Organization (WTO) — an international organization designed to reduce trade barriers, including barriers to trade based on intellectual property — must comply with TRIPS. TRIPS binds a substantial number of countries — over 130 countries are presently members of WTO; these countries constitute 90 percent of world trade and include countries that utilize or expect to utilize biotechnology (2). Compliance with TRIPS requires complying with provisions of other international agreements that TRIPS incorporates, such as the Paris

Convention for the Protection of Industrial Property (13). TRIPS came into effect on January 1, 1995. However, the TRIPS requirements were not immediately binding; all members were provided time to enact laws to comply with TRIPS, with more time allotted for developing countries (many of which previously provided limited, or even no, patent rights whatsoever). The first date by which any WTO member had to comply with TRIPS was January 1996; at this time all developed countries were supposed to be in compliance with all provisions of TRIPS. In addition, as of January 2000, all developing countries (and those whose economies are in transition from a centrally planned economy into a free-market economy) should be in compliance with most portions of TRIPS. However, "least-developed" countries have until 2005 to provide patents for any subject matter that was previously deemed unpatentable. Moreover less-developed countries (LDCs) do not have to comply with any of the minimum patentability standards until 2005, with the possibility for a further extension of time.

Compliance with TRIPS requirements is aided by the potential of some type of retaliatory behavior by other WTO member nations. In particular, if a member state fails to comply with TRIPS, another member country may challenge the noncompliance. Members must initially try to resolve issues through a dispute resolution process pursuant to the Dispute Settlement Understanding (DSU) (1,14). If members fail to reach a mutually satisfactory resolution regarding any alleged noncompliance with TRIPS, a complaining member may request a WTO panel to examine the issue in an adjudicatory proceeding. The WTO panel must issue a decision within a relatively short timeframe, after which parties must comply or appeal to an appellate body of the WTO and subsequently comply with any decision by the appellate body. If a member fails to comply, sanctions may ultimately be imposed, including retaliatory action that includes withholding benefits under TRIPS or other WTO agreements such as the General Agreement on Tariffs and Trade (GATT). Although additional information on the DSU is beyond the scope of this chapter, the WTO Web site provides a current overview of active disputes; also additional information on dispute settlement is readily available from other sources (14–19). However, it should be noted here that the DSU rules have already been effectively applied to enforce protection of biotechnology. For example, India's compliance with the transitional provisions of TRIPS was challenged by the United States and the EU. After failing to convince both the WTO panel and Appellate Body that its laws were in compliance, India took action to amend its laws (20–21). The India dispute illustrates the effectiveness of TRIPS in effectuating real change to protection of biotechnology, as well as the relative rapidity under which such change occurs.

The TRIPS requirements relating to patents focus on three main issues: what must be patented, the scope of patent protection for issued patents, and the enforcement of patent rights. Although all three of these are important, this section will primarily focus on the first two issues since their impact is more particular to the field of biotechnology; information concerning enforcement of all patent rights under TRIPS, including patent rights for biotechnology patents, is available from other sources (19). The TRIPS requirements are actually minimum standards of protection that must be satisfied by all members. Because of the substantial number of countries that must comply with TRIPS, the TRIPS requirements are very important.

Eligible Subject Matter. The requirements for what must be patented are set forth under Article 27 of TRIPS. In particular, TRIPS first sets forth that:

> 1. Subject to the provisions of paragraphs 2 and 3, [p]atents shall be available for any inventions, whether products or processes, in all fields of technology, provided that they are new, involve an inventive step and are capable of industrial application. Subject to paragraph 4 of Article 65, paragraph 8 of Article 70 and paragraph 3 of this Article, patents shall be available and patent rights enjoyable without discrimination as to the place of invention, the field of technology and whether products are imported or locally produced.

This provision imposes several requirements. First, it states that patents must be available for "inventions" subject to two limitations (in paragraphs 2 and 3) which exist for particular categories of subject matter as well as exceptional situations. Second, it states that a patentable invention must be "new," involve an "inventive step," and be "capable of industrial application." Third, it states that except for countries to whom a transitional period is allowed (under Articles 65 and 70 of TRIPS), patents and patent rights must be provided without discrimination; countries cannot discriminate in issuing patents or providing patent rights based on where the invention was made, the type of invention, and whether products are imported or locally produced. As the last requirement relates to both patentability and scope of patent rights, it will be discussed with respect to both issues separately.

Although these requirements are not presently binding on all WTO members, they must eventually be complied with. The TRIPS requirements of patentability indicate what types of biotechnology are presently patentable or will be in the future. TRIPS requires that patents be granted to all "inventions" that satisfy technical patentability requirements, subject to three specific exceptions. Table 2 summarizes these exceptions. As shown in this table, subject matter may be excluded if it is (*1*) not considered an "invention, or (*2*) specifically delineated within TRIPS as excludable, or (*3*) falls within the unspecified exclusion for inventions contrary to *ordre public* or "morality." Each one of these three requirements allows member states the opportunity to narrow the scope of patentability. The scope of patentable subject matter may be limited based on a narrow interpretation of undefined TRIPS terms; this table indicates such undefined terms by placing them all in quotes.

The requirement that patents be available for all "inventions" without regard to the type of technology involved and whether the invention is a product or a process is important for biotechnology. Prior to TRIPS more than 50 countries did not provide patent protection

Table 2. Permissible Exclusions of Patentable "Inventions" Pursuant to TRIPS[a]

Anything excluded from the definition of an "invention"[b]
Specific subject matter exclusions (Article 27(3))
• Methods of treatment of humans and animals
• Plants[c] and animals (other than microorganisms)
• "Essentially biological processes" for making plants and animals (other than microorganisms)
Unspecified exclusion for "*ordre public*" or "morality" (Article 27(2)) (subject matter may be excluded if preventing commercial exploitation is necessary to protect "*ordre public*" or "morality")[d]
• Protecting human, animal or plant life
• Protecting health
• Avoiding serious prejudice to the environment

[a]TRIPS *permits* members to exclude these categories from patentability, and members are free to provide patent protection for these categories.
[b]The word "invention," as well as other quoted words on this table, are undefined in TRIPS and may be subject to differing interpretations (which can affect the scope of patentable biotechnology).
[c]TRIPS does require that plant "varieties" (also undefined within TRIPS) be protected either under the patent system or a sui generis system.
[d]N/A where commercial exploitation is prohibited by local law.

for pharmaceutical and other products relating to health and medicine (22,23). In fact, the disparate treatment of such products was a prime consideration in enacting this language (24). Still it is unclear whether this provision mandates a general principle of patentability of "inventions" (25). In addition, even if a general presumption of patentability is established, the meaning of the word "invention," as well as the allowable exceptions from patentable "inventions," can substantially narrow what types of biotechnology are patentable (26).

Identifying what is an appropriate "invention" is the first point at which subject matter may be narrowed. TRIPS itself does not define an "invention." Although this is consistent with prior patent law in many countries, it leaves the scope of patentable subject matter particularly uncertain in the biotechnology area where there is not always a marked distinction between unpatentable "discoveries" and patentable "inventions." Countries generally consider discoveries and naturally occurring substances to be unpatentable inventions, but the genetic manipulation of naturally occurring substances has arguably blurred the line between discovery and invention. Some countries take a relatively expansive view of the term "invention" to include products that are found in nature, so long as they are isolated by an unnatural (biotechnological) process; isolated genes and gene sequences may be considered patentable under such a definition of invention. On the other hand, countries may elect to *exclude* isolated products of nature as noninventions (22,24). Because the term is undefined in TRIPS, countries are free to adopt either view. However, even if isolated biological material is considered a patentable invention, problems with other patentability requirements such as "industrial application" may exist (e.g., if there is no known use for the material other than as a general probe, as will be later discussed).

Specific Exclusions. The three types of subject matter that are specifically set forth as excludable subject matter under TRIPS all impact the scope of biotechnology that may be patented. TRIPS allows the following categories of subject matter to be excluded from patentability:

1. Methods of treating and diagnosing humans and animals
2. Plants and animals other than microorganisms
3. "Essentially biological processes" for creating plants and animals

It is important to note that these are *optional* exclusions from patentability. Accordingly, countries may adopt some, all, or none of these exceptions in their patent laws and still be in compliance with TRIPS. In addition, it should be noted that even if a country enables a broad range of subject matter to be considered an "invention," adoption of these exclusions, or a broad interpretation of these exclusions could significantly narrow the field of what is considered patentable subject matter. Moreover, even if none of these exclusions are adopted, what is actually patentable may nonetheless be a narrow range of inventions if the criteria of patentability are applied strictly, as will be explained later.

1. *Methods of Treatment and Diagnosis.* TRIPS allows methods of treating or diagnosing humans or animals to be excluded from patentability under Article 27(3)(a). This prohibition is similar to the language of many European and Latin American countries' patent laws (22). Application of similar prohibitions under national laws have shown that the scope of such an exclusion may be a function of the interpretation of individual terms, such as the meaning of "therapy," "method of treatment," or "diagnosis." (27–32). This exclusion is of great importance for biotechnology as the exclusion of "therapeutic and surgical methods" may exclude patents for gene therapy. In addition the bar on "diagnostic methods" may result in the exclusion of genetic testing or screening kits from patentability. Countries have taken varying stances in applying similar statutory exclusions to gene therapy; for example, some countries only exclude in vivo, but not ex vivo methods of treatment, whereas other countries exclude treatment involving humans but permit patenting of the identical treatment on animals.
2. *Animals, Plants, and Essentially Biological Processes.* TRIPS also states under Article 27(3)(b) that members may exclude from patentability "plants and animals other than micro-organisms, and essentially biological processes for the production of plants or animals other than non-biological and microbiological processes." However, protection of "plant varieties" must be provided either under the patent system or a sui generis system (1). This provision allows countries to disallow patents on transgenic plants and animals, even if they meet the technical requirements of patentability. Although genetically

modified plants may be entitled to some protection if considered a plant "variety" (which is not defined under TRIPS), exactly what protection is uncertain. While TRIPS requires an "effective" sui generis system, there is no definition of what this system requires. This may include one system for protecting plant varieties that existed at the time TRIPS was enacted — UPOV (24); however, not all commentators agree that UPOV is consistent with this provision (33). Moreover, even if applicable, UPOV does not provide the identical scope of protection as that required under TRIPS, as will be noted in the later discussion of UPOV. In addition it is uncertain what other system besides UPOV would be adequate. Moreover, TRIPS does not require any protection for transgenic animals. While there is the possibility that the process of creating a transgenic animal could be patented (assuming the technical patentability requirements are met), this provides less protection than if the animal itself were patentable. In addition patentability of a process for creating a transgenic animal is not certain as members need not protect "essentially biological processes" for the creation of animals. TRIPS does not define "essentially biological" other than to state that it does not include processes that are "nonbiological" or "microbiological." Accordingly member countries could define "essentially biological" to exclude any process that involved at least one biological step (i.e., one that does not involve genetic engineering). This exclusion will likely generate continued debate as it was controversial during the initial TRIPS negotiation and must be reviewed every four years after TRIPS enters into force (1,26,33).

Nonspecific Exclusion. In addition to the foregoing specific exclusions that implicate biotechnology, a potentially major hurdle for biotechnological inventions may be the unspecified exclusion under Article 27(2) that allows "inventions" to be barred from patentability if preventing "commercial exploitation" of the subject matter is necessary to protect "*ordre public* or morality." However, TRIPS also states that this exclusion is not satisfied by the mere fact that exploitation of an invention will be in violation of a national law. TRIPS does not provide additional guidance on the definition of *ordre public* or *morality* other than to indicate that it may involve the following broad categories: protecting human, animal or plant life, protecting health, and avoiding serious prejudice to the environment (1). Further, because this exclusion refers to "inventions" rather than categories of subject matter as in Article 27(3), it appears that it may need to be applied on a case-by-case basis (24).

This provision stands as a potentially large barrier to patenting new types of biotechnology as patenting biotechnology raises issues of morality even in countries where patenting of biotechnology is allowed. Similar provisions have been utilized, or noted as capable of being utilized, to deny patents on transgenic plants and animals as well as isolated gene sequences (27). However, interpretation of what constitutes ordre public or morality in the context of patenting biotechnology has been elusive. This exception requires that the violation of "ordre public" or "morality" be the result of "commercial exploitation." This may pose some interpretative difficulty as commercial use or exploitation is not required to be disclosed in a patent application and may not be known. Accordingly, although a patent office is charged with determining whether commercial exploitation of an invention would violate these undefined terms, the necessary information may not be available at the time of examination to enable an educated decision.

The scope of this exception may be limited by prior interpretations of similar terms, albeit in contexts other than TRIPS, or even patent law. In particular, prior WTO dispute panels have examined and interpreted the meaning of "necessary" and "morality" under the General Agreement of Trade and Tariffs (GATT), which governs trade of goods (24,34–36). In that context, it has been found that the term "necessary" requires objectively justifiable measures and that there be no alternative measure available (24). Although prior WTO panels may be relied on by subsequent panels, it is unclear what weight a panel would give to a previously interpretation made in a different context (37). In addition it has been suggested that the definition of *ordre public* can be derived from case law by the European Court of Justice, the court that arbitrates issues among the EU countries (38). This may however be similarly unhelpful as there are no such cases interpreting *ordre public* in the context of patentability. Finally, countries or the WTO could look to prior interpretation of similar provisions, such as in Europe; although there are no decisions finding that inventions violate such a provision, suggestions as to the parameters of such a provision may be considered. For example, European courts have interpreted a similar provision under EPC article 53(a). The EU Biotechnology Directive also provides explicit examples of types of commercial exploitation that would violate an identically worded provision (5,27).

Illustrations of Patentable Subject Matter. Because TRIPS is a minimum standards framework with terms that are subject to differing interpretations, the scope of patentable subject matter may vary among different countries that comply with TRIPS. The diversity of approaches may be illustrated by a comparative view of specific types of patentable subject matter. A recently compiled report by the Working Party of the Trade Committee of the Organisation for Economic Cooperation and Development (OECD) provides a good source of such information. In response to an OECD questionnaire to individual countries concerning intellectual property practices in the area of biotechnology, information from patent offices of 22 OECD countries was obtained. Although the information was provided in varying degrees of detail and does not necessarily constitute official policy of individual patent offices, it is nonetheless useful for comparison.

Table 3 presents a summary of information submitted by individual patent offices to the OECD concerning the types of inventions in the area of biotechnology that are considered patentable subject matter. This information was submitted roughly in the time period 1998 to 1999. Accordingly, to the extent that laws have been amended

Table 3. Patentable Subject Matter

	Nucleic acid sequences	Amino acid sequences	Isolated materials	Living unicellular organisms	Plants, parts, varieties	Animals, organ varieties	Humans, organ, human-derived products	Methods of treatment	Methods of diagnosis
Australia	Y	Y	Y	Y	Y	Y	N	Y, except essentially biological processes that produce humans	Y except essentially biological processes that produce humans
Canada	Y	Y	Y	Y	N (but PVR exist)	N	Human-derived only	No, except ex vivo	Y, unless surgery or therapy
Japan	Y	Y	Y	Y	Y	Y	Human-derived; human organs if not immoral	Only animals or ex vivo	Animals only
Korea	Y	Y	Y	Y	N, except asexual plants	Y	Human-derived only	Animals only	Animals only
New Zealand	Y	Y	Y	Y	Y	Y	Human organs; human-derived (if process to create not immoral)	Animals only	Y, unless surgery involved
Norway	Y	Y	Y	Y	N, except plant parts that can't differentiate to whole plants	Animal organs only	N, except isolated elements	Ex vivo only	Ex vivo only
EPC/EPO	Y	Y	Y	Y	Y, except "varieties"	Y, except "varieties"	Human-derived only	N	N
EU	Y	Y	Y	Y	Y, if not technically limited to specific variety	Y, unless animal modified to cause suffering without substantial medical benefit	Human-derived if isolated or produced by "technical process"	N/A (national/EPC law to apply)	N/A (national/EPC law to apply)
United States	Y	Y	Y	Y	Y	Y	Human-derived only	Y	Y

since that time, this information would not reflect that. The table includes countries in which patents are often sought such as Japan, Canada, and Australia. In addition, the EPO is included because it issues EP patents that can become transformed into national patents. Information from individual countries of EPC are not included in this table as their laws largely mirror that of the European Patent Office (EPO), the governing office of EPC. However, some of these countries did provide separate information to the OECD survey, and this information is available from the OECD Web site (40). In addition some countries are included for comparison such as South Korea, New Zealand, Norway, and the United States. For information concerning all 22 countries that responded to the OECD questionnaire, the actual report should be consulted, as well as the responses of individual countries (39–41).

As noted on Table 3, there are several categories of biotechnological matter that are universally considered patentable subject matter. In particular, chemical structures composed of nucleic acid sequences and gene sequences are considered patentable subject matter in all responding countries; this is true regardless of whether the sequences correspond in part or whole to information found in living organisms. Material isolated from living organisms, other than such sequences, are all considered patentable subject matter as well as living unicellular organisms. It should be noted, however, that despite the uniform patent eligibility of such inventions, actual patents nonetheless may not issue because of how the technical requirements of patentability are applied.

More variation in patentable subject matter exists with respect to multicellular organisms and subparts of those organisms. The only common denominator among approaches is that patents on humans are universally rejected. However, "human-derived products," including cell lines, genes, and nucleic or amino acid sequences are considered patentable, for example, in Canada, Japan, and the United States. However, even to the extent that multicellular organisms and their parts are not categorically excluded from protection, they may nonetheless be denied protection on grounds similar to TRIPS 27(2)—namely they may violate a provision of the patent act that precludes patents on inventions that are unethical or immoral. Exclusions on the basis of morality could be a factor in the denial of patents for inventions by the EPO as well as by the patent offices of Japan, and Canada as all of them have such a provision.

Similarly there is wide variation with respect to what types of methods of treatment or diagnosis would be considered patentable. For the purposes of comparison, the methods of treatment column in Table 3 represents responses to both whether methods of treatment are patentable and whether methods involving genetic engineering for purposes other than surgery, therapy or diagnosis (e.g., for experimentation or research) are patentable. Some countries appear to take a fairly strict interpretation of statutory provisions against methods of treatment and diagnosis such that gene therapy is excluded. Others will allow methods of treatment or diagnosis only with respect to animals, or with respect to methods occurring outside the body (ex vivo). Finally, just as with patents on multicellular organisms, patentability of biotechnological methods are also impacted by provisions in patent acts that preclude patents that are immoral or contrary to *ordre public*.

Table 4. Technical Patentability Requirements

A patent application must disclose an "invention" (see Table 2) that is: • "new" • "industrial application" or "useful" • "inventive step" or "nonobvious"
Patent application (Article 29) must disclose invention adequately for duplication by one "skilled in the art"

Patentability Criteria. There are two basic requirements for patentability, as shown in Table 4. An application must establish that an invention meets technical requirements of patentability. The application must further sufficiently disclose the invention such that someone who is similarly skilled could replicate the invention. Article 27(1) of TRIPS requires that technical requirements include that an invention be new (novel), nonobvious (or have an inventive step), and be useful (or be industrially applicable) (1). Although these requirements are ones that are common to most patent laws, countries have differed in their definition of these terms, as well as in the application of these terms to biotechnology (22,24). For example, there is no consensus on whether patents on gene sequences (that are considered "inventions") may be denied for lack of "industrial applicability" if the function of gene sequences is unknown other than as a probe. Similarly there is no consensus on whether isolated biological matter is "new." While application of patentability requirements to biotechnology are developing with respect to all the requirements, novelty and adequate description will be focused on here because application of these standards to biotechnology raises unique issues. As was seen with patentable subject matter, the undefined patentability criteria allows wide variation among countries; each requirement is an additional juncture at which the scope of patentable subject matter may be further limited.

Novelty. The requirement that an invention be new is fundamental to the patent system policy of providing an incentive to produce things that would not otherwise be known to the public. An invention is generally new or novel if it was not known or available to others prior to the application of the patent; it is often stated that an invention is new if it is not previously known in the "prior art." Prior art may include descriptions in printed publications (e.g., foreign patents and published patent applications) as well as oral information and actual use. In all cases, to constitute prior art, the information must be such that the invention is essentially known to those in the same field (i.e., the description must be adequate to "teach" the invention to one in the field).

In determining novelty, it is important to note that pursuant to TRIPS, all WTO members must now recognize a "right of priority" for filing of patent applications (1,13,42). This means that if a patent application is filed in a member country, that date of filing may be relied on for subsequent filings of applications for the same invention in

other member countries, so long as subsequent filings are within 12 months of the original application (1,13,42). This right of priority is very important to determining novelty because, without such a right, an applicant's own application could bar a patent in another country unless patent applications were simultaneously filed in every country in which protection were desired. Therefore recognition of priority date is important as it gives an applicant 12 months within which to decide where to file and actually complete the formal filing requirements.

There are different interpretations of the novelty requirement taken by countries to determine whether an invention is sufficiently deserving on patent protection. The different approaches reflect different perspectives of how new something must be before a patent is awarded. Under a "strict" approach to novelty, an invention is not novel if it was known in any way, in any country, prior to the date of the patent application — regardless of whether the applicant was aware of it, or if the applicant could have reasonably been aware of it; this strict approach is often referred to as "absolute novelty." However, another approach is to reward an applicant for bringing some invention to the particular territory in which the patent is sought that likely would not otherwise have been known. Under this "relative novelty" approach, nonwritten knowledge or use of an invention outside the territory is not considered prior art for purposes of novelty; this is the approach taken by the United States. The rationale for this scheme is that a printed publication is accessible to all even if printed in another country, whereas a use of an invention is much more difficult to ascertain and thus a patent should be granted to one who brings the invention to the public's attention.

In the biotechnology area, the distinction between the absolute and relative novelty standards are demonstrated by some biotechnology methods. For example, new uses of previously known products (often referred to as "second use" or "second medical use") may be disallowed under a strict novelty criterion. Although such uses are considered "new" under both United States law as well as the EPC, TRIPS certainly does not mandate such an interpretation. Similarly new products that result from previously known processes may be disallowed under a strict novelty criterion. Again, the United States provides patent protection for such processes.

It is possible to also provide a grace period in terms of novelty. In particular, a grace period may allow an applicant to disclose an invention to the public prior to filing the patent application without sacrificing patent rights. It is common for the grace period to last only six months from the triggering event and to be no longer than a year; for example, the United States has a one-year grace period, while Canada and EPC provide a more limited period of six months. The rationale behind such a system is that it enables speedier public access without removing the incentive of patent protection. Countries vary with respect to what activities are allowable during the grace period. For example, in the United States a grace period is provided for not only publications but also prior use, sale, and offers for sale. However, in other countries, a grace period may only be available for activity of an inventor, and the activity may be limited to disseminating information only at specified conferences; some illustrative subscribers include Canada, Australia, and EPC.

Adequate Description. TRIPS requires that for a patent to issue, an application must first disclose the invention "in a manner sufficiently clear and complete for the invention to be carried out by a person skilled in the art." (1). Once again, the TRIPS agreement has only mandated the result but has not dictated how countries must meet it. Countries vary with respect to whether a deposit of a sample of a biotechnological invention is required, as well as when and where this must be done (4,5,39). The Budapest Convention provides some guidance to its members (many of whom include WTO members) on when deposits are required, and the relevant procedures that must be followed (10).

Patent Rights

Patent Term. TRIPS requires that members provide a minimum patent term of 20 years, calculated from the date of filing of the patent application (1). An applicant generally has no rights during the application process. No extension of the term is required by TRIPS even if the effective term has been substantially reduced by the patent examination process; this is problematic for patents that involve biotechnology as they often take longer to examine than more traditional inventions. In addition no extension of term is provided for delays in the sale of a patented invention that often result from regulatory approval for pharmaceuticals. Accordingly the effective term of pharmaceutical patents may be considerably shorter than that provided to other subject matter. Although many developed countries, including the United States, and more recently Japan, provide an extended term of protection for such patents, TRIPS does not require it. Accordingly, although TRIPS mandates patent protection for all inventions and nondiscrimination in terms of patent rights, it does not mandate an effectively equal patent term.

Scope of Protection. Consistent with prior patent law doctrine, patents granted pursuant to TRIPS do not provide an affirmative right of use to the patent owner. Rather, patents only provide a right to exclude others. To be entitled to use the patented invention, the owner must determine if there are additional laws with which compliance is necessary. For example, a newly patented pharmaceutical typically cannot be sold without governmental approval and separate applications are required to obtain patent rights and the right to sell. Moreover the patent owner may need to determine if permission from another patent owner is required to avoid infringement. For example, the owner of a patent on a new use of a previously known and patented compound would need permission from the owner of the patented compound to actually use the newly patented invention (as the owner of the patented compound has the right to exclude others from making the compound).

TRIPS provides that a patent confers on its owner "exclusive rights" to prevent unauthorized persons from certain activities. The activities that may be excluded differ with regard to whether the subject matter of the

patent is a product or process. If the patented subject matter is a product, the owner can exclude others from "making, selling, offering for sale, selling, or importing, the patented product. On the other hand, if the patented subject matter is a process, the owner is entitled to exclude others from using the patented process as well as offering for sale or selling the process; in addition the owner can exclude others from importing the product obtained "directly" from the patented process. Most of the exclusive rights established under TRIPS have been generally recognized in industrialized countries. It should be noted that many industrial countries also allow rights against persons who contribute or induce these activities, although TRIPS does not require this; the United States, Japan, and certain European countries allow such rights.

Although TRIPS specifies that certain activity is within the patent owner's "exclusive rights," the nature of that exclusivity may be limited pursuant to other articles of TRIPS (as will be discussed in the next section), as well as by a nation's interpretation of these provisions. For example, what constitutes a patented product or process must first be determined in order to determine what may be excluded. Generally, patents include one or more claims, which are sentences at the end of a patent that define the scope of the invention, and accordingly define the scope of patent rights. In addition patents generally can include claims to both products and processes; some countries allow claims to not only chemical structures but also claims to function and products made by certain processes. To determine what constitutes the patented invention, the claims most likely will be examined (often referred to as "claim interpretation"). However, countries differ vastly in their approaches to claim interpretation. Countries may look not only at the patent claims but also the rest of the patent, its history, and other extrinsic evidence to determine the "true meaning" of the claims. Some countries are more liberal in interpreting patent claims, and they find patent infringement even when a defendant is not precisely within the literal scope of patent claims (sometimes referred to as infringement under the "doctrine of equivalents" as the defendant's actions are considered equivalent as a matter of equity, to what is literally stated in the claims). Equivalents are considered either at the time the application is filed or at the time of the infringement. All of the above variations allow member countries substantial leeway in the extent of protection granted.

Regardless of how a country interprets the patent, TRIPS provides a higher threshold of protection for patented processes than was previously recognized in most countries. In particular, TRIPS provides the right to exclude unauthorized persons from products obtained "directly" from a patented process in addition to excluding persons from the patented process itself. The additional protection against products "directly" obtained by a patented process is intended to provide protection in the case where the patented process is used in a country where no protection exists, and the resulting product is imported into the country in which patent protection exists (24). The extension of protection to direct products also aids owners of such patented processes to establish unauthorized use of a patented process as access to an infringing process is often impossible. In addition TRIPS now provides that judicial authorities may require the alleged infringer to rebut a presumption of infringement in the case of patented process if the infringer's product is identical to that produced by the patented process (1). However, it is only required in one of two circumstances: (*1*) where the product obtained by the patented process is "new," or (*2*) when there is a "substantial likelihood that the identical product was made by the process and the owner of the patent has been unable, through reasonable efforts, to determine the process actually used" (1).

Although the scope of the patent owner's exclusive rights is fairly clear, one exception is the scope of the importation right. In particular, TRIPS specifies that for patents on both products and processes, the owner has the right to exclude others from importing the patented product or the product obtained directly from the patented process. The right to preclude importation of a patented product is a right that was not generally recognized in industrialized countries prior to TRIPS. However, TRIPS does not clarify what this new right should include; there is a footnote after the importation right that cross-references an earlier article of TRIPS that states that for purposes of dispute settlement, TRIPS is considered not to address the issue of exhaustion of patent rights. The principle of "exhausting" a patent right is that if a patented product is legitimately sold or otherwise conveyed, no further patent rights exist with regard to that article. To the extent that there is an argument that imports of patented products fail to impinge on a patent owner's exclusive rights because of the principle of international exhaustion, such arguments will not be recognized in official dispute settlement proceedings.

Exceptions. Although the previously discussed provision of TRIPS provides "exclusive rights" to the patent owner, TRIPS also clearly contemplates that the "exclusive right" may be more circumscribed. Two separate articles of TRIPS provide exceptions to the "exclusive right" that is provided in Article 28. In particular, Article 30 provides a "limited exception" to the exclusive right, while Article 31 contemplates that member countries may allow unauthorized "other use" of a patent. As these exceptions are critical to defining the full scope of patent rights — just as possible exclusions to patentable subject matter were important to determining what inventions are patentable — the exceptions to the patent owner's exclusive rights will be discussed here.

"Limited" Exceptions. Although TRIPS provides that members may provide "limited exceptions" to the patent right, it does not clearly define what constitutes a "limited exception." In particular, TRIPS Article 30 states that "[m]embers may provide limited exceptions to the exclusive rights conferred by a patent, provided that such exceptions do not unreasonably conflict with a normal exploitation of the patent and do not unreasonably prejudice the legitimate interests of the patent owner, taking account of the legitimate interests of third parties" (1).

Although the scope of this exception is not explicity clear, some countries have tried to incorporate the exception into national law. Some countries have incorporated

the language of TRIPS Article 30 wholesale into their national laws, without attempting to define any terms. Other countries, while not using the literal language in Article 30, have adopted language that is equally vague. For example, Argentina has amended its patent laws to state that the patent office is permitted to establish limited exceptions in "sectors of vital interest to the socio-economic and technological development of the country" (43). It remains to be seen whether countries with established patent systems, including exceptions to patent rights, are in fact TRIPS-compliant; this cannot be clearly established in the absence of challenge by another country. However, even in the absence of an official challenge, countries are becoming more cognizant of TRIPS requirements and the possibility of a challenge. In fact there has been some discussion in the United States concerning whether a provision of the U.S. patent laws that precludes patent owners from obtaining relief against infringing doctors of medical procedure patents either violates the TRIPS rights or is exempt under this provision (44–46).

The proper interpretation of the "limited exception" provision will probably be a major issue in international patent protection. One panel recently addressed WTO whether Article 30 permits (*1*) the manufacture of patented product for the purpose of obtaining regulatory approval prior to the patent expiration or (*2*) the manufacture and storage ("stockpiling") of generic drugs such that they can be sold the day the patent expires (15). In particular, Canada's patent laws were alleged to infringe on patent rights during the patent term (15,47). Canada argued that Article 30 provides it with an exception to what would otherwise technically be a violation of patent rights mandated by TRIPS (48). In particular, Canada believed that its actions were allowable as a "limited exception" in the case of patented pharmaceuticals where generic drug companies need to infringe a patented drug to obtain regulatory approval prior to the expiry of the patent (48). The rationale for such an exception was that without it, a patent owner would effectively be given an extension of patent term due to the generic manufacturer's need to wait for regulatory approval; allowing limited infringement by the generic manufacturer was considered to be in the public interest because it enables cheaper drugs to be provided to the public sooner. The WTO found that the regulatory approval exception was an allowable exception to the usual patent owner rights pursuant to article 30. However, the WTO panel found that the stockpiling provision was not a "limited exception" in accordance with article 30. Although the decision is not binding on other parties, it is important to the many nations which have regulatory approval exceptions similar to Canada's such as the United States and Japan; these countries are now more assured that their patent law exceptions are TRIPS-compliant. The WTO decision is also important beyond the specific subject matter of regulatory approvals. In particular, the WTO panel decision explained how article 30 should be interpreted in relation to patent rights under article 28 (47). The WTO panel underscored that a patent owner is entitled to the full panoply of patent rights listed under article 28 (and not just the right to sell as Canada had previously asserted) (47). In addition, the WTO panel clarified that to be entitled to an exception under article 30, three requirements must be met:

1. there must be a "limited exception" to the exclusive rights;
2. the exception must not "unreasonably conflict" with the "normal exploitation" of a patent; and
3. the exception must not "unreasonably prejudice" the "legitimate interests" of the patent owner, taking into account the "legitimate interests" of "third parties."

The WTO panel decision itself may be consulted for the detailed explanation of each of these requirements (47). Even without a complete explanation of the panel's opinion, however, it should now be appearent that future disputes concerning exceptions to patent rights will be carefully evaluated by WTO panels in accordance with the specific facts of each case.

In addition, what other exceptions to a patent owner's exclusive rights may exist remains an issue. Some countries and commentators have assumed that certain activity is covered by Article 30 based on prior drafts of Article 30 (22,24). For example, activity that has been assumed to be allowable includes private noncommercial use, use of the invention for research, experimental use to test or improve the invention, use for teaching purposes, preparation of medicine, prior use of the invention by a third party before the date of application of the patent (22,26). However, whether such uses are actually consistent with article 30 is unknown. Although a WTO panel recently had occasion to address this issue, it specifically declined to comment on whether national patent laws providing exceptions for experimental use were consistent with TRIPS ariticle 30 (47).

Compulsory Licensing. Compulsory licenses are an important issue as they substantially limit a patent right. The term "compulsory licensing" refers to a license of the patent owner's invention that is involuntary, or compulsory. Compulsory licensing is important to biotechnology patents because even if patents are appropriately granted, the patent right is substantially diluted if compulsory licensing is allowed. Moreover, prior to TRIPS, it was not uncommon, particularly in developing countries, to require compulsory licensing of pharmaceutical patents, or other patents relating to health on the ground that it was necessary in the interest of public welfare. In addition compulsory licenses have been granted in some countries where the patent owner is not using the invention in the country (usually referred to as not "working" the invention).

Article 27 prohibits certain types of "discrimination"; in particular, it requires that all patents, once issued, be entitled to the same right to exclude without regard to the subject matter of the patented technology; in particular, it states that there should be no discrimination based on whether the invention is "imported or locally produced." The negotiating history of this provision indicates that developed countries intended this to exclude compulsory licenses for nonworking as previously allowed

under the Paris Convention (24,49). Some interpretations of this provision interpret it as eliminating such a requirement (50–51). However, although this may have been the intent of some countries in drafting this provision, the correct interpretation is unclear. For example, Brazil explicitly provides compulsory licenses for failure to exploit an invention locally (52–54).

An important issue under TRIPS is determining what types of compulsory licensing are permissible. Several articles of TRIPS may relate to compulsory licensing. Article 31, which places procedural limitations on licenses, is often presumed to apply to compulsory licenses; indeed, although Article 31 presently refers to "other use" that is unauthorized by the patent owner, prior drafts of Article 31 used the term "compulsory licenses"(24). Two additional articles of TRIPS recognize that intellectual property rights, including patent rights, are not absolute. In particular, Article 7 states that:

> protection and enforcement of intellectual property rights should contribute to the promotion of technological innovation ... to the mutual advantage of producers and users of technological knowledge and in a manner conductive to social and economic welfare, and to the balance of rights and obligations.

Article 8 provides that:

> 1. Members may ... adopt measures necessary to protect public health and nutrition, and to promote the public interest in sectors of vital importance to their socio-economic and technological development, provided that such measures are consistent with the provisions of this agreement.
> 2. Appropriate measures ... may be needed to prevent the abuse of intellectual property rights by right holders or the resort to practices which unreasonably restrain trade or adversely affect the international transfer of technology.

It is unclear whether only situations described in Article 31 are subject to compulsory licensing, or if Article 31 imposes requirements on compulsory licenses of any subject matter. TRIPS implicitly provides certain ground for issuing compulsory licenses; in particular, compulsory licenses are deemed proper for a national emergency, noncommercial use, anticompetitive practice, if a patent owner has refused to license, and where necessary to practice another patented invention (1). However, TRIPS does not indicate what other grounds would be permissible for compulsory licensing, or whether this is the only provision that relates to such licensing. Although all compulsory licenses are an imposition on a patent right, at least those provided under Article 31 provide the owner with reasonable remuneration and are not to be granted for entire classes of inventions. However, if another provision of TRIPS is interpreted as justifying compulsory licensing — such as Article 8 — the protection provided to patent owners under Article 31 would essentially fail to exist. Interpreting Article 8 to establish an independent basis for compulsory licenses has been suggested by commentators, usually in the context of trying to accommodate developing countries who are attempting to comply with TRIPS (55–56).

Determining whether Article 31 is the only applicable provision to compulsory licensing is important because of the many limitations placed on licensing under Article 31. In particular, Article 31 requires that authorization be "considered on its individual merits." This seems to suggest that compulsory licensing of all patents within a certain category would be impermissible. In addition Article 31 requires that compulsory licensing should only be permitted if a patent owner has first denied a request; although the attempt to obtain permission is waived for a national emergency or public noncommercial use, even in these situations, the patent owner must be informed promptly after the use has begun. Also Article 31 places limitations on the scope and duration of the use; it must be nonexclusive, nonassignable, preferably for a domestic market, and subject to termination once the conditions that necessitated the unauthorized use cease. The decision to grant a license must be subject to judicial review, and the holder of the patent right must be entitled to "adequate remuneration," which takes into account the "economic value of the authorization," The holder of such a right is also entitled to judicial review of any decision regarding remuneration.

All of the above requirements raise issues of interpretation. For example, it is unclear whether compulsory licenses must be granted on an individual basis, or if categories of inventions may be considered together. TRIPS also does not specify any criteria to determine whether remuneration is "adequate" other than to state that it must take into account the "economic value of the authorization." Countries could continue to apply the average royalty rate paid in voluntary licenses within a given industry rather than consider the royalty that would have been paid with regard to the particular invention.

Transitioning to the Future. Although developing countries and "economies in transition" should have in compliance been with most TRIPS requirements as of January 2000, they need not provide product patents until the year 2005 in any areas that they had previously not patented. This is important to international biotechnology because many countries do not provide product protection of pharmaceutical and/or agricultural products. However, TRIPS does provide some protection for such products during this intervening period.

For countries that did not provide patents for pharmaceutical and agricultural chemical products on terms consistent with Article 27 as of January 1995, a system for filing what has been referred to as "mailbox" applications must exist as of that date (1,20–21). The mailbox applications consist of patent applications that will not be examined until the country must comply with the remainder of TRIPS provisions; however, the mailbox system provides patent applicants with a "date of application" on the day the application is deposited for purposes of novelty and priority (1,20). Patents will issue based on mailbox applications after the expiration of the transitional period.

For products that are subject to mailbox applications, an exclusive marketing right (EMR) must be granted

to provide patentlike protection in the interim period before mailbox applications are examined. An EMR is available where (*1*) a patent on the same invention was granted in another WTO country as of January, 1995 and (*2*) marketing approval for the invention was granted in the same country (1,20). The EMR must last for the shorter of (*1*) five years after the EMR is granted or (*2*) the grant or rejection of the patent (1,20). Thus an EMR enables exclusion of commercial use even in the absence of a patent. However, an EMR provides no possible recourse to oppose noncommercial use.

Even after the transitional period expires, issues are likely to linger concerning the TRIPS requirements. Not only are some present requirements subject to interpretation, but there are also suggestions to modify the TRIPS requirements. For example, it has been suggested, mostly by developing countries, that the scope of items that may be excluded from patent protection should be broadened to include microorganisms, microbiological processes, gene sequences and essential drugs listed by WHO (55). Similarly African countries and Venezuela have argued that compulsory licensing should be allowed for essential drugs. On the other hand, developed countries want to expand patentable subject matter by eliminating the permissible exclusion of plants and animals. Moreover there is some desire to provide protection for holders of "traditional knowledge" or to disallow patents on items available to the public; this desire derives from the belief that TRIPS is not consistent with the Convention on Biological Diversity and International Undertaking on Plant Genetic Resources (32). However, while this issue continues to be studied, the TRIPS agreement is unlikely to be radically changed in the near future as any actual reconciliation with other international agreements is likely to require substantial time and negotiation. While all these issues will likely be discussed either within the context of TRIPS or some new international agreement, TRIPS is likely to be the relevant framework for the near future, considering its long path to enactment and the difficulty in establishing enough consensus to change the agreement.

Plant Protection Under UPOV

Overview. Intellectual property protection of plants should be examined outside the TRIPS framework because TRIPS expressly allows members to exclude plants from patentability, so long as an effective sui generis system exists to provide protection. Although TRIPS does not specify what such a system would require, patentlike rights have been provided to those who develop new plant varieties. In particular, plant breeders have been provided certain exclusive rights of exploitation when they create new varieties of plants; such rights have been referred to as "plant breeders' rights" (PBR) or "plant variety rights" (PVR). The rationale for such a system is very similar to the rationale for the patent system, albeit limited to one particular area. Namely breeders' rights are justified as providing an incentive to develop new plant varieties that further the development of agriculture; also the PVR enables the breeder to obtain a reasonable return on developing a new variety.

International agreement on the importance of such a right has existed since 1961 when UPOV was first signed. UPOV represents a union of member states that have agreed to provide a breeder's right in accordance with the terms of UPOV. Each member country enacts its own national laws to implement at least the minimum standards stated under UPOV. Accordingly, as with patents, an applicant seeking a PBR must seek protection from each country where protection is desired.

Although UPOV was revised in 1972, 1978, and 1991, there are primarily two versions under which countries are currently operating — the 1978 and 1991 versions of UPOV. Thirty countries operate under the 1978 UPOV, while only 12 countries use the 1991 version, which generally provides greater rights to breeders. Countries who were members of the 1978 UPOV had the option, but were not required, to sign the 1991 Convention. As of May 2000 the following countries had signed the 1991 UPOV: Bulgaria, Denmark, Germany, Israel, Japan, the Netherlands, Republic of Maldive, Russian Federation, Slovenia, Sweden, the United Kingdoms and the United States Although the EU is not a formal member of UPOV, it has enacted legislation consistent with the 1991 UPOV (100–101). As countries are currently operating under both systems, the highlights of both will be discussed here, as well as the critical distinctions; additional information is available elsewhere (57–59). Table 5 provides a summary and comparison of the requirements of the 1978 and 1991 UPOV; the distinctions between the two are noted in italics.

Eligible Subject Matter. The fundamental premise of all versions of UPOV is that breeders should be entitled to some exclusionary right with respect to the varieties they create; however, the different versions of UPOV vary in the type of right that must be provided. Under both the 1978 and 1991 UPOV, members may protect a particular variety through either patent protection or a breeder's right consistent with the provisions of UPOV. However, an important distinction is that under the 1978 UPOV — the version that is most widely followed at present — only one type of protection is available for the same genus or species such that dual protection under the patent system and a UPOV-type system is prohibited (2, Article 2(1)). However, there is a grandfather clause that allows members who had previously provided double protection to continue to do so. This provision was included chiefly to accommodate the United States which provided dual protection prior to the 1978 UPOV and wanted to continue do so (2, Article 37). In contrast, under the 1991 UPOV, dual protection is explicitly allowed, namely members may grant the same plant variety a patent as well as a breeder's right (3, Article 40). Accordingly, whereas dual protection was prohibited under the 1978 UPOV, it is now permissible under the 1991 UPOV but still not mandatory.

The scope of varieties that a member must protect also differs between the 1978 and 1991 versions of UPOV. The 1978 UPOV only requires 24 genera or species to be granted protection, and only after eight years of joining UPOV. On the other hand, the 1991 UPOV requires that *all* genera and species eventually be protected; any

Table 5. Plant Protection Under UPOV (non-patent protection)

	1978 UPOV	1991 UPOV
Available protection	Plant *or* UPOV protection *only*; no concurrent protection allowed except for grandfather clause	Plant *and* UPOV *possible*; members can elect for concurrent protection
Scope of varieties entitled to protection	*Requires* that *24* genera or species be granted protection within 8 years of member joining UPOV; *allows* protection for *all* genera or species	*Requires* protection for *all* genera or species within 5 years of signing 1991 UPOV if already a member, or within 10 years if new member
Technical requirements	• Distinct • Uniform • Stable • Commercially novel; grace period *optional*	• Distinct • Uniform • Stable • Commercially novel; grace period *required*
Excluded activities	• Production of the propagating material for purposes of commercial marketing • Offering the propagating material for sale • Marketing the propagating material • Repeated use of the plant variety to commercially produce another variety • Commercial use of ornamental plants as propagating material in the production of ornamental plans or cut flowers	• Production or reproduction • Conditioning for the purpose of • Offering for sale • Selling or other marketing • Exporting • Importing • Stocking for any of the above purposes
Excluded subject matter	Plant variety and its propogating material	• Protected variety • Varieties that are not "clearly distinguishable • Essentially derived varieties • Variety whose production requires repeated use of a protected variety
Exceptions to UPOV right	• Noncommercial use • Experimental use • Use for further breeding • Farmers saving seed from the harvest of a protected variety (farmer's exception)	• Noncommercial use • Experimental use • Use for further breeding *unless it creates an essentially derived variety* (members *may* provide a farmer's exception subject to the interest of the breeder; otherwise, activity will constitute infringement)
Rights during pendency of application	Members *may* provide provisional protection against the "abusive rights of third parties"	Members *must* provide provisional remedies during this time; at a minimum the breeder must obtain equitable remuneration for any activities within the right to exclude
Term of protection	15 years from grant of breeders right; 18 years from grant for vines, fruit trees, forest trees, and ornamental trees	20 years from grant of breeders right; 25 years for trees and vines
Sample countries operating under the convention	Australia, Austria, Belgium, Canada, Finland, France, Ireland, Italy, New Zealand, Norway, Portugal, South Africa, Spain, Switzerland	Denmark, Germany, Israel, Japan, Netherlands, Sweden, United Kingdom, United States

member who was already a member to the 1978 UPOV must do so within five years of signing the 1991 UPOV whereas those who are joining UPOV for the first time have an additional five years to do so (3, Article 3).

A UPOV right can only be obtained after an examination in a member state that determines whether the application satisfies the UPOV technical criteria. UPOV requires that breeders be entitled to a right of priority similar to that under TRIPS; the breeder may utilize the date of the first application in a UPOV member state in determining novelty of subsequent applications for the same variety in other member states. However, UPOV does not mandate how member states must perform the examination. Nonetheless, it does contemplate that the breeder applicant provide information, including propagating material and seed (2, Article 7). In addition the 1991 UPOV suggests that the examination include growing the actual variety (3, Article 12). Accordingly, some countries compare the grown variety with the closest reference variety in the applicant's submitted description as well as a standard benchmark variety (26). The formalities of the application, including whether a description of the breeding process is necessary and whether it needs to be publicly disclosed, are left to individual member countries to determine.

Requirements for UPOV Protection. UPOV rights are provided to (*1*) a plant "variety" in accordance with the

UPOV definition that (2) meets the technical criteria under UPOV of being distinct, uniform, stable, and commercially novel (2, Article 6; 3, Article 5). Although the technical criteria are essentially the same under the 1978 and 1991 versions, the 1991 UPOV provides a definition of a "variety," whereas the 1978 UPOV did not. In particular, the 1991 UPOV explicitly defines a qualifying variety as "a plant grouping within a single botanical taxon of the lowest known rank which grouping can be defined by features characterizing a given genotype or combination of genotypes, and is distinguished from any other plant grouping by the expression of at least one of the said characteristics." (3, Article 2(i)).

Both Acts require that the variety be sufficiently distinct. They both provide that a variety meets this requirement if it is "clearly distinguishable" from another variety that is "common knowledge" at the time (2, Article 6; 3, Article 7). A variety is likely to be "clearly distinguishable from another" if the variety is different from one or more morphological characteristics (e.g., leaf shape and flower color) or physiological characteristics (e.g., disease resistance and hardiness) of other varieties known at the time of the application (26). The more difficult issue is to determine what was known at the time of the application. What is common knowledge appears similar to the concept of "prior art" under patent laws. Just as TRIPS does not define what should be considered prior art in determining novelty, UPOV does not define "common knowledge." However, both Acts include applications filed in the definition of "common knowledge" although such applications are not publicly available (2, Article 6(1); 3, Article 7). In addition the 1978 UPOV provides that common knowledge may include description in prior publications or commercial sale (2, Article 6(1)(a)).

Besides being distinctive, the new variety must be both "uniform" and "stable." The variety must be uniform with respect to features of its reproduction or propagation (26). The variety must maintain its essential characteristics after repeated reproduction or propagation, or if the breeder has defined a limited cycle or reproduction, it must remain true at the end of each cycle (2, Article 6(d); 3, Article 9). The amount of stability required is a function of the species and variety. However, it is generally assumed that a variety is stable if it exhibits a reasonable level of uniformity in its essential characteristics for a minimum of two successive growing seasons (26).

Finally, the breeder must establish "novelty" of the variety. Under the 1978 UPOV, novelty exists as long as the breeder does not authorize the variety to be offered for sale or market in the state where the breeder seeks protection prior to application. Novelty may also exist even after the breeder authorizes the variety to be offered for sale or marketed in a state where protection is sought if that state provides a one-year grace period after the commercial activity to file an application; states may provide a grace period of up to four years for commercial activity that occurred in other states. In contrast, the 1991 UPOV provides that members must provide a one-year grace period.

Commercial activity that affects novelty differs between the 1978 and 1991 versions of UPOV. Under the 1978 UPOV, the only commercial activity that defeats novelty is activity pursuant to the authority of the breeder (2, Article 6(1)(b)). However, under the 1991 UPOV, protection is barred if the "propagating or harvested material" of the variety was "sold or otherwise disposed of to others ... for the purpose of exploitation of the variety"; as this no longer mentions consent of the breeder, it includes all commercial activity regardless of whether the breeder consents (3, Article 6).

UPOV Right

Scope of Protection. Although member states can always provide additional protection, at a minimum they must enable the breeder to exclude certain activities by others. Under the 1978 UPOV (Article 5(1)), the breeder can exclude others from the following:

1. Production of the propagating material for purposes of commercial marketing,
2. Offering the propagating material for sale,
3. Marketing the propagating material,
4. Repeated use of the plant variety to commercially produce another variety,
5. Commercial use of ornamental plants as propagating material in the production of ornamental plants or cut flowers.

Accordingly, under the 1978 UPOV, the reproductive material of a protected variety may be used to produce another variety as long as it is not for the purpose of selling the protected variety and as long as there is no repeated use of the material of the protected variety for the commercial production of the new variety. Thus a second breeder could generally breed and commercialize a new variety without providing any compensation to the initial breeder of the protected variety.

The activities that the breeder can exclude under the 1991 Act are even broader than those under the 1978 Act. The 1991 Act provides seven acts of exploitation for which authorization from the breeder is required (3, Article 14):

1. Production or reproduction
2. Conditioning for the purpose of propagation
3. Offering for sale
4. Selling or other marketing
5. Exporting
6. Importing
7. Stocking for any of the above purposes

The breeder must authorize these seven activities with respect to propagating material of the protected variety as well as varieties that are not clearly distinguishable from the protected variety (in accordance with the UPOV definition of distinctiveness). In addition the breeder must authorize these activities with regard to harvested material (including entire plants or parts of plants), if the harvested material has been obtained through the unauthorized use of propagating material and the breeder has had no reasonable opportunity to exercise

his right in relation to the propagating material (3, Article 14(1)–(2)). The 1991 UPOV permits, but does not require, that member states provide the breeder with the right to exclude products made directly from harvested material of a protected variety through unauthorized use of harvested material where the breeder has had no reasonable opportunity to exercise his right in relation to the harvested material (3, Article 14(3)); a member state that adopts this provision would preclude farmers from saving harvested seeds. However, member states can instead elect to provide a farmer's right similar to that which is available under the 1978 Act. If the breeder is not provided rights with regard to farm-saved seeds (if a farmer's exception is provided), the breadth of any such farmer's exception is within a member state's discretion as long as it protects the breeder's "legitimate interests" (3, Article 15(2)). For example, the breeder's interests may be safeguarded by only providing a farmer's right to small farmers, or farmers of certain crops.

Under both the 1978 and 1991 UPOV, provisional protection may be granted to the breeder during the pendency of the UPOV application. Under the 1978 UPOV such protection may, but need not, be provided against "abusive rights of third parties" during the pendency of the application (2, Article 7(3)). Under the 1991 UPOV, however, provisional remedies during this time period are required. In particular, members to the 1991 UPOV *must* provide measures to safeguard rights during pendency so that at a minimum the breeder obtains equitable remuneration for any activities within the right to exclude. However, member states can reduce the impact of this provision by only applying it against those who knew of the application (3, Article 13).

Exceptions. A major distinction between the 1978 and 1991 UPOV is that under the 1978 Act the breeder cannot prevent farmers from saving products of their harvest to replant, namely farmers who save seed from the harvest of the protected variety do not infringe when they later replant the seed. However, under the 1991 Act, although a member state *may* provide a farmer's exception subject to the interests of the breeder, it can also allow such activity to constitute infringement — something that was not a possibility under the 1978 Act.

Other than the farmer's privilege, the exceptions to the breeder's right under the 1978 and 1991 Acts are similar. Both provide for three exceptions to the breeder's rights: private noncommercial use, experimental use, and use of the protected variety for further breeding (3, Article 1(i)–(ii)). However, with respect to use for further breeding, the breeder has more rights under the 1991 UPOV. In particular, the breeder has a right against those who use the protected variety to breed additional varieties if the new variety is "essentially derived" from the protected variety (3, Article 15(1)(iii)). A variety is "essentially derived," and subject to the breeder's right to exclude, when it is predominately derived from the initial variety and, except for differences that result from the act of derivation, it displays the same essential characteristics that result from the genotype, yet is clearly distinguishable from the initial variety (3, Article 7(5)(b)). This new requirement is intended to remove the unfairness that operates under the 1978 Act, whereby a genetic modification of a protected variety could enable a second breeder to obtain rights without providing any recognition or compensation to the original breeder. UPOV explicitly provides examples of what should constitute essentially derived varieties: those obtained by selection of a mutant, of a somaclonal variant, or of a variant individual from plants of the initial variety, backcrossing, or transformation by genetic engineering (3, Article 14(5)(c)). Thus, although protected varieties may continue to be used as a source of initial variation, if the resulting variety falls within the definition of an essentially derived variety, authorization from the initial breeder is required.

Two final limitations to the PBR under UPOV are compulsory licensing and exhaustion of the PBR. Compulsory licensing is permissible under both versions of UPOV. However, it is only possible if it is in the public interest and equitable remuneration is provided to the breeder (2, Article 9; 3, Article 17). Only the 1991 UPOV provides for a principle of exhaustion. The UPOV exhaustion principle is analogous to that previously discussed with regard to patent protection; in particular, it provides that the breeder's right will not extend to acts concerning material of the protected variety (or essentially derived variety) if sold or otherwise marketed by the breeder unless further propagation is involved or export of the variety is involved (3, Article 16).

Comparison of UPOV and Patent Protection. It is important to note the distinctions between UPOV and patent protection, particularly for countries where only one type of protection is allowed such that an informed decision on the optimal type of protection may be made. First, it should be noted that whereas patents protect all "inventions" (whether the inventions are products or processes) UPOV seeks to protect one type of product — plant varieties. UPOV can only protect the variety itself, whereas patent protection, if available, can also cover a transformed gene or process of making such a gene. Second, rights against imports are only available under the 1991 UPOV, whereas any country that is a WTO member can prevent imports of patented products. Third, both UPOV Acts explicitly allow use of a protected variety for not only noncommercial purposes but also for the purpose of breeding other varieties. Although there is the possibility that such use could be permissible as a "limited exception," to patent rights under Article 30 of TRIPS, there is at least uncertainty as to whether that is possible, unlike the explicit exception provided under UPOV.

In one respect, the 1991 UPOV may provide more extensive protection than patent rights because of the requirement that states provide protection during the pendency of the UPOV application. Although WTO members may provide similar protection, TRIPS does not require any rights be provided prior to the issuance of a patent. In addition the requirements necessary to obtain UPOV rights potentially require less disclosure than a patent application and potentially no disclosure to the public. Whereas TRIPS mandates disclosure of the invention to obtain patent protection, UPOV does not require a description of the "invention," and does not

require any disclosure that would allow replication by another. There is no requirement of either a deposit or of public access to the deposit. Although a member state may impose such requirements on an applicant, UPOV itself does not.

The term of protection may also be different if patent protection is chosen versus UPOV protection. Notably TRIPS calculates the term of patent from filing date, whereas UPOV calculates it from the date of grant of the right. While TRIPS requires a minimum period of 20 years from the filing of an application, the term is 15 years from the grant of the breeder right under the 78 UPOV (18 for vines, fruit trees, forest trees, and ornamental trees) or 20 years from grant of the breeder right under the 91 UPOV (and 25 years for trees and vines).

PRESENT SYSTEMS OF PROTECTION

Europe

Relevant Laws. Most European patents are sought from the EPO rather than from patent offices of individual countries. EPO law includes (*1*) the EPC, which is an agreement among European countries (24), (*2*) regulations implementing the EPC, and (*3*) judicial interpretation of the EPC by EPO courts (60,61). EPC provides a streamlined process for obtaining patents in Europe — a single application to the EPO may lead to national patents in every member state; EPC thus provides a shortcut to obtaining patent protection in multiple European countries.

Eighteen months after an application to EPO, a European Patent (EP) application is published. Subsequently EPO will determine whether to grant a patent (an EP patent). If EPO decides to grant a patent, the EP patent is published and available for opposition on almost any ground on which EPO could have denied patentability. If the EP patent survives the opposition period (is not revoked in its entirety), it will become transformed into individual national patents once certain procedural requirements are fulfilled. Once those requirements are met, EPC ceases to govern the EP patent, and national laws control how the national patents are enforced in each country (4,27).

In addition to the EPO law on patentability, patentability requirements in the European Union (EU) will be discussed. The EU is a union of presently 15 countries. It was founded to further political, economic, and social cooperation; formerly, the EU was known as the European Community (EC) or the European Economic Community (EEC). Unlike EPC, the EU is an organization that governs many facets of national law through a governmental structure including a Parliament, Council, and Commission that roughly represent branches of an executive government. These EU bodies can require member states to take actions through the issuance of regulations (immediately and directly binding on member states) or directives (binding as to result only and usually not immediately effective). In the area of biotechnology, the EU has enacted a directive concerning patent protection for biotechnology that includes both patentability and enforcement of such inventions (5). Member states are to be in compliance with this by July 20, 2000 (5). It is unlikely that all members will actually be in compliance by this time since the EU directive is presently subject to legal challenge by several member countries (62,63). However, it is perhaps more important to note that the EPO has issued regulations that adopt many of the articles of the EU Biotechnology Directive. Thus, even if EU member states have not yet altered their national laws, that may be of minor import because most patent applicants utilize the EPO's streamlined application system, rather than applying to individual national patent offices. This is particularly true because all of the current EU member states are also members of the EPC. Most importantly, all of the current EU member states are also members of EPC. Accordingly, because the EPO has made its laws consistent with the EU Directive, the laws and policies of the EPO are presently of paramount importance and will be the focus of this section. More information on patent laws of individual European countries is available from other sources (64–67).

Patentable Subject Matter. The first hurdle to patenting biotechnology is to establish that there is a patentable "invention." Natural discoveries are excluded from the scope of patentable inventions under EPC. Accordingly an issue that has been raised is whether certain types of biotechnology, including isolated natural substances, are natural discoveries and thus unpatentable. Under current EPO policy, "biological material," which is defined as "any material containing genetic information and capable of reproducing itself or being reproduced in a biological system" may be patentable regardless of whether it previously occurred in nature if it is "isolated from its natural environment or produced by means of a technical process" (5, art. 2; 40–41). Isolated elements may be patentable even if the elements are isolated from the human body such as gene sequences or partial gene sequences (5, art. 3(2); 29). However, the mere discovery of the elements of the human body, including embryonic stages of the human body is not considered to be patentable (5, art. 5; 40–41,61).

Specific Exclusions

1. Methods of treatment. The patentability of gene therapy and associated technology is an issue due to a provision of EPC that defines methods of treatment and diagnosis of either humans or animals to lack "industrial applicability," therefore making such methods unpatentable. In particular, Gene therapy can be considered a method of treatment; diagnostic kits utilizing genetic engineering can be considered as methods of diagnosis. However, EPO courts could interpret this exclusion narrowly to allow some methods to be patentable. For example, although the EPC excludes diagnostic methods, the EPO has interpreted this exclusion narrowly to only exclude diagnostic methods whose results can be immediately used to decide on a course of medical treatment; if the method provides interim results in the course of making a diagnosis, the method is patentable (67). Methods of diagnoses

are only excluded if actually carried out on the body; accordingly diagnosis of body tissues or fluids removed from the body may be patentable (67).

Even without relying on judicial construction of the method exclusion, some patent protection may be available pursuant to a related part of the EPC statute. In particular, the EPC explicitly states that the medical treatment prohibition "shall not apply to *products*, in particular substances or compositions, for use in any of these [prohibited] methods." (68, Article 52). Therefore, even though patent protection is excluded for the actual method of treatment, products used for such treatment may still be patentable subject matter. Thus, although both somatic and germ-line therapies are considered unpatentable methods of medical treatment, the EPO has indicated that products such as genetically modified cells intended for use in somatic gene therapy are patentable and would not be automatically barred (39). However, patentability may be ultimately precluded pursuant to another provision of the EPC or the EU directive that bars patenting of inventions that are contrary to *ordre public* or morality. Products for use in germ-line gene therapy, in particular, may be found in violation of such a provision (39).

2. Plant and animal varieties. EPC Article 53(b) declares unpatentable "plant or animal varieties or essentially biological processes for the production of plants or animals; this provision does not apply to microbiological processes or the products thereof." This provision does not bar the patenting of transgenic animal (65,80–81). However, other provisions of EPC, in particular, the exclusion of inventions that raise morality concerns could pose a problem even if this provision does not (4, art. 53(a); 80–81). The patentability of plants, on the other hand, is less certain.

   The patentability of plant varieties is presently unclear because of conflicting EPO case law and the yet to be implemented EU Directive. The EU Directive as well as recent amendments to the regulations implementing EPC provide that plants are patentable "if the technical feasibility ... is not confined to the particular plant or animal variety" (7,61). This would suggest that transgenic plants could be patented and patent claims could cover varieties so long as it covered more than a single variety. Not all EPO case law is consistent with this — recent cases have held that so long as a claim to a transgenic plant may cover a plant varieties, it should be precluded from patentability (27,69–73). However, in the most recent EPO case T1054/96, In re Novartis, the Enlarged Board of Appeal declared that a claim to a transgenic plant could be acceptable even though it may include specific plant varieties (74). Although this would appear to make EPO practice consistent with the EU Directive, whether this will continue remains to be seen.

   It should also be noted that article 53(b) states that although processes for creating the ambiguous "varieties" are unpatentable, other types of biotechnological processes may be patentable. In particular, microbiological processes and the resulting products are patentable — so long as they do not include the improper "varieties" and also meet the technical requirements of patentability. A microbiological process consists of "any process involving or performed UPOV or resulting in microbiological material" (7,61). A process for the production of plants or animals is deemed to be "essentially biological" if it "consists entirely of natural phenomena such as crossing or selection" (7,61). Accordingly only traditional breeding of plants and animals are excluded "essentially biological" processes under Article 53(b). Potentially patentable microbiological processes could thus include methods of obtaining, transforming and using microorganisms such as viruses and bacteria.

Nonspecific Statutory Exclusion. Inventions, whose publication or exploitation would be contrary to *ordre public* and morality, are excluded from patentability even if they otherwise constitute an invention and meet the technical patentability requirements. The important query here is the meaning of *ordre public* and morality. One EPO board has considered the term *ordre public* to include the protection of public security and the physical integrity of individuals within the society, including protection of the environment (71). However, there is no single unitary concept of either *ordre public* or morality in all members of EPC. Moreover it is unclear whether the accepted standards of conduct to which the invention is to be compared are those within one member state or all member states. An invention could be deemed to be lacking morality only if it was contrary to the accepted standards for all member countries, or, an invention could be deemed to be lacking morality if it was contrary to any one country. In addition it has been noted by certain EPO courts that there must be an "overwhelming consensus" of opinion before Article 53(a) will bar patentability (76,77). In any event, there have been instances where this has been raised as a ground for denying a patent on genetic engineering inventions. In particular, patents on transgenic animals, plants, and isolated gene sequences have been opposed based on this provision (76,77–78).

Although not an absolute bar to patenting transgenic animals, the *ordre public* and morality exclusion may preclude patents on some animals. To determine whether a transgenic animal is barred by this provision the EPO has previously used a balancing test that takes into account the suffering of animals and possible dangers to the environment, on one hand, with the potential benefits to humans, on the other. Accordingly, in the first and most famous case where this was applied, the Harvard OncoMouse application, it was held that because a mouse genetically engineered to be predisposed to cancer had such substantial utility to humankind, that it outweighed the suffering of the individual animal (81). However, not all transgenic animals have been held to meet this test. For

instance, the EPO has denied applications of transgenic animals whose utility is to study baldness (27).

In the future a modified balancing test may be applied as the EU Directive, as well as the amended EPO regulations, state a differently worded test. In particular, patents on processes for creating transgenic animals as well as the resulting animals are precluded if it is "likely to cause them suffering without any substantial medical benefit to man or animal" (7,61). It literally seems to require some showing of "*substantial* medical benefit" where a process is "likely" to cause animal suffering. However, it is unclear what would constitute a "medical" benefit, let alone a "substantial one." For example, cows that are genetically engineered to produce more milk would appear to be of questionable "medical" benefit. Moreover it is unclear what "suffering" must consist of. Some might even argue that the cow is "suffering" by having its genes altered. On the other hand, unlike the OncoMouse, the altered cow is not programmed to a hastened death.

The appropriate test to apply to plants is unclear. It has been noted that the OncoMouse balancing test is not the only way to evaluate violation of Article 53(a) (71). However, in the absence of a balancing test, the vaguer standard of *ordre public* and morality must be applied. This appears difficult to establish. For example, in a case where a transgenic plant engineered to be pesticide-resistant was opposed as violating Article 53 (a) (among other provisions), the EPO court dismissed alleged danger to the environment on the ground that surveys of the general public showing opposition to genetic engineering were inadequate; actual danger, rather than the "mere possibility" of danger was stated to be necessary and infringement of environmental regulations alone was not considered adequate (71).

Although the interpretation of the morality requirement is still unclear, the EU Directive and the EPO regulations provide a noninclusive list of what will per se violate morality or *ordre public*. Current examples of biotechnology that are unpatentable because of this include processes for cloning humans, processes for germline gene therapy, and use of human embryos for industrial or commercial purposes (7). Processes for modifying the genetic identity of animals which are likely to cause them suffering without any substantial medical benefit to human or animal, and animals resulting from such processes" (7). Because EPO has amended the regulations to EPC to essentially adopt the EU Directive wholesale, the categorical bars to patentability stated in the Directive presumably apply (61). In addition, although EPO has noted that "the provisions of Article 27(2) of TRIPS will be considered" (39), it is presently unclear how that will affect the analysis, if at all, as there are notable similarities in the language, as indicated on Table 6. In fact, it may be that the EPO case law is drawn upon by the TRIPS council and member countries as the most pertinent and largest body of law in interpreting what is otherwise vague language.

Patentability Requirements. As the focus of this article is on issues relating to patenting biotechnology, major issues regarding technical requirements will be noted. However, more detailed information on the specifics of each technical requirement under EPC alone is discussed elsewhere (27,64).

Novelty. EPO takes an "absolute novelty" approach to determining whether an invention is patentable.

Table 6. Exclusions from Patentability

	EPC/EU[a]	TRIPS
Categorical exclusions of specific subject matter	• Methods for treatment or diagnosis of human/animal body • Plant and animal "varieties" • "Essentially biological processes" for the production of plants or animals except "microbiological" or other technical processes • "Human body, at the various stages of its formation and development"	• Diagnostic, therapeutic and surgical methods for the treatment of humans or animals; • Plant and animals other than micro-organisms, • "Essentially biological processes" for the production of plants or animals other than "nonbiological" and "microbiological processes"
Nonspecific exclusions	Inventions whose commercial exploitation would be contrary to "*ordre public* or morality" Listed examples: • Processes for cloning humans • Processes for germ line gene therapy of humans • Commercial uses of human embryos • Methods of genetically modifying animals that are "likely to cause them suffering without any substantial medical benefit to man or animal," and any animals resulting from such methods	Inventions whose commercial exploitation must be prevented to protect "ordre public or morality," including to protect human, animal or plant life or health or to avoid serious prejudice to the environment

[a]Mandatory exclusions.
[b]Permissive exclusions.

Accordingly relevant prior art includes any written publications, abstracts, or drawings; prior use anywhere in the world; and prior, though later published, patent applications (3, Article 54). However, EPO does recognize a very limited grace period of six months before the filing of the application for disclosure made by the applicant at an official, or officially recognized, international exhibition (3, Article 55). In addition there has been some discussion concerning whether a more expansive grace period should be provided.

Novelty may be satisfied for isolated products found in nature so long as it is done through a technical process. Thus a new substance discovered as being produced from a microorganism, or a previously unknown gene or protein, would be considered novel and patentable if the other patentability requirements were satisfied (40).

Novelty may be established in cases of newly discovered uses of previously patented compounds. In particular, a compound may be patented with a use limitation even if the compound itself was previously known as long as its use for treatment or diagnosis was not previously known (first medical use). A second patent can be obtained for the newly discovered use of a previously known compound, even if the compound was previously known to have a use, if it is newly discovered to be useful for additional therapeutic purposes (second medical use). In either event, it should be recalled that only the compound, and not the process of using the compound, can be patented so that the medical treatment prohibition is avoided. This often results in just a technical distinction of how the invention is claimed — a claim reciting a "method of treating X using substance Y" would be considered unpatentable, whereas a first medical use claim reciting "substance Y for use as an active pharmaceutical substance" would be acceptable, as would a second medical use claim stating "use of substance Y for the preparation of a pharmaceutical compound for the treatment of X."

Industrial Applicability and Inventive Step. The invention must also have "industrial applicability," which requires that the invention be useful in "any kind of industry, including agriculture." (3, Article 57). In general, this is a fairly easy requirement to meet. However, industrial applicability may be an issue with some nucleic acid sequences. The EPO has noted that it is "questionable" as to whether sequences that have no known use other than as probes satisfy the industrial applicability requirement (40). Similarly most inventions will be found to have met the "inventive step" requirement unless the invention is "obvious to try" by a person skilled in the art with a reasonable chance of success. However, in the area of biotechnology, the application of the inventive step requirement is uncertain as different tests have been applied by EPO and national courts (82–83).

Description and Deposit. The disclosure requirement merely requires sufficient disclosure such that the invention can be carried out by a skilled person for the claimed subject matter using common general knowledge and the information provided in the application. No specific examples are required in the application if a skilled person could carry out the invention without undue experimentation. In addition, pursuant to the EU Directive, the patent application for a sequenced or partially sequenced gene must be disclosed; however, that alone may not be sufficient — there may be other issues with patentability requirements for expressed sequence tags (ESTs) pursuant to Articles 52(1) and 56 (40).

A deposit is required if necessary to enable reproduction of the invention claimed in an application. A deposit is required, for example, if a specific microorganism cannot be reproduced based on the application alone. However, if a protein or nucleic acid is sufficiently defined in the application such that it may be synthesized by one of skill in the art, it need not be deposited.

If a deposit is required, it must be made no later than the filing date. The deposit may be made to any institution recognized under the Budapest Treaty as well as other institutions as noted in the Official Journal of EPO. The deposit may become available to others after the patent is published. However, if the applicant informs EPO prior to the publication of the application of a preference to limit access, there will be no public access until an EPO patent is granted, or 20 years after the filing date if the application is refused or withdrawn. If public access is not specifically requested to be restricted, the deposit will be available after the application is published to any person who requests it on the condition that the person does not make the material available to a third party and if the person is using the deposited material for experimental uses only.

Patent Rights

Challenging an EP Patent. Even after an EP patent has been issued, there is some uncertainty involved because the patent is subject to a period of opposition during which anyone, including competitors or special interest groups, may file an opposition against the patent. Oppositions may be filed within nine months after the patent grant is published. There are many grounds for contesting a published EPO patent, including lack of novelty (Articles 54–55), lack of inventive step (Article 56), lack of industrial application (Article 57), improper subject matter (Article 53), noninvention (Article 52) and inadequate disclosure (Articles 100(b) and 83). Opposition may result in amendment of the patent scope or complete revocation of the patent.

It should be noted that opposition proceedings are conducted by EPO before the EP patent becomes a collection of individual national patents. Accordingly successful opposition of an EP patent is a much cheaper and more efficient to challenging parallel national patents in different countries. However, even if a patent withstands opposition in EPO, it may still be subject to revocation in national proceedings. An EPO decision in opposition proceedings has no binding effect, even on the same parties in later national proceedings for which national law, rather than EPC, will govern.

There may be inconsistency between the EPO system and national courts as well as inconsistency within courts of same nations. For example, in the recent biotechnology case of *Biogen v. Medeva*, several different courts reached different decisions on whether the Biogen patent was valid, and even courts that reached the same result had

different reasoning (82–87). The Biogen patent was first subject to opposition within EPO, after which the EPO Technical Board of Appeal maintained the patent (84). However, once the Biogen EP patent was transformed into national patents, trouble began. Biogen first sued Medeva for using its patented composition. In its defense, Medeva asserted that Biogen's patent was invalid on a number of grounds. In the initial court, Biogen's patent was considered valid on the same grounds as the EPO had found (85). However, the next court to hear the case, the High Court of Justice, came to the opposition conclusion — namely that Biogen's patent was invalid for all asserted grounds — that it was insufficient, obvious, and possibly not even an "invention" (86). Finally, the House of Lords also declared Biogen's patent invalid, but for different grounds than the High Court of Justice; in particular, the House of Lords declared the critical issue was one that had not been formally raised in any of the decisions of the prior courts (87). The House of Lords determined that the claimed invention was too broad and held that the patent from which priority was claimed could not support the patent at issue (87).

Scope of Protection. Granted EP patents have the same scope of protection within a given country as do patents granted by the individual country's patent office (4, Article 2(2)). EPC generally does not govern activity after the opposition period with two exceptions. EPC provides that there should be uniform protection conferred by European patents pursuant to Article 69. EPC provides that there should be protection for the "direct product" of a patented process under Article 62(2). However, because national courts may interpret identical provisions differently, there is bound to be some inconsistency despite the requirement of "uniform protection."

The scope of protection for biotechnological matter is still under development. Although EPO issued regulations in June 1999 implementing some of the provisions of the EU Directive, it did not include any of the EU Directive provisions concerning the scope of protection of biotechnological inventions (61). However, the EU Directive will be considered here as all EU members are EPC members. The EU Directive provides that patents on biological material provide protection on any biological material derived from the initial patented material through propagation or multiplication; accordingly offspring of a patented transgenic animal would be within the scope of protection (7). In addition, for patented processes that produce biological material with specific characteristics as a result of the invention, protection is extended to any biological material "directly obtained" from the patented process as well as any biological material "derived from the directly obtained" material through propagation or multiplication (7).

The EU Directive provides limitations on the scope of protection for certain biological material. In particular, it is stated that no protection is provided for material obtained from material placed on the market by the holder of the patent with his consent (7). In other words, once the patent owner sells the patented product, rights for subsequent products are extinguished. In addition the Directive provides a farmer's rights similar to the rights provided under UPOV as well as a corollary right for animal breeders (7). Member states have control over how much of an exception to provide with respect to the newly created animal breeder's right (7). Finally, compulsory licensing is also specifically provided for, although most of the specifics of such licensing are left up to individual countries (7).

Another limitation on infringement that has been established in some countries is an exception for experimental use. Many EU and EPC countries exempt from infringement "acts done for experimental purpose relating to the subject matter of the patented invention" (64,88); this is often referred to as an experimental use exception. It is unclear whether experiments for market approval constitute infringement and whether there should be any difference when testing is done by manufacturers in preparation for the sale of generic drugs (89–90). However, the most recent case on this issue in Germany found that clinical trials on a patented compound to determine its properties and effects were protected by the experimental use exception even if conducted with the goal of acquiring marketing approval; the only limit to the exclusion was stated to be if the tests were solely directed at determining commercial facts such as market needs and price acceptance (91). Even if national courts uniformly interpret a broad exception, such an interpretation could be challenged as violating TRIPS and not protected by the "limited exception" provision of Article 30 of TRIPS.

Enforcement Issues. An important enforcement issue in Europe is that in recent years, some courts have issued "cross-border" or "pan-European" injunctions against defendants in patent infringement actions (92,93). An injunction is an order issued by the court; in the context of a patent infringement action, an injunction is often issued against a defendant who is infringing. The injunction orders a defendant to stop infringing or be subject to court sanctions. What is notable about cross-border injunctions is that typically courts only issue injunctions within their territorial boundaries, that is, within the nation in which the court sits. However, cross-border injunctions are injunctions in which the court orders a defendant not to infringe in other states or countries (across its own borders).

The availability of cross-border injunctions are particularly useful to patent owners desirous of relatively inexpensive means of stopping a defendant from infringing in multiple nations. While a patent owner attempting to enforce an EP patent would traditionally have to pursue litigation in multiple jurisdictions with potentially different verdicts, a cross-border injunction enables a single litigation to potentially enjoin a defendant's action in all European countries.

The legal basis for cross-border injunctions is based on a provision of the Treaty of Brussels (94) that allows a case to be brought before the court of any defendant's residence when there are multiple courts. Through a liberal interpretation of this provision, courts, and particularly the district court of the Hague, have found themselves competent to hear cases against defendants in patent cases as long as one defendant was Dutch (92–94,97). Jurisdictions that have issued cross-border injunctions

include the Netherlands, England, Germany, Belgium, and France, although the majority of such injunctions have been issued by Dutch courts (92–94,97). Although subsequent case law has limited application of this principle in cases against foreign defendants (95), the potential for such injunctions still exists in the absence of an effective means to enforce patents issued by EPO. Indeed, commentary on cross-border injunctions has noted that the only long-term solution is a more uniform enforcement mechanism (97). Such a possibility is in fact a potential reality.

Additional Protection

Pharmaceuticals. In addition to patent protection, manufacturers of patents on pharmaceuticals in the EU are entitled to some patentlike exclusionary rights after the term of a patent via a supplementary protection certificate (SPC) (98–99). A SPC allows a limited exclusivity right after the ordinary 20-year patent term expires. The term of SPC is calculated as the time elapsed between the date of filing the application for the patent and the date of the first marketing authorizing minus five years, up to a maximum of five years. There are four requirements that must be met to obtain a SPC: (*1*) The product sought to be protected by an SPC must be protected by a basic patent in force, (*2*) the product must not have been granted marketing authorization, (*3*) the product must not have already been the subject of an SPC, and (*4*) the marketing authorization is the first to place the product on the market (Article 3). Only the patentee, not a licensee, is entitled to apply for a SPC.

A SPC allows its owner the right to exclude others from the commercial sale of covered goods. However, because a SPC is not a patent, the owner of SPC cannot exclude others, such as generic drug manufacturers, from testing the drug that is no longer patented. SPC effectively achieves the same result as an infringement exception provided to manufacturers of generic drugs under the U.S. system.

Plants. Regardless of the status of patenting plant "varieties" under EPC or the EU Directive, an alternative means for protection exists both under the EU and under national laws of individual countries. Plant Variety Rights within the Community are governed by an EU regulation that provides a communitywide right; this right is additional to any rights available under national regimes. This means that for plants, protection is potentially available through EPO (subject to the bar on plant varieties), through national patent systems of EPC or EU countries; in addition PVR may be available either from individual countries or from the EU. The benefit of obtaining rights under the EU system is that the breeder obtains an exclusionary right that applies throughout the entire EU system with only one application. Plant variety protection pursuant to the Community regulation is consistent with protection available under the 1991 UPOV (100–101).

For patented plant-related inventions, additional protection is available in the form of a patentlike right along the same lines as for pharmaceuticals. Since 1996, an SPC has also existed for "plant protection products" that are protected by a patent (102). The term and procedural requirements of an SPC for such products is identical to the one provided for pharmaceuticals. What differs is the product that is protected. In relation to plant products, those that qualify for SPC protection are defined as "active substances and preparations containing one or more active substances ... intended to": (*1*) protect plants against harmful organisms, (*2*) influence the life processes of plants, such as a growth regulator, (*3*) preserve plant products, (*4*) destroy undesirable plants, or (*5*) destroy parts of plants or otherwise minimize undesirable growth (102).

Future Developments. Patent protection for biotechnology in Europe is likely to continue to be in a state of flux in the near future due to the existence of different systems for obtaining and enforcing patents. A uniform system for obtaining as well as enforcing patent rights throughout the EU member states has been previously envisioned, although it has not yet become a reality.

In particular, although the Community Patent Convention (CPC) provides for such a system, it has not been ratified by all EU members. Because of political and constitutional reasons, CPC has not come into force, although it was first signed in 1975 (65). At this point, it is anticipated that the EU will cease efforts toward effectuating CPC, and instead work toward enacting a regulation to create a unitary EU patent (103); the EU Commission is planning to propose such a regulation in the year 2000 but no such regulation has been proposed as of June 2000. The EU has been clear that such a system would be largely consistent with the EPC provisions and at least for a transitional period, coexist with the present two-tiered system of national and EP patents (103). However, even if such a regulation were adopted, the protection of biotechnology is likely to continue to be remain unsettled for quite some time.

Japan

Japanese patent law has changed markedly in the last 10 years, both because of the TRIPS agreement and because of international pressure to conform its laws to those of Europe and the United States. Some doctrines that have been established for decades in the United States have only recently been established in Japan. It is likely that amendments and clarifications to its patent laws will continue. Current Japanese Patent Office (JPO) practice with respect to biotechnology can be determined by examining the JPO Guidelines, as well as the JPO's response to some hypothetical biotechnology patent examples, all of which are publicly accessible through the JPO web site (104,105). Because of the availability of this information, as well as the fact that Japanese patent law will likely continue to evolve, this discussion will provide an overview of recent changes to Japanese patent law that affect biotechnology rather than attempt to provide a comprehensive description of all laws concerning biotechnology.

Patentability. The JPO has clarified that biotechnology inventions are not per se precluded from patentability (39–40,105). Therefore the usual patentability standards

apply to such inventions. However, application of those standards has posed some difficulties in the area of biotechnology. For example, it is sometimes difficult to determine when genetic engineering inventions are novel or have an inventive step. Nonetheless, the JPO has indicated a wide variety of biotechnological innovations to be considered patentable subject matter, so long as they can still meet the technical patentability requirements. In particular, living unicellular organisms, animals and animal parts, plants, and plant parts are all considered patentable subject matter. As a "general rule," human-derived products are said to be eligible for patentability, although humans themselves may be barred under a morality-based exclusion (40). New uses of known compound are also considered patentable (40).

However, "industrially applicable" has been statutorily defined to exclude certain subject matter that narrows the scope of patentable biotechnology. As in EPC certain types of "medical activity" are considered not industrially applicable. Of particular relevance to the pharmaceutical area is that methods of treating humans by surgery or therapy, as well as diagnostic methods practiced on the human body are all excluded as being not industrially applicable; however, pharmaceutical compositions for use in the unpatentable methods can be industrially applicable. In addition methods of treating animals are patentable so long as there is no attempt to patent treatment on the human body. The relevance of these exclusions to biotechnology is that gene therapies are not patentable because they are considered a method of treatment within this exception. However, pharmaceutical products produced by gene therapy techniques may be eligible to be patented (40).

Another bar to patenting biotechnology is that inventions "liable to contravene public order, morality or public health" may be considered unpatentable under Section 32 of the Japanese Patent Laws (106). This language has been the basis for denial of patents on new medical treatments as well as methods of breeding new plants or animals. In addition JPO has indicated that this may also preclude patents on humans as well as human organs (40).

While not a bar to patentability, it is important to be aware of deposit requirements for biotechnology. For inventions concerning microorganisms, the microorganism must generally be deposited with an institution designated by JPO or international depositary authorities. This requirement can be met simultaneously with application but is improper after filing. There are certain situations where microorganisms need not be deposited, for example, if it can be created by a person skilled in the art based on the specification description alone.

Patent Rights

Opposition and Patent Term. Prior to 1996 the JPO published applications for opposition by third parties after examination, but prior to the issuance of an actual patent and tied the patent term to this opposition period. The pre-1996 patent term was 15 years from the date of publication of the postexamination (and pre-grant) application, with a maximum duration of 20 years from filing date (106, Article 67(1)). This scheme has been substituted with an opposition procedure *after* the patent grant, similar to the situation described with regard to EPO (106, Article 113). Now, any third party can file an opposition, and there is a uniform patent term of 20 years from the date of filing of the patent application.

Patent terms of pharmaceutical or agricultural inventions that are subject to pre-marketing administrative approval may be extended for up to an additional five years. However, there are certain limitations to obtaining such an extension in addition to qualifying subject matter. For instance, the request must be submitted within three months of regulatory approval of the patented invention, and no request may be filed within six months from the end of the patent term. The request must show that the patent could not be commercialized for at least two years after issuance due to delay necessitated by regulatory approval (from either the day the patent issues, or the day approval is first sought, whichever is *later*).

Scope of Protection. Infringement of a patent can occur through both direct and indirect means. However, the infringement must be commercial and public, unlike infringement under U.S. law. It should be noted that pursuant to TRIPS requirements, the rights of patent owners were recently extended to also include the ability to exclude unauthorized offers for sale of the patented invention. In particular, literal infringement occurs when an authorized party:

1. commercially makes, uses, sells, offers to sell, or imports a patented product;
2. commercially uses a patented process; or
3. commercially makes, uses sells, offers to sell or imports a product made by a patented process.

Under a recent revision to Japanese patent laws, unless an alleged infringer (the defendant in a patent infringement suit) shows proof to the contrary, if a defendant's product is identical to one obtained by the patented process, it is presumed to be manufactured by such process (107).

Infringement under the doctrine of equivalents (DOE) has been a relatively recent phenomenon in Japan. Although courts had applied the DOE sporadically in recent years, it was not until 1998 that the Supreme Court affirmatively embraced the doctrine and set forth clear standards for its application (108–109). The Supreme Court clarified that the DOE analysis should be considered *at the time of the infringement* (rather than at the time of filing as was previously held by lower courts) (108); this change is consistent with U.S. law and favors patentees and particularly pioneering patents that can now potentially cover related advances that develop after a patent issues. Although courts were initially hesitant to find infringement under DOE, at least one district court has done so as of 1999 (110).

Under the DOE law in Japan, infringement may be found even if the patent has an element that does not cover the accused product, if the accused product is regarded as being equivalent. Equivalence is determined based on a multifactor test. The essential inquiry in finding infringement via DOE is whether the accused product or process contains elements identical or equivalent to each

claimed element of the patented invention; again, this central inquiry echoes current U.S. law requiring that DOE be conducted on an element-by-element analysis. Infringement under DOE may be found where:

1. portions are not essential to the patentability of the patented invention;
2. objectives and effects of the patented invention are still attained after replacing the portions with their counterparts in the accused product;
3. the replacement would have been obvious to those skilled in the art at the time of manufacturing the accused product;
4. the accused product was neither anticipated nor obvious to those skilled in the art as of the filing date of the application for the patented invention; and
5. no special conditions exist, such as the accused product being intentionally excluded from the scope of claims during the prosecution of the application for the patented invention (no prosecution history estoppel).

There are several exceptions to the basic patent right to exclude others from commercial use of the invention that may affect biotechnological inventions. No patent rights are available with respect to products that existed in Japan prior to the filing of the patent application (106). Also no patent rights exist for acts of preparing medicines in accordance with medical prescriptions (106). The Japanese Patent Act has long since established that there is no right to exclude those who use a patented invention for "purposes of experiment or research" (105, Article 69(1)). However, this has been recently interpreted to cover the manufacture of patented products for pre-marketing approval by companies seeking to manufacture generic versions of patented drugs (110). This judicial interpretation brings Japanese patent law in line with practice under U.S. law. However, as noted with the similar provision in the EU, it is unclear whether this interpretation is consistent with TRIPS.

Compulsory licenses of patented inventions are also available in certain circumstances under Japanese law. In particular, one who has independently created a patented invention and has been commercially using it or made preparation to do so, can obtain a nonexclusive license (Article 79). JPO has noted that compulsory licenses are available in the cases of nonworking of a patent for at least three years (Article 83), dependent patents (Article 92), and public interest (Article 93). However, it is unclear whether the provision of such compulsory licenses are entirely compatible with TRIPS.

Enforcement Issues. Finally, damages are important in examining patent protection of biotechnology in Japan because damages enable a patent owner to effectively enforce rights and preclude infringement. An examination of damages under Japanese Patent law is particularly relevant because of recent changes that substantially improve, recovery for prevailing plaintiffs (i.e., patent owners). Previously only limited damage awards were available (at least in comparison to U.S. awards). Plaintiffs generally did not even attempt to seek lost profits as courts rarely granted them; the primary reason for this was that plaintiffs could not establish causation between infringement and damages because plaintiffs had no access to confidential information of the defendant and Japanese patent law did not provide for any inferential causation of lost profits. However, under the present law, which became effective as of January 1, 1999, the types of damages available have been explicitly broadened in the patent statute to include, for the first time, lost profits as a measure of patent infringement damages.

Under the new section, the plaintiff is relieved of the causation burden as the statute provides a presumptory amount that the defendant must then rebut. The presumptory amount of lost profits consists of the infringer's sales volume multiplied by the patentee's profit rate, as long as such amount does not exceed the amount a patentee would be able to obtain based on its own manufacturing capacity. The burden is on the infringer to establish that the presumed damages are incorrect; the infringer would need to show that the actual number of infringing products was lower.

In addition to providing for lost profits, the amended patent law also allows the potential for increased awards based on a reasonable royalty calculation. Previously damage awards were calculated based on industry-standard royalty rates and rates for licensed government-owned patents; this resulted in royalty rates ranging from 2 to 4 percent, which was in stark contrast to a more typical rate of 8 to 10 percent in Europe (112). This stringent calculation was in part derived from the fact that the word "normally" qualified the term "reasonable royalty." The amended law removes the qualification "normally," such that courts should be able to grant higher amounts than previously. The cap for damages available from an infringing corporation has been raised from 5 million yen ($36,500) to 150 million yen ($1.09 million) (106, Article 201).

There are additional provisions in the Japanese patent laws that may assist plaintiffs in patent infringement actions to recover more damages. In the past plaintiffs were forced to extrapolate sales of defendants infringing products based solely on public documents because there was no requirement that parties produce documents other than those that it intends to rely on at trial. However, under a new provision of the patent laws, if a party so requests, a court may order the opposing party to produce documentation necessary to determine damage caused by an infringement (barring some legitimate reasons for failing to produce such documents) (106, Article 105). Moreover a new provision of the Civil Procedure Code, provides judges with discretion to determine the appropriate amount of damages where "it is extremely difficult to prove the amount of damages from the nature of such damage" (114, Article 248). Although this provision was intended to assist in calculating intangible damages such as emotional distress, it has been relied upon to discount a defendant's calculation of lost profits and to instead adopt a calculation closer to the plaintiffs when the defendant's internal documentation was not made available (112,115).

In the future, it is possible that even more monetary compensation may be available for plaintiffs, which would

bring Japanese patent law closer to the situation in the United States In 1998 the JPO proposed to provide punitive damages and partial attorney fees for successful plaintiffs — both of which are available remedies to prevailing patent owners in the United States and serve as a deterrent effect against infringement. Although no such legislation was passed in 1998, JPO, as well as the Ministry of International Trade and Industry, continue to lobby for such changes.

BIBLIOGRAPHY

1. Agreement on Trade-Related Aspects of Intellectual Property Rights (April 15, 1994), Marakesh Agreement Establishing the World Trade Organization, Annex C, Legal Instruments — Results of the Uruguay Round, 33 I.L.M. 81, Available at: *http://www.wto.org*
2. International Convention for the Protection of New Varieties of Plants (1978 Act), Available at: *www.upov.int/eng/convntns*
3. International Convention for the Protection of New Varieties of Plants (1991 Act), Available at: *www.upov.int/eng/convntns*
4. Convention on the Grant of European Patents (October 5, 1973), 13 I.L.M. 270.
5. Directive 98/44/EC of the European Parliament and of the Council on the Legal Protection of Biotechnological Inventions (July 6, 1998), OJ (L 213) 13.
6. EPO Member States, Available at: *http://www.european-patent-office.org/epo/members.htm*
7. Information on the European Union, Available at: *http://www.cali.co.uk/bis/europe.htm*
8. About the WTO, Available at: *http://www.wto.org/wto/about/organsn.htm*
9. States Party to the International Convention, Available at: *http://www.upov.int/eng/ratif/index.htm*
10. Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure (April 28, 1977), 32 U.S.T. 124, Available at: *http://www.wipo.int/eng/iplex/index.htm*
11. Status of Member States to Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure, Available at: *http://www.wipo.org/eng/ratific/q-budpst.htm*
12. Europa, Key Issues: Enlarging the EU, Available at: *http://europea.euint/geninfo/key-en.ht*
13. Paris Convention for the Protection of Industrial Property (March 20, 1883), *as last revised*, Stockholm, July 14, 1967, arts. 1-12, 21 U.S.T. 1583, 828 U.N.T.S. 305, Available at: *http://www.wipo.int/eng/iplex/index.htm*
14. Understanding On Rules and Procedures Governing the Settlement of Disputes (April 15, 1994), WTO Agreement, Annex 2, 33 I.L.M. 1226.
15. Overview of the State-of-Play of WTO Disputes (November 26, 1999), Available at: *http://www.wto.org/wto/dispute/bulletin.html*
16. E.-U. Petersmann, in E.-U. Petersmann, ed., *International Trade Law and the GATT/WTO Dispute Settlement System*, Kluwer Law International, London, 1998.
17. T.P. Stewart and M.M. Burr, *N.C. J. Int. Law Com. Reg.* **23**, 431 (1998).
18. J.H. Reichman, *Virginia J. Int. Law* **37**, 334, 337 (1997).
19. R.C. Dreyfuss and A.F. Lowenfeld, *Virginia J. Int. Law* **37**, 275 (1997).
20. India, Patent Protection for Pharmaceutical and Agricultural Chemical Products (December 19, 1997), AB 1997-5, W.T./DS50/AB/R.
21. C. MacDonald-Brown and L. Ferera, *Eur. Intell. Prop. Rev.* **2**, 69 (1998).
22. C. Correa, in C. Correa and A. Yusuf, eds., *Intellectual Property and International Trade: The TRIPS Agreement*, Kluwer Law International, Brentford, UK, 1998.
23. World Intellectual Property Organization, *Existence, Scope and Form of Generally Internationally Accepted and Applied Standards/Norms For the Protection of Intellectual Property*, Doc. No. MTN.GNG/NG11/W/24/Rev.1, Annex II (September 1988).
24. D. Gervais, *The TRIPS Agreement: Drafting History and Analysis*, Sweet & Maxwell, London, 1998.
25. C. Correa, *Eur. Intell. Prop. Rev.* **8**, 435, 437–438, 441 (1997).
26. C.M. Correa, *J. World. Intell. Prop.* **1**, 75–98 (1998).
27. E.S. Van de Graf, *Patent Law and Modern Biotechnology*, Sanders, Rotterdam, 1997.
28. T820/92, *Contraceptive Method*, 1995 Off. J. Eur. Pat. Off. 113; 26 Int. Rev. Indus. Prop. & Copy. L. 543 (1995).
29. 290/86 (*Cleaning Plaque/ICI*), 1992 Off. J. Eur. Pat. Off. 414; 23 Int. Rev. Indus. Prop. & Copy. L 815 (1991).
30. T 144/83 (*Appetite Suppressant*), 1986 Off. J. Eur. Pat. Off. 301; 18 Int. Rev. Indus. Prop. & Copy. L 258 (1987).
31. *T 74/93, Contraceptive Method/British Technology Group*, 1995 Off. J. Eur. Pat. Off. 712, *reprinted in* 27 Int. Rev. Indus. Prop. & Copy. L. 99 (1996).
32. R. Moufang, *Methods of Medical Treatment under Patent Law*, 24 Int. Rev. Indus. Prop. & Copy. L. 18-49 (1993).
33. P. Cullet, *J. World Intell. Prop.* **2**, 617 (1999).
34. S. Charnovitz, *Virginia J. Int. Law* **38**(6), 639–744 (1998).
35. C.T. Fedderson, *Minn. J. Global Trade* **7**, 75–122 (1998).
36. *Analytical Index of the GATT*, pp. 562–596 (1994).
37. D. Palmeter and P. Mavroidis, *Am. J. Int. Law* **92**, 398 (1996).
38. T. Ackerman, *Tex. Int. Law J.* **32**, 489–510 (1997).
39. Organisation for Economic Cooperation and Development, *Intellectual Property Practices in the field of Biotechnology*, TD/TC/WP(88) 15/final (1999), Available at: *www.oecd.org*
40. Organisation for Economic Cooperation and Development, *Responses from Member Countries to Working Part of the Trade Committee of the OECD Questionnaire on Intellectual Property Practices in the Field of Biotechnology* (1999), Available at: *www.oecd.org*
41. Responses from EPO and JPO, Available at: *http://appli1.oecd.org/ech/BiotechIPR.nsf*
42. G.H.C. Bodenhausen, *Guide to the Application of the Paris Convention for the Protection of Industrial Property as Revised at Stockholm in 1967*, United International Bureau for the Protection of Intellectual Property, Geneva, 1968.
43. Argentina Patent Law art. 44.
44. *Cong. Rec.* **142**, S11843–11844 (September 26, 1996).
45. American Bar Association, Section of Intellectual Property Law, *Annual Report 1996–1997*, ABA Sect. I.P. Law Rep. 70, 103 (1997).
46. C. Ho, *U.C. Davis Law Rev.* **3** (2000).
47. WTO, Canada-Patent Protection of Pharmaceutical Products Report of the Panel, WT/DS114/R (March 17, 2000).

48. Canada Department of Foreign Affairs and International Trade, *Trade Negotiations and Agreements — Panel Cases to which Canada is a Party* (1999), Available at: *http://www.dfait-maeciigc.ca/tna-nac/summary-e.as*
49. Draft of July 23, 1990 (Doc. MTN.GNG/NG11/W/76), reprinted in D. Gervais, *The TRIPS Agreement: Drafting History and Analysis*, Sweet & Maxwell, London, 1998.
50. M. Halewood, *Osgoode Hall Law J.* **35**, 243, 264–265 (1997).
51. R. Weissman, *Univ. Ponn. J. Int. Econ. Law* **17**(4), 1069–1125 (1996).
52. Brazilian Industrial Property Law (May 14, 1996).
53. J. Giust, *Hastings Int. Comp. Law Rev.* **21**, 597–637 (1998).
54. I. Ahler, *Int. Rev. Indus. Prop. Copy. Law.* **28**(5), 632–671 (1997).
55. E. Duran and C. Michalopoulos, *J. World Intell. Prop.* **2**, 865–873 (1999).
56. J. Watal, *J. World Intell. Prop.* 281–307 (1999).
57. International Convention for the Protection of New Varieties of Plants (1961 Act), Available at: *www.upov.int/eng/convntns*
58. S.K. Verma, *Eur. Intell. Prop. Rev.* **6**, 281–289 (1995).
59. B. Greengrass, *Eur. Intell. Prop. Rev.* **12**, 466–472 (1991).
60. EPO Guidelines for Substantive Examination, Available at: *www.european-patent-office.org/guidelines/english/*
61. Decision of the Administrative Council Amending the Implementing Regulations to the European Patent Convention, 1999 Off. J. Eur. Pat. Off. 43; 29 Int. Rev. Indus. Prop. & Copy. L. (1998).
62. *Netherlands v. Parliament*, Case C-377-78, 1998 O.J. C 378/13.
63. A. Scott, *Eur. Intell. Prop. Rev.* **4**, 212–215 (1999).
64. I. Muir, M. Brandi-Dohrn, and S. Gruber, *European Patent Law: Law and Procedure under the EPC and PCT*, Oxford University Press, Oxford, UK, 1999.
65. G. Van Overwalle, *The Legal Protection of Biotechnological Inventions in Europe and in the United States: Current Framework and Future Developments, Technical and Ethical Approaches*, Leuven University Press, Leuven, Belgium, 1997.
66. K. Goldbach, H. Vogelsang-Wenke, and F. Zimmer, *Protection of Biotechnological Matter under European and German Law: A Handbook for Applicants*, VCH Publishers, New York, 1997.
67. T 385/86 (Noninvasive measurement/Bruker), 1988 Off. J. Eur. Pat. Off. 308 (September 25, 1987); 1988 Eur. Pat. Off. Rep. 357.
68. T 143/94, *Trigonelline*, 1996 Off. J. Eur. Pat. Off. 430, reprinted in 28 Int. Rev. Indus. Prop. & Copy. L. 95 (1997).
69. T 49/83, Propagating material/Ciba Geigy, 1984 Off. J. Eur. Pat. Off. 112; [1979-85] Eur. Pat. Off. Rep. vol. C. 758.
70. T 320/87, Hybrid plants/Lubrizol, 1990 Off. J. Eur. Pat. Off. 71.
71. T 356/93, Plant Cells/ Plant Genetics Systems, 1995 Off. J. Eur. Pat. Off. 545.
72. R. Nott, *Eur. Intell. Prop. Rev.* **21**, 33 (1999).
73. I. Voelker, *Europe Won't Reverse Controversial EPO Ruling*, IP Worldwide, July/August 1997, Available at: *http://www.ljx.com/patents/7-8europe.html*
74. T 1054/96, In re Novartis, 1998 Off. J. Eur. Pat. Off. 511 (Tech. Bd. Appeal 1997); 30 Int. Rev. Industrial Property and Copyright Law 78 (Tech. Bd. App. 1997).
75. G. Van Overwalle, *J. L. & Tech.* **39**(2), 143–194 (1999).
76. G 1/98, Novartis, 2000 Off. J. Eur. Pat. Off. 111–171 (Enlarged Board of Appeal, December 20, 1999), Available at: *http://www.european-patent-office.org/epo/pubs/oj000/3_00/index_e.htm*
77. Relaxin, 1995 Off. J. Eur. Pat Off. 388 (Exam. & Opp. Div. 1994).
78. *Greenpeace UK v. Plant Genetic Systems NV*, 24 Int. Rev. Indus. Prop. & Copy. L. 618–625 (Opp. Div. 1992).
79. H.R. Jaenichen and A. Schrell, *Eur. Intell. Prop. Rev.* **9**, 345 (1993).
80. Harvard/Onco-mouse, 1990 Eur. Pat. Off. Rep. 501 (Tech. Bd. App.).
81. Harvard/OncoMouse, 1991 Eur. Pat. Off. Rep. 525 (Examining Division).
82. I. Karet, *Eur. Intell. Prop. Rev.* **1**, 21–26 (1997).
83. M. Spence, *Law Q. Rev.* **113**, 368–374 (1997).
84. *Biogen v. Medeva*, 1977 R.P.C. 1 (House of Lords).
85. *Biogen v. Medeva*, 1995 R.P.C. 25 (Patent Court).
86. *Biogen v. Medeva*, 1995 R.P.C. 68 (Ct of Appeals); 27 Int. Rev. Indus. Prop. & Copy. L. 856 (1996).
87. R. Crespi, *Int. Rev. Indus. Prop. Copy. Law* **28**(5), 603–622 (1997).
88. Community Patent Convention, Art. 27.
89. ICI/Medicopharma, Sup. Ct, December 18, 1992, The Netherlands.
90. Ars/Organon, 1997 Int. Rev. Indus. Prop. & Copy. L 702 (Ct of Appeal 1994), affd 1998 Int. Rev. Indus. Prop. & Copy. L 702 (Sup Ct 1996).
91. Clinical Trials II (kirin Amgen/Boehringer Mannheim), 1998 RPC 423 (April 1997).
92. W. Meibom and J. Pitz, *Eur. Intellec. Prop. Rev.* **8**, 469–478 (1997).
93. G. Gauci, *Eur. Intellec. Prop. Rev.* **9**, 361–362 (1998).
94. J. Thomas, in K. Hill and T. Takenaka, eds., *International Perspectives on Intellectual Property: The Doctrine of Equivalents, Enforcement in East Asia and Other Issues*, CASRIP, Seattle, WA, 1997, pp. 189–233.
95. Treaty of Brussels par. 6(1).
96. *EGP v. Boston Scientific* (April 24, 1998).
97. P. Coletti, *J. Pat. Off. Soc.* **81**, 351 (1999).
98. Council Regulation No. 1768/92 of 18 June 1992, Concerning the Creation of A Supplementary Protection Certificate for Medicinal Products, 1992 OJ (L 182) 1-5.
99. H. Von Morze and P. Hanna, *J. Pat. Off. Soc.* **77**, 479 (1995).
100. Council Regulation No. 2100/94 of 27 July 1994 on Community Plant Variety Rights, 1994 OJ (L 227) 1-30.
101. P.A.C.E. van der Kooj, *Introduction to the EC Regulation on Plant Variety Protection*, Kluwer Law International Brentford, UK, 1997.
102. Council Regulation No. 1610/96 of the European Parliament and the Council on the Creation of a Supplemental Protection Certificate for Plant Protection Products, (July 23, 1996), OJ (L 198) 3-35.
103. EU Green Paper on Patents, IP/97/558; Memo 97/65 (June 1997).
104. Trilateral Project B3, Available at: *http//www.european-patent-office.org/tsw/sr-3-b3b.html*
105. Implementing Guidelines for Inventions in Specific Fields: Biological Inventions (Ch. 2), Available at: *http://222.*

jpo-miti.go.jp/guidee/sisine.html (published February 27, 1997).

106. Japan Patent Law reprinted by J.P. Sinnott, *World Patent Law and Practice*, Matthew Bender, New York, 1968.
107. *SmithKline Beecham v. Fujimoto*, 30 Int. Rev. Indus. Prop. & Copy. L. 457 (Tokyo Dist. Ct. October 12, 1998).
108. THK v. Tsubakimoto ("Ball-Spine case"), No. 3292, March 29, 1996.
109. S. Yamamoto and J.A. Tessensohn, *J. Pat. Off. Soc.* **81**, 483 (1999).
110. *Pharmacia & Upjohn Aktiebolag v. Nippon Eli Lilly* (Osaka District Ct., May 27, 1999), Available at: *http://courtdomino.courts.go.jp/chiteki.nsf/caa027de696a3bd349256795007b825/7da56fc63308e18a492567a20013a5c8? Open document*.
111. Clinical Trial III, 30 Int. Rev. Indus. Prop. & Copy. L. 448 (Japan Sup. Ct. 1999).
112. R. Gunther and D. Yoshida, *Pro-Patent Trend Signaled in Japan, N.Y. Law J.*, S3 (July 26, 1999), citing Kogyo Shoyuken Shingikai [Industrial Property Advisory Committee (IPAC)], Tokkyoho nado no kaisei ni kansuru toshin [Report on Reforms to the Patent Law] (December 16, 1997) [1997 IPAC Report].
113. A. Hirai, *IP Worldwide* 8 (July/August 1998).
114. Japan Civil Procedure Code 220.
115. K. Hagberg, E. Doi, and J. Ishiguro, *Nat. Law J.* C7, (September 14, 1998).

See other INTERNATIONAL ASPECTS entries; MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT; OWNERSHIP OF HUMAN BIOLOGICAL MATERIAL; see also PATENTS AND LICENSING entries; SCIENTIFIC RESEARCH, POLICY, TAX TREATMENT OF RESEARCH AND DEVELOPMENT.

M

MEDIA COVERAGE OF BIOTECHNOLOGY

DOROTHY NELKIN
New York University
New York, New York

OUTLINE

INTRODUCTION

In 1924 Edwin E. Slosson, editor of the first science writing syndicate in America, described his view of science journalism. "The public that we are trying to reach is in the cultural stage when three-headed cows, Siamese twins and bearded ladies draw the crowds to the side shows." That is why, he explained, science is usually reported in short paragraphs ending in "-est." "The fastest or the slowest, the hottest or the coldest, the biggest or the smallest, and in any case, the newest thing in the world" (1).

In some respects little has changed in the world of science reporting. In the 1990s journalists still play up the hottest scientific discoveries, the riskiest technologies, and the latest miracle cures. The media coverage of rapid advances in biotechnology and especially in human genetics illustrates these tendencies: extravagant reports about wonder therapies, genetically engineered pigs, and cloned sheep attract readers and sell magazines. "The gene of the week" is "hot" news even though the significance of particular discoveries has often been questioned. The controversial theories emerging from research in behavioral genetics that purport to explain human behavior and social differences are especially newsworthy though their validity is often in doubt.

The media reporting on biotechnology and genetics is mainly promotional, directly reflecting the influence of scientific sources of information. But it is also polarized, swinging from enthusiastic promises to warnings of peril. Journalists are especially attracted to controversies, and new developments in biotechnology have created many provocative and disputed issues. Biotechnology products, intended to control pests or plant diseases and to increase agricultural yields have been controversial because of potential risks to environmental quality or human health. The bioengineering of animals and the creation of new transgenic foods has raised moral and aesthetic objections as well as concerns about risk.

Media coverage of science and technology provides a useful window on public attitudes, but science reporting also has an important influence on public perceptions. The way people perceive research in human genetics or developments in agricultural biotechnology — the way they interpret their costs and benefits — may be influenced less by the details of scientific evidence than by the repeated messages conveyed in the popular press. These media messages help to create the beliefs and assumptions that underlie personal decisions, social policies, and institutional practices.

Following a brief overview of the history of the media coverage of biotechnology, this article will illustrate its most important characteristics by describing five areas that have received considerable media attention. The first widely reported development in biotechnology was the discovery and synthesis of interferon. The media's presentation of news about this therapeutic product was volatile and polarized, a style that has since characterized many reports about biotechnology events. The coverage of research on gene therapy suggests the important influence of scientific sources in shaping the content and tone of science news. The coverage of cloning demonstrates the appeal of drama, myth, and image to the journalists reporting on science events. The media attraction to behavioral genetics suggests the appeal of scientific theories that conform to popular social stereotypes or support prevailing policy agendas. Finally, reports on biotechnology risks, and the problems that may follow from genetic predictions illustrate the media's attraction to policy disputes. A persistent theme pervades this media coverage of biotechnology issues — a concern about the social implications of the growing ties between science and commercial interests.

HISTORICAL OVERVIEW

"After years of being a dowdy old lady, biology has become belle of the ball." Its revolutionary potential has attracted researchers "in droves," and "bankers [are] in hot pursuit." To the media in the early 1980s, biotechnology was expected to become the next economic miracle. But, ironically, only a few years earlier the reports on biotechnology had been more about risks than revolutions. In the mid-1970s molecular biologists held an international meeting at the Asilomar conference center in California to assess the potential risks of recombinant DNA research (2). This was mainly a technical discussion, but the press evoked images of Frankenstein monsters and Andromedalike strains spreading like an incurable disease. Some reporters worried about "warping the genetic endowment of the human race"; others about "biological holocaust." The message? Runaway science needs to be controlled (3).

Yet, only a few years later, questions of safety ceased to be news. Journalists dropped the subject, turning their attention to promises and applications. The discovery of ways to synthesize interferon brought reports of miracle cures. Techniques of gene splicing, once represented as dangerous, became "a mundane tool," news and headlines began to tout the potential applications of biotechnology research as "miracles," as the key to better health. In the space of only a few years, media attitudes had markedly changed. In 1977, for example, *Time* magazine had run a cover story called "The DNA Furor: Tinkering with Life" Only three years later a *Time* cover story was called "DNA: New Miracle." Similarly in 1976 the *New York Times* magazine section published an article called "New Strains of Life or Death?" Later a 1980 article in the same section of the *New York Times* was called "Gene Splicing: The Race towards Better Human Health" (4).

By the early 1990s the "runaway science of genetic engineering" had become a "technological frontier." Stories welcomed the patenting of genetically engineered products, the implications of these products for resolving medical, agricultural, and industrial problems, and the proliferation of genetics R&D firms. Local papers touted the importance of these firms to the regional economy. Journalists described geneticists as pioneers, "unlocking the basic laws of nature," discovering the "secrets of life," solving the problems of devastating disease. The biotechnologists working on "high tech veggies" will "do wonders" to help meet nutritional requirements and to enhance the economy. Genomic researchers are "riding the DNA trail." Geneticists are "relentless hunters of genes," involved in a "race" to find the markers for disease. Media accounts are reverent — almost religious: A 1994 cover of *Time* depicts a figure on a pedestal, his arms extended in a Christlike pose, his torso inscribed with a double helix. The caption reads: "Genetics — The Future Is Now." The image, of course, is the Ascension.

Encouraged by the enthusiastic response to research in medical genetics, behavioral psychologists in the 1990s began to publish and to publicize their long controversial studies on the genetic basis of behavioral conditions and personality traits. Always seeking provocative copy, journalists reported these claims, focusing especially on the most controversial — those concerning predisposition to antisocial behavior. And, often uncritically, they drew conclusions about the implications of this research for social policy (5).

Despite the general media enthusiasm about biotechnology, media reports are frequently tempered by doubts. Many journalists, for example, have been critical of the growing links between the biotechnology industry and universities for their affect on open research and they have called attention to the conflicts of interest that are endemic to this field where profits and ethics collide. The media have amplified the critical views of activist Jeremy Rifkin, a persistent and media-savvy biotechnology gadfly. They have extensively reported the protests of animal rights groups against the creation of transgenic animals, and the concerns of religious groups about genetic engineering. In response to the rapid development of genetic tests, media attention turned to the issue of genetic discrimination as tests reveal information about individual predispositions that could influence access to insurance or to jobs (6). And the old Frankenstein metaphors have reappeared in reports about genetically modified foods.

Reporting on biotechnology — from the development of interferon to the creation of clones — the press has batted readers back and forth from biotechnology miracles to visions of apocalypse, from celebrations of progress to warnings of peril, from optimism to doubt. These dramatic shifts in media reports take place as journalists respond to the promotional hyperbole of scientists and technical institutions, but also to protestors and changing popular fashions. This fickle and volatile style of reporting serves the interests of the media in their endless quest for newsworthy and dramatic material. The discovery and synthesis of interferon provided the first opportunity for this style of science journalism.

INTERFERON

Interferon, a protein manufactured in the body when a virus invades a cell, was discovered in 1953 as a natural therapeutic agent, a so-called interfering protein that inhibits infection. The possibility of isolating the protein raised hopes for eventually developing a cancer cure, and this caught media attention. However, the scarce supply of the agent at that time limited scientific progress and clinical possibilities and journalists lost interest until 1975 when Mathilde Krim, a politically astute geneticist, organized a conference intended to publicize the potential of interferon and to win public support for research (7).

Krim's persistent efforts and the growing interest of the American Cancer Society (ACS), which began to sponsor clinical trials, brought a deluge of media coverage in the late 1970s. The research articles in scientific journals explicitly qualified the promises of interferon, indicating the tentative nature of existing studies, the high cost of isolating the protein, and its therapeutic limits. In the popular press, however, interferon became a "magic bullet," a miracle cure for everything from cancer to the common cold.

In 1980 Biogen, a new biotechnology firm, developed a DNA clone for this protein, opening the possibility of producing large quantities of the product at low cost. Uncritically accepting promotional information provided by the company at a press conference, journalists welcomed this new technological development as still another miracle. "Like the genie in a fairy tale," the *Detroit Free Press* told its readers, "science came up with the key to the magic potion." *Reader's Digest* wrote about a "wonder therapy," *Newsweek* about "cancer weapons" and "the making of a miracle drug." Business journalists focused on interferon as a profitable commodity, calling attention to the dramatic increase in stock prices of biotechnology firms. *Business Week* described the efforts to synthesize the substance as a "race" to capture the market: "We have just passed the quarter mile pole and all the horses are in a bunch." *Time* wrote described interferon as a "gold mine" for patients and for biotechnology firms.

The *New York Times* science writer Harold Schmeck, however, broke this promotional pattern. Writing cautious

reports, he suggested that while interferon was promising, there was no definitive evidence of its effectiveness. He also reported on possible harmful effects, suggesting that "the seemingly ideal weapon" was less of a panacea than anticipated, and he reported on interferon studies that "put cancer use in doubt" (8). Emphasizing the "modest, controversial, and even negative results of research," Schmeck observed that the promise of a scientific advance was raising research money, but also raising false hopes. In response to this article four scientists from the Sloan Kettering Institute for Cancer Research wrote a letter to the *New York Times*, expressing concern that such qualified reporting could undermine public support of interferon research (9).

The difficulties of using interferon as a therapeutic agent became public in 1982 when four patients in France died after they were treated with the drug. Abruptly, the tone of reporting changed from exaggerated optimism to disillusionment: "From wonder drug to wallflower." The wonder drug was demoted from a magic bullet to a "research tool." Newspapers and magazine articles assessed the situation pessimistically: "Jury's out on interferon as a cancer cure"; "Studies cast doubt on cancer drug"; "It's a hard row to hoe." Research continued, but little more appeared in the press until a series of patent disputes turned media attention to the question of proprietary interests in commercially promising biotechnology products.

These reports on interferon research demonstrate several themes that have since characterized media reports on biotechnology. First, the content is limited. Little appeared in the press coverage of interferon about the actual nature of the research; instead, most articles appealed to public concerns about cancer and the hopes of soon finding a cure for this dread disease. While interferon's short-term usefulness as a therapeutic agent was problematic, the research did yield significant scientific understanding of basic biological concepts (e.g., the control of gene expression in mammalian cells and the regulation of immunity) that in the long term have affected the practice of medicine. But those readers who followed the interferon story learned little about such developments.

Second, the media coverage of interferon placed great emphasis on scientific and technological competition. Scientists and the firms developing interferon were in a "race" for breakthroughs, for solutions to a dread disease. The gradual accumulation of information that is inherent to the research process was not considered news. Whether the goal is to discover a new genetic marker or to clone a sheep, the media are attracted to the competition in science, the race to be the first to get results.

The media reports about interferon were also volatile. Readers were mainly treated to hyperbole — to a promotional coverage designed to raise their expectations and whet their interest. Scientists played an important role in shaping this coverage. Far from being neutral sources of information, they actively sought a favorable press, equating public interest with research support, and journalists for the most part were inclined to accept the claims of scientific sources. However, when predictions about interferon's curative powers failed to materialize, unqualified optimism in the press quickly shifted to the opposite extreme. This pattern of premature promotion followed by a bitter backlash when optimistic promises fail has continued to be a striking characteristic of the media coverage of gene therapy.

PROMISES OF GENE THERAPY

The media have welcomed claims about gene therapy with extravagant headlines and promotional hype as reporters convey — often uncritically — an array of futuristic scenarios presented by enthusiastic scientists. In the future, said a geneticist to *Discover*, "present methods of treating depression will seem as crude as former pneumonia treatments seem now" (10). In the future, said another scientist, food companies will sell infirmity related breakfast cereals targeted to aid those with genetic predispositions to particular diseases. "Computer models in the home will provide consumers with a diet customized to fit their genetic individuality, which will have been predetermined by simple diagnostic tests" (11). In the future, geneticist French Anderson told a *Time* Reporter, "Physicians will simply treat patients by injecting a snippet of DNA and send them home cured" (12).

To the media, the National Institutes of Health (NIH) Human Genome Initiative has replaced the NASA Space program as a new frontier, the cutting edge of high-technology exploration, and just as in the heyday of NASA, journalists are receptive to scientific enthusiasm. On their part, scientists seek media coverage as a form of public relations, regarding public visibility as a means to attract funds for their research. Thus scientists and the press offices of their institutions are inclined to turn precliminary experimental findings into magic bullets. For example, in 1995 the Scripps Research Institute issued a press release announcing that researchers had found a cure for cancer through a small injection of a protein that would cut off the blood supply from tumors and cause them to shrink while leaving normal tissue intact. The announcement about a cancer cure, designed to attract media attention, was, it turned out, only about a laboratory observation; there had been no experimental trials testing the relevance of the observation to human pathology. But as anticipated, the press release was covered as welcome news in the press (13).

Traditionally working in a context where success is measured by the judgment of peers, scientists have long assumed that a record of accomplishment is sufficient to maintain research support. Thus information, the scientist's "stock-in-trade," has been directed primarily toward professional colleagues. Most scientists have not been interested in public visibility; on the contrary, they have feared it could result in external controls on their work. And "visible scientists" — those who seek media attention — have often been marginalized or disdained (14). But attitudes in the scientific community have changed. Dependent on direct congressional appropriations or, increasingly, on corporate support of research, many scientists now believe that scholarly communication is no longer sufficient to maintain their enterprise. They

see gaining national visibility through the mass media as crucial to securing the financial support required to run major research facilities and to assuring favorable public policies toward science and technology.

Geneticists seeking to maintain public support have become skilled in rhetorical strategies designed to attract the media. In media interviews they have described the genes as "master molecules:" We are but "readouts" of our genes. They describe the body in deterministic terms — as a set of instructions, a blueprint, a map, or a program that is transmitted from one generation to another. They suggest that by deciphering the text, classifying the markers on the map, and reading the instructions, they will unlock the key to human ailments and human nature, revealing the secrets of human life. Molecular biologist Walter Gilbert, introduces his public lectures on gene sequencing by pulling a compact disk from his pocket and announcing to his audience: "This is you." Scientists, he claims, will "provide ultimate answers to the commandment, know thyself" (15).

Geneticists also emphasize the predictive powers of their science by calling the gene "a Delphic oracle," "a time machine," "a trip into the future," "a medical crystal ball." James Watson, the first director of the Human Genome Project, has announced in frequent media interviews that "our fate is in our genes." The metaphors scientists use to describe their work convey several messages about the meaning of genetics that have been widely disseminated through a receptive press: a definition of the gene as an essentialist and deterministic entity, and a promise that genetic research will enable the prediction of future behavior and disease and thus lead to therapeutic solutions.

The biotechnology industry has further encouraged media hype about gene therapy. The industry has made a major financial commitment to gene therapy, expecting this will be the basis of future medicine. Over 60 percent of gene therapy studies are directly financed by industry in anticipation of a profitable market in the near future. Corporate advertisements announce "a great leap in the treatment of disease," and promise "a healthy future one gene at a time." Immediately following the discovery of the mutation in the BRCA-1 breast cancer gene, one pharmaceutical company announced in a newspaper advertisement that they had made progress in finding a "breakthrough . . . new treatments and ultimately the cure for breast cancer." Another advertisement appearing in sports magazines, said it all: "'Bad Genetics?' Use Optigenetics — 'The first genetic optimizer.'"

The development of pharmaceutical products and the proliferation of clinical trials on new therapeutic procedures have encouraged the tendency towards technological optimism in the press. For example, the introduction of therapeutic molecules, especially TPA (tissue plasminogen activator) for dissolving clots, and the use of human growth hormone to treat dwarfism became newsworthy issues. Then the first FDA-approved gene therapy experiment in 1990 — the injection of cells containing ADA genes in a child with an immune system malfunction — became a major news event. "The long awaited era of genetic therapy has at last arrived" said a writer in *The Sciences* (16). *Discover* called gene therapy, "The Ultimate Medicine." Writing on gene-transfer techniques, the reporter proclaimed, "Genetic surgeons can now go into your cell and fix those genes with an unlikely scalpel: a virus." Interviewed for this article, molecular biologist Richard Mulligan declared that "We can use gene transfer to make a cell do whatever we want. . . .We can play God in that cell." Similarly *US News* told its readers that gene therapy is the medicine of the future. "No disease has given up more of its secrets to genetic sleuths than cancer." Genetics, promised the writer, will allow doctors to "do something" about the disease. The isolation of the colon cancer gene in 1993 prompted an enthusiastic scientist to tell a *New York Times* reporter of its implications: "Deaths are entirely preventable" (17).

When they report on complex scientific issues, journalists rely heavily on press releases, often responding with uncritical enthusiasm to promotional hype. "Genetic Research Leaves Doctors Hopeful for Cures," "New Hope for Victims of Disease," "Genetics, the war on aging . . . [is] the medical story of the century . . . Genetic technologies will dramatically curtail heart disease, aging, and much more" (18). In a story called "the Age of Genes," *US News* reported that "advances bring closer the day when parents can endow children not only with health but also with genes for height, good balance, or lofty intelligence" (19).

While promises, backed by scientific authority, raise hopes of instant cures, aside from controlling reproduction there is little that can be done to cure genetic disease. The gap between identifying a predisposition to a genetic condition and finding a therapeutic solution is very wide. The problems of clinical application follow in part from the complexity of genetic diseases. Some are caused by the absence of the activity of a gene product, others by altered proteins that disrupt cellular function, and still others by alterations of chromosomal structure. The early expectations about a successful therapy for cystic fibrosis, reported widely in the press, were confounded by the fact that this disease has many more mutations than originally anticipated.

There are also problems in finding safe vectors capable of transporting genes into targeted cells. Most gene therapies use viruses as the carrier mechanism, for this is the most feasible way of targeting appropriate cells. But there are risks. Inserting a gene in the wrong place along a strand of DNA could cause an undesirable effect. And the immune system may attack cells treated with gene therapy, responding to them as foreign or infected. Gene transfer experiments on monkeys were found to cause malignant T-cell Lymphoma. Clinical trials of Genentech's promising drug called Pulmozyme, developed to treat cystic fibrosis, were halted when they found significant mortality rates among treated patients.

Discovering risks and side effects is, of course, the purpose of animal research and clinical trials. But in an area hyped by both scientists and the media as "the medicine of the future," failures become more than routine science; they also become a newspeg for journalists who seldom convey to their readers the fact that a failed clinical trial can itself yield useful information. The death of a gene therapy patient in 1999 brought an abrupt end to the gene therapy hype.

The growing realization of the practical difficulties of extending laboratory studies to clinical applications is puncturing inflated expectations, and this is reflected in skeptical media reports. Describing the research on the unusual frequency of a mutation in the BRCA1 gene among Ashkenazy Jewish women, reporters commented again and again on the absence of effective therapies. "Does it make sense to screen healthy women for the defect given that there is no good therapy to offer those in whom it is found?"

Some media reports associate experiments in gene therapy with genetic engineering or "tampering" with genes. "Lurking behind every genetic dream come true is a possible Brave New World nightmare," says a *Time* reporter. "To unlock the secrets hidden in the chromosomes is to open up the question of who should play God with man's genes." An accompanying image portrayed scientists balancing on a tightrope of coiled DNA (20). And an illustration for a *New York Times* article on gene therapy and the potential of genetic engineering featured a drawing imitative of the famous Edvard Munch painting, "The Scream." A figure stands, horrified, mouth ajar, eyes wide open, her hair a mass of coiled DNA (21). Such images proliferated in the reports about the creation of Dolly, the first cloned sheep.

CLONING

This section "Cloning" has been adapted from Ref. 22. The consequences of cloning have long captured the popular imagination: cloning has been a major theme in horror novels and science fiction films. Genetic engineering research that has been associated with cloning has met a critical press. For example, in 1993 scientists at a George Washington University laboratory conducted a genetic engineering experiment that "twinned" a nonviable human embryo. The purpose was find a way to create additional embryos for in vitro fertilization, but major newspapers, popular magazines and talk shows covered the experiment as if it had actually yielded a cloning technology for the mass production of human beings. The media response was remarkable and diverse. The *Los Angeles Times* announced the glorious news that "infertility, virginity and menopause are no longer bars to pregnancy" (23). But also envisioned were embryo and selective breeding factories, cloning on consumer demand, breeding of children as organ donors, a cloning industry for selling multiples of human beings, and even a "freezer section of the biomarket" (24). *Time*, wrote of the "Brave New World of cookie cutter humans" (25). And repeatedly, scientists in media reports were accused of "playing God."

Then, in February 1997, scientists at the Roslin Institute in Edinburgh cloned Dolly, after 276 attempts, from the genetic material of a six-year-old sheep. The media response to the production of a sheep by cloning a cultured cell line reflected futuristic fantasies and Frankenstein fears about science and especially about genetic engineering. The meaning of Dolly that was conveyed by the media reflected a pervasive assumption of "genetic essentialism" — that human beings in all their complexity are simply readouts of a molecular text, that human identity is contained entirely in the sequences of DNA in the human genome (26). Thus, speculated journalists, why not clone great athletes like Michael Jordan, or great scientists like Albert Einstein, or popular politicians like Tony Blair, or less popular politicians like Newt Gingrich, or wealthy entrepreneurs like Bill Gates. Some reporters lauded cloning as a way to assure immortality. Again and again, media stories predicted that cloning will allow the resurrection of the dead (e.g., bereaved parents might clone a beloved deceased child). Or the technology could provide life everlasting for the deserving (narcissists could arrange to have themselves cloned).

But there were also anxious scenarios developed in the press, including futuristic stories about making new Frankenstein monsters, or creating Adolph Hitler clones, or producing "organ donors" only to harvest their (fully compatible) viscera (27). While cloning could theoretically make both sexes irrelevant to reproduction, the technology appeared as a threat to the male of the species — men would no longer be necessary! It also held a promise of creating perfect cows, sheep, and chickens, or perhaps even perfect people. If sperm banks (as portrayed in some women's magazines) were a place to "shop for Mr. Good genes," why not, asked reporters, use cloning to produce and reproduce perfect babies?

Journalists elicited views from people in various professions about the implications of cloning for their fields. A divorce lawyer predicted the doubling of his business. Historians wondered if the founding fathers could be cloned for display in a "living history" exhibit in a theme park: They suggested that the park might be called "Clonial Williamsburg." Some facetious policy commentators announced that cloning experiments could be developed to solve social problems: The race problem could be resolved by manipulating the balance between melanin and IQ genes. The age-old nature-nurture dispute could be definitively settled by creating clones and raising them systematically in different environments.

News articles covered religious perspectives on cloning (28). One writer quipped that cloning offered a "second chance for the soul." If you sin the first time, try again. But a theologian, Rabbi Mosher Tendler, a professor of medical ethics at Yeshiva University in New York City, warned in a news interview that "whenever man has shown mastery over man, it has always meant the enslavement of man." Other theologians, long concerned about the implications of genetic engineering, worried that the scientists who experimented with cloning were "playing God" and "tampering with God's creation." Articles in evangelical magazines such as *Christianity Today* or *The Plain Truth* have regular articles opposing genetic engineering as "tampering" with genes." Pope John Paul II has taken a position on genetic manipulation, arguing that: "All interference in the genome must be done in a way that absolutely respects the specific nature of the human species, the transcendental vocation of every being and his incomparable dignity. . ." (29).

In his scientific paper itself, Dr. Wilmut called attention to the problem of whether "a differentiated adult nucleus can be fully reprogrammed." He called the lamb in question

6LL3 rather than Dolly, and made it clear, in diagrams and illustrations of gels, that there is some question about the precise genetic relationship between Dolly and the "donor" (30). Somatic DNA, which was the source of Dolly's genes, is constantly mutating. Dolly, in fact, may not be genetically identical in every way to her "mother," a point that is of some importance for the possible agricultural applications of cloning techniques.

For the media, however, such technical details were less important than symbolic associations. The cloning of a lamb was immediately set in a context of other fears about genetics and genetic manipulation, and even in a context of more general fears about science and its applications. One journalist compared cloning to weapons development. Another worried that the shortage of organs for transplantation would be resolved by cloning anencephalic babies (who are born without a brain but are otherwise normal) so that their organs could be harvested for patients in need. And a writer for *Newsweek* related the creation of Dolly to broader concerns about food biotechnology by speculating about "cloned chops" (31).

Dolly also evoked an amazing range of media humor. A *New York Times* journalist interviewed writer Wendy Wasserman who wondered what you would say to your shrink if you are your own mother (32). A cartoonist in the London *Guardian* depicted a women comforting a cab driver who had just run over her husband: "That's alright, I have another one upstairs." A writer predicted a new action movie called "Speed Sheep" in which thousands of cloned sheep clogged Interstate 95. Headlines of cloning stories revelled in puns: "An udder way of making lambs," "Send in the clones," "Little Lamb, who made thee?" "Will there ever be another ewe?" and "Getting stranger in the manger." And inevitably there was the anticipation of "Double Trouble."

More pointed jokes — as well as serious editorial commentaries — expressed the growing tensions over commercial control of biotechnology and its implications for the commodification of the body. Just as the GWU experiment evoked images of a cloning industry and breeding factories, so Dolly evoked cynical references to "test tube capitalists," and sardonic queries about a market for genetic "factory seconds" and "irregulars." Meanwhile *Business Week* anticipated "The Biotech Century" in which cloning animals is just the beginning: "It's all happening faster than anyone expected" (33).

As in the media coverage of gene therapy, the messages evoked by Dolly have ranged from promises of miracles to portents of disaster. Editorial appeals called for regulation and for a moratorium on cloning experiments. As political and social pressures began to grow, scientists responded, defending the importance of the work. Media images were "selling science short." The calls for regulations and restrictions, some argued, ignored the medical benefits that could follow from cloning experiments and their potential contribution to the development of life-saving treatments and the testing of new drugs. We are not interested in playing God, said James Geraghty, president of the biotechnology firm, Genzyme, but in "playing doctor" (34). Mammalian cloning could help to generate tissue for organ transplantation and encourage transgenics experimentation. And certainly research using cloning would enhance scientific knowledge about cell differentiation. The politicians who sought a ban on cloning research, said the scientists, were "shooting from the hip."

But media coverage continued to reflect mistrust of this kind of science, concern that commercial interests would ignore social considerations, and fear that the outrageous possibilities suggested by a cloned sheep will eventually, perhaps inevitably, be realized. News reports and media headlines suggested that "Science fiction has become a social reality." "Whatever's Next?" And, inevitably, "Pandora's Box."

Dolly, for a very brief period had reinforced media myths about science — evoking both euphoric fantasies and horrible nightmares and eliciting a fear of science out-of-control. Yet, only a few months after the media blitz, Dolly and the problems of cloning ceased to be news. By June 1997, when the National Bioethics Advisory Commission appointed by President Clinton reported its recommendations to continue the moratorium on federal research funding in this area, and to consider federal legislation, the media had lost interest and paid little attention to the report. In July 1997 the British scientists who had cloned Dolly cloned Polly and three other lambs, this time from fetal rather than adult cells. Polly, a transgenic lamb, carries a human gene in her cells. But cloning was already old news, and this event was reported as merely one more technical advance — the fusion of a fibroblast cell from a fetus to an egg cell. For the media, cloning, even when it involved a human gene, was accepted as routine.

Dolly in the media had been more than a biological entity; she became, ever so briefly, a cultural icon, a symbol, a way to define the meaning of personhood and to express concerns about the forces shaping our lives. She provided a window on popular beliefs about human nature, on public fears of science and its power in society, and on concerns about the human future in the corporate-driven climate of the biotechnology age. A more lasting preoccupation for the media has been research on the genetics of human behavior, research that purports to explain age-old questions about human differences and offers tantalizing, if problematic, prospects for developing technical solutions to social problems.

GENETICS OF HUMAN BEHAVIOR

The language of biological determinism is pervasive in the press. A media survey found references — some more plausible than others — to the genetic basis of shyness, directional ability, aggressive personality, caring tendencies, exhibitionism, homosexuality, dyslexia, job success, arson, traditionalism, preferred styles of dressing, tendencies to tease, political leanings, religiosity, criminality, intelligence, social potency, and zest for life (26). The media refers to selfish genes, pleasure-seeking genes, criminal genes, celebrity genes, homosexual genes, couch potato genes, depression genes, genes for genius, genes for saving, and even genes for sinning. They are presented

in the media as if they are simple Mendelian disorders, directly inherited like brown hair or blue eyes.

Theories of behavioral genetics seek to explain human differences, and like other theories purporting to explain race or gender differences, they have enjoyed a very active press. The idea of biological determinism had attracted considerable news coverage following the controversy over Jensen's claims about the relationship between race and IQ. It later reappeared as sociobiology in media reports that were less concerned with substance than with provocative images, social implications, and policy applications. In selecting this subject for extensive coverage, journalists in effect have used a controversial theory to legitimize a particular point of view about the importance of biological determinism.

Sociobiology is a field devoted to the systematic study of the biological basis of social behavior. Its premise is that behavior is shaped primarily by genetic factors, selected over thousands of years for their survival value. Its most vocal proponent, E.O. Wilson from Harvard University, contends that genes create predispositions for certain types of behavior and that a full understanding of these genetic constraints is essential to intelligent social policy. He believes that sociobiology is "a new synthesis," offering a unified theory of human behavior. "The genes hold culture on a leash," he wrote in his book *On Human Nature*. "The leash is very long but inevitably values will be constrained in accordance with their effects on the human gene pool" (35).

During the early 1980s Wilson's arguments about human behavior, extrapolated from his research on insect behavior, were attacked by other scientists for their purported justification of racism and sexism, their lack of scientific support, and their simplistic presentation of the complex interaction of biological and social influences on human behavior (36). But the 1980 publication of Wilson's first book on the subject, *Sociobiology, A New Synthesis*, was welcomed in the *New York Times* as a "long awaited definitive book." Subsequently sociobiological concepts appeared in newspaper and magazine articles about the most diverse aspects of human behavior. They were used, for example, to explain:

- Child abuse. "The love of a parent has its roots in the fact that the child will reproduce the parent's genes." (*Family Week*)
- Machismo. "Machismo is biologically based and says in effect: 'I have good genes, let me mate.'" (*Time*)
- Intelligence. "On the towel rack that we call our anatomy, nature appears to have hung his-and-hers brains." (*Boston Globe*)
- Promiscuity. "If you get caught fooling around, don't say the devil made you do it. It's your DNA." (*Playboy*)
- Selfishness. "Built into our genes to insure their individual reproduction." (*Psychology Today*)
- Rape. "Genetically programmed into male behavior." (*Science Digest*)
- Aggression. "Men are more genetically aggressive because they are more indispensable." (*Newsweek*)

The press has been especially attracted to sociobiology's controversial implications for understanding stereotyped sex differences. The theory, we are told, directly challenges women's demands for equal rights, for differences between the sexes are innate. *Time*, for example, tells its readers that "Male displays and bravado, from antlers in deer and feather-ruffling in birds, to chest thumping in apes and humans, evolved as a reproductive strategy to impress females" (37). And a *Cosmopolitan* reporter, citing the "weight of scientific opinion" to legitimize his bias, writes: "Recent research has established beyond a doubt that males and females are born with a different set of instructions built into their genetic code" (38).

The media have readily picked up on every research project suggesting there might be a genetic basis of sex differences. In 1980, for example, two psychologists, Camille Benbow and Julian Stanley, published a research paper in *Science* on the differences between boys and girls in mathematical reasoning (39). Their study, examining the relation between Scholastic Aptitude Test scores and classroom work, found that differences in the classroom preparation of boys and girls were not responsible for differences in their test performance. The *Science* article qualified the implication of male superiority in mathematics: "It is probably an expression of a combination of both endogenous and exogenous variables. We recognize, however, that our data are consistent with numerous alternative hypotheses." But the popular press was less qualified, writing up the research as a strong confirmation of biological differences and a definitive challenge to the idea that differences in mathematical test scores are caused by social and cultural factors. The newspeg was not the research but its implications.

The authors themselves encouraged this perspective in their interviews with reporters, where they were less cautious than in their scientific writing. Indeed, they used the press to push their ideas as a useful basis for public policy. According to the *New York Times*, they "urged educators to accept the possibility that something more than social factors may be responsible. . . .You can't brush the differences under the rug and ignore them" (40). The media were receptive. *Discover* reported that male superiority is so pronounced that "to some extent, it must be inborn" (41). *Time*, writing in 1980 about the "gender factor in math," summarized the findings: "Males might be naturally abler than females" (42).

The most striking feature of the media articles on sociobiology was how easily reporters slid from noting a provocative theory to citing it as fact, even when they knew that the supporting evidence was flimsy. A remarkable article called "A Genetic Defense of the Free Market" that appeared in *Business Week* clearly illustrates this slide. While conceding that "there is no hard evidence to support the theory," the author wrote: "For better or worse, self-interest is a driving force in the economy because it is engrained in each individual's genes. . . .Government programs that force individuals to be less competitive and less selfish than they are genetically programmed to be are preordained to fail." The application of sociobiology that he calls "bio-economics" is controversial, he admitted; nevertheless, it is "a powerful defense of Adam Smith's laissez-faire views" (43).

The journalists who write about the genetics of behavior recognize, indeed rely on, the existence of controversy to enliven their stories. Yet most articles convey a point of view by giving space to advocates and marginalizing critics — often described as "few in number but vociferous," or people who are "unwilling to accept the truth." In 1976, for example, *Newsweek* suggested that Wilson was a victim like Galileo: "The critics are trying to suppress his views because they contradict contemporary orthodoxies" (44). In 1982 *Science Digest* compared the criticism of sociobiology to the attacks by religious fundamentalists on the theory of evolution: "Like the theory of evolution, sociobiology is often attacked and misinterpreted" (45). This a comparison places sociobiology's scientific critics, such as Stephen J. Gould and Richard Lewontin of Harvard University, in the same league as William Jennings Bryan.

In a seamless transition, the sociobiological ideas that appealed in the early 1980s have drifted into genetic explanations of social behavior and human differences. During the 1990s the media have disseminated the ideas about the inherited basis of behavior that were generated by studies of identical twins reared apart. These controversial studies gained both media attention and public credibility as part of the growing popular interest in genetics. Offering a simple explanation of complex behavior and a reinforcement of prevailing stereotypes, these ideas appealed to the media that began to attribute an extraordinary range of behaviors to "the genes." *US News and World Report* published an authoritative-looking table providing precise percentages of how much personality traits were determined by heredity rather than culture: extroversion 61 percent, conformity 60 percent, worry 55 percent, creativity 55 percent, optimism 54 percent, and so on (47). In the same issue, *US News* published an article called "How Genes Shape Personality," claiming that "solid evidence demonstrates that our very character is molded by heredity." It suggested that the future of Baby M, the child in a controversial surrogacy dispute, may not rest on which family got her, but in her genes (46). In 1992 *Time* once again offered an explanation of sex differences: "Nature is more important than nurture" and it is just a matter of time until scientists will prove it (47).

The concept of genetic predisposition has appeared to explain a range of personality characteristics. A 1993 *New York Times* article was called: "Want a room with a view? Idea may be in the genes" (48). But especially attractive to the media are explanations associating aggression and violent behavior with biological predisposition. Throughout the coverage of behavioral genetics are references to "bad seeds," "criminal genes" and "alcohol genes." To a *New York Times* writer "evil is embedded in the coils of chromosomes that our parents pass to us at conception" (49). And in a news report of a murder involving the arrest of a 14-year-old high school boy from a "good home," the *New York Times* interpreted the event as a key piece of evidence in "the debate over whether children misbehave because they had bad childhoods or because they are just bad seeds." The reporter used the power of inheritance to explain the incident: "Raising Children Right Isn't Always Enough" read the headline. The implications? There are simply "bad seeds" (50).

The acceptance, indeed promotion of genetic explanations of behavior, reflects in part the media's idealization of science as an ultimate authority. But it also reflects the tendencies of science to justify social stereotypes and popular policy agendas. By its selection of what theories to champion, the press in effect uses the imprimateur of science to support a particular world view. It does so, however, with little attention to the substance of science, its slow accumulative process, and the limits of these theories as an adequate explanation of complex human behavior, shaped by multiple genetic and environmental influences.

THE REPORTING ON BIOTECHNOLOGY RISK

The media coverage of biotechnology has been, in large part, enthusiastic, optimistic, and, indeed, promotional. Yet there remains a persistent and pervasive ambivalence about the implications of this rapidly developing field. Though journalists raise few questions about the ultimate benefits of genetics research or the credibility of research claims, they frequently question the potential abuses of genetic information. We read of the importance of genetic explanations of disease, and then are warned that the ability to identify genetic predisposition is far ahead of therapeutic possibilities. Stories extol the benefits of genetic research, but then decry the risk of gathering genetic information. We are told that this is the dawn of a new genetic era, and then cautioned about an impending eugenic nightmare.

Journalistic attention has often focused on the risks of biotechnology, especially in the area of agriculture and food production where biotechnology applications have been a target for public demonstrations. Some of this reporting is futuristic — abstract speculations about the possible harm of bioengineered products that are yet to appear. But journalists have mainly reported on existing controversies. One of the earliest biotechnology disputes focused on the field testing of Ice Minus, the genetically altered microbes that were developed to inhibit ice crystallization so as to protect strawberries and other fragile crops from injury from frost. Environmental groups, concerned about health hazards, were strongly opposed to these tests. Attracted to a growing controversy and to public demonstrations, news reports about the Ice Minus field tests included striking and provocative photographs of the workers who were spraying the fields, wearing protective clothing that resembled the moon suits associated with the cleanup of toxic chemicals and nuclear wastes (51).

Opposition to the bioengineered Flavr Savr tomato also gained considerable media attention. As in the case of cloning, the issue appealed to journalists as much for the irresistible potential for puns as for evidence of real risk. The genetically engineered tomato, introduced by the biotechnology firm Calgene in late 1991, was initially welcomed in the press as a fruit that would not rot on the way to the market. The product generated media stories on the "wonders" of high technology foods — leaner meat, celery sticks without strings, crisper and sweeter vegetables — and the press supported Calgene's effort to

classify its product as a food rather than a drug subject to FDA regulations. But then, as critics of biotechnology moved in, skepticism became fashionable, and journalists began to write about the tomato as a "frankenfood," a "killer tomato." There was a "tomato war" and a "tomatogate" (52). The idea of injecting mouse genes into food, the spectacle of restaurant chefs boycotting a tomato, the concern about "safe soup," attracted reporters who covered this product as an example of the risks of biotechnology. The business press responded by denouncing "crackpots and scaremongers" who hold back the "wheels of progress" by playing on public fears.

The bioengineering of transgenic animals was also controversial especially among animal rights activists whose colorful antics have often attracted the attention of reporters (53). The media uncritically followed the antics of animal rights groups who projected images of composite cattle, "geeps" (half goat, half sheep) and grossly oversized, distorted pigs. Reporters also addressed the concerns of small farmers who believed that costly applications of biotechnology advances would give further economic advantage to agribusiness, and the opposition of religious groups who worried about the meaning of scientists tampering with nature and "playing God." Here again, media-savvy Jeremy Rifkin was able to use the press to attract publicity for his antibiotechnology campaign. But, as we saw in the evolution of media coverage of cloning disputes, reporting on transgenic foods and animals is usually focused on the newest or most dramatic events. The media, for example, only briefly mentioned Polly, the ultimate transgenic animal who had been cloned in 1997 with human genes.

In some striking ways the images pervading the media coverage of biotechnology risks are remarkably similar to those that were projected during the nuclear power controversy — Frankenstein monsters, mutant animals, mad scientists, and consumers without choices who are captive to an industry portrayed as out of control. Reporting in both of these areas has suggested that risk in the media is often a surrogate issue. Fears of biotechnology are linked to ethical and religious issues, to concerns about economic inequities, and to deep mistrust of a commercially driven science. And media reporting reflects the sensitive question of consumer choices about food and environmental quality — controversial issues of considerable interest to newspaper readers.

CONCLUSION

The media serve, in effect, as brokers between science and the public, framing social reality for their readers and shaping the public consciousness about science-related events. They are for most people the only accessible source of information about important scientific and technical choices. Through their selection of news about science and technology, the media help set the agenda for public policy. Through the information they convey about biotechnology risks, they may affect stock market prices in a volatile industry and influence product sales. And through their presentation of science news, they shape personal attitudes and public actions.

Thus scientists and the companies involved in biotechnology research have been extremely sensitive to their image in the press. And like advocates in any field, they are prone to overestimate the benefit of their work and minimize its risks. Hoping to shape that image, they have become adept at packaging information for journalists. It was not journalists, but scientists who initially employed attention-seeking metaphors to describe the genome as a "blueprint of life," a "Book of Man," a "medical crystal ball." Geneticists themselves have promoted the "gene of the week" and touted the latest therapeutic possibilities. And agricultural researchers were the first to promote the economic benefits of cloning. Courting media attention, those engaged in biotechnology and genetics research have also helped to evoke premature enthusiasm and optimistic expectations. The media have not created the science news; they have mainly amplified and disseminated the messages conveyed by scientists themselves.

Most journalists have limited knowledge about science, and they are vulnerable to manipulation by their sources of information. But as the above review suggests, the media are conveying mixed messages about the costs and benefits of biotechnology. Journalists seem to welcome the notion of biological determinism — simple, startling, and easy to convey — yet many writers remind their readers of the history of eugenics and place current research in this threatening historical context. While journalists report with enthusiasm and wonder the promises of gene therapy, they also warn about potential abuses of genetic manipulation. The media have welcomed agricultural innovations, but they have also warned about potential health or environmental risks.

Perhaps the most striking feature of the media reporting on biotechnology is a pervasive concern about the social and economic context in which this field is developing. The ties between the science of genetics and its commercial applications have invited widespread media cynicism. Reporters have repeatedly called attention to the nonscientific interests — the investments and profits in an intensely competitive field — that are driving biotechnology and its clinical and agricultural applications. Science journalists have long maintained an image of academic science as a pure and unsullied profession, a neutral source of authority, and an objective judge of truth (4). They are skeptical of corporate driven science. These days, according to some disillusioned journalists, scientists working in biotechnology and genetics are "greedy entrepreneurs" or "molecular millionaires" driven by economic interests that threaten their objectivity and override concerns about abuse. Political cartoons portray geneticists in less than flattering terms as bumbling, naive, and unaware of the social implications of their discoveries. And news reports and editorials repeatedly call attention to troubling aspects of the growing links between science and industry, the conflicts of interest that are inevitable when profits and ethics collide.

The context of science, especially in the fields of biotechnology and genetics, has radically changed in recent years. And media coverage, appropriately, is beginning to reflect the implications of these changes.

The media have retained their history of technological optimism, but they are also expressing a growing concern that the expanding commercial interests in these profitable areas of science will overide important social considerations.

BIBLIOGRAPHY

1. E. Slosson, quoted in D.J. Rhees, Master's Thesis, University of North Carolina, Raleigh, 1979.
2. S. Krimsky, *Genetic Alchemy*, MIT Press, Cambridge, MA, 1982; M. Rogers, *Biohazard*, Knopf, New York, 1977.
3. M. Altimore, *Sci. Technol. Hum. Values* Fall, 24–31 (1982).
4. D. Nelkin, *Selling Science: How the Press Covers Science and Technology*, 2nd ed., Freeman, New York, 1995.
5. D. Nelkin, in M. Rothstein, ed., *Biology and Culture*, Johns Hopkins University Press, Baltimore, MD, 1998.
6. D. Nelkin and L. Tancredi, *Dangerous Diagnostics*, 2nd ed., University of Chicago Press, Chicago, IL, 1994.
7. S. Panem, *The Interferon Crusade*, Brookings, Washington, DC, 1984.
8. *N.Y. Times*, May 27, 1980.
9. *N.Y. Times*, Letter, June 17, 1980.
10. L. Wingerson, *Discover*, February, 60–64 (1982).
11. F. Clydesdale, *Food Technol.*, September, 134–136 (1989).
12. *Time*, January 17, 1994.
13. L. Altman, *N.Y. Times*, January 10, 1995.
14. R. Goodell, *The Visible Scientists*, Little, Brown, Boston, MA, 1977.
15. W. Gilbert, *Current State of the HGI*, Harvard University Dibner Lecture, Harvard University, Cambridge, MA, 1990.
16. K.W. Culver, *Sciences* **1**, 18–24 (1991).
17. N. Angier, *N.Y. Times*, December 3, 1993.
18. A. Rosenfeld, *Longevity*, May, 42–53 (1992).
19. S. Brownlee and J. Silberner, *U.S. News and World Rep.*, November 4, 64 (1991).
20. P.E. Dewitt, *Time*, March 20, 70 (1989).
21. Cartoon by S. Goldenberg, *N.Y. Times*, September 16, 1990.
22. Adapted from D. Nelkin and S.M. Lindee, *Cambridge Rev.* (1998).
23. *Los Angeles Times*, October 27, 1993.
24. *Review of Media Coverage of Human Embryo Cloning*, Newsletter, January 1, Center for Biotechnology Policy and Ethics, Texas A&M University, College Station, 1994.
25. *Time*, November 8, 1993.
26. D. Nelkin and S. Lindee, *The DNA Mystique: The Gene as Cultural Icon*, Freeman, New York, 1995.
27. R. Langreth, *Wall Street. J.*, February 24, 1997.
28. G. Niebuhr, *N.Y. Times*, March 1, 1997.
29. Address to the Pontifical Academy of Sciences, October 8, 1994, quoted in *Family Resource Center News*, Winter (1996).
30. I. Wilmut et al., *Nature (London)* **35**, 810–812 (1997).
31. L. Reibstein and G. Beals, *Newsweek*, March 10, 58–59 (1997).
32. Quoted in *N.Y. Times*, February 27, 1997.
33. Cover story of *Business Week*, March 10, 1997.
34. J. Geraghty, quoted in *Gene. Eng. News*, April 1, 10 (1997).
35. E.O. Wilson, *On Human Nature*, Harvard University Press, Cambridge, MA, 1980.
36. U. Segerstrale, *Biol. Philos.* **1**, 53–87 (1986).
37. *Time*, August 1, 1977.
38. *Cosmopolitan*, March 1982.
39. C. Benbow and J. Stanley, *Science* **210**, 1262–1264 (1980).
40. *N.Y. Times*, December 7, 1980.
41. P. Weintraub, *Discover*, April, 15–20 (1981).
42. *Time*, December 15, 1980.
43. *Business Week*, April 10, 1978.
44. *Newsweek*, April 12, 1976.
45. *Science Digest*, March 1982.
46. *US News and World Report*, April 13, 1987.
47. *Time*, January 20, 42 (1992).
48. *N.Y. Times*, November 30, 1993.
49. D. Franklin, *N.Y. Times Mag.*, September 3, 36 (1989).
50. M. Newman, *N.Y. Times*, December 22, 1991.
51. S. Krimsky, *Biotechnology and Society*, Praeger, New York, 1990.
52. *Barrons*, May 31, 1993, and the *N.Y. Times Mag.*, November 21, 24 (1993).
53. J. Jasper and D. Nelkin, *Animal Rights Crusade*, Free Press, New York, 1992.

See other entries EDUCATION AND TRAINING, PUBLIC EDUCATION ABOUT GENETIC TECHNOLOGY; INTERNATIONAL INTELLECTUAL PROPERTY ISSUES FOR BIOTECHNOLOGY; PUBLIC PERCEPTIONS: SURVEYS OF ATTITUDES TOWARD BIOTECHNOLOGY.

MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT

ROBERT MULLAN COOK-DEEGAN
Georgetown University
Annapolis, Maryland

OUTLINE

INTRODUCTION

National policies, including but not restricted to government rules and actions, have profoundly influenced the pace and direction of biotechnology. Medical products and services were among the first and most significant applications of biotechnology. The economic impact of new biotechnologies was felt first in pharmaceuticals, even

before other sectors that saw early applications, such as agriculture and environmental applications. Many nations have pursued policies to cultivate the growth of biotechnology. Many of the first applications were developed in the United States. This was not because the United States had a coherent policy to promote biotechnology. It did not. The United States does, however, have a set of policies to promote health research, and in particular a political structure that has supported the consistent growth of the National Institutes of Health (NIH) since the end of World War II. Those health research policies spawned medical biotechnology.

Development of diagnostic and therapeutic products depends to an unusual extent on government and private funding for research, intellectual property protection, norms governing academic science, product regulation by government, and historical and cultural factors that influence how national governments frame issues arising from medical biotechnology. This entry reviews how national policies influenced the development of medical biotechnology, using the United States as a case example.

DEFINITIONS

Biotechnology is the practical use of living things. At this level of generality, however, agriculture, forestry, fishing, ranching, and many other activities would be included, whereas most intend to refer to the practical applications of molecular biology, and in particular the structural analysis of DNA and proteins. In its most common sense, biotechnology was a term used first by stock analysts to describe a set of companies that began to form late in the 1970s to exploit recombinant DNA, cell fusion, and other methods of molecular biology (1).

One of the first and most influential reports on biotechnology, *Commercial Biotechnology, An International Analysis*, was completed in January 1984 by the Congressional Office of Technology Assessment (OTA) (2). That report distinguished so-called new biotechnology from the old, and focused mainly on "dedicated biotechnology firms," those largely or solely devoted to using the new molecular biological techniques. OTA's definition, modified to accommodate other new techniques of molecular biology, remains useful and formed the basis for another OTA report eight years later, *Biotechnology in a Global Economy* (3).

Technologies to analyze the structure of DNA and proteins have evolved rapidly, so the methods that are new change each year, but the term biotechnology has continued to refer mainly to academic and commercial activities that depend on the structural analysis of DNA and proteins. This definition also includes some activities of major pharmaceutical firms, companies, and research groups that develop instruments used in biology. Activities with direct practical relevance to use of information from cellular and molecular biology have come into being since biotechnology became a widely used term, such as computer analysis of DNA and protein structure and large-scale genetic analysis of organisms (genomics). Medical biotechnology, as used here, refers to the use of molecular biological techniques to develop drugs and diagnostic services in established firms as well as those founded for this purpose. Biological instrumentation and informatics firms (or activities) are also often counted as biotechnology, but their focus is generally on markets to supply research tools. They are generally excluded here, except in a section on patent policies for research tools.

FACTORS AFFECTING THE COMMERCIAL DEVELOPMENT OF BIOTECHNOLOGY

The 1984 and 1992 OTA reports analyzed several industrial sectors for which biotechnology was relevant. Among the areas considered, only the pharmaceutical sector was mainly focused on human medical applications. OTA identified 10 factors that influenced the commercial development of biotechnology (Fig. 1). The first nine are listed in descending importance as judged by OTA. OTA judged the tenth factor, public perception, more variable and unpredictable, at times playing a major role in policy and at other times taking a back seat to the other factors:

1. Financing and tax incentives for firms
2. Government funding of basic and applied research
3. Personnel availability and training
4. Health, safety, and environmental regulation
5. Intellectual property law
6. University-industry relationships
7. Antitrust law
8. International technology transfer, investment, and trade
9. Targeting policies in biotechnology
10. Public perception

All 10 factors are significantly affected by social values and government policies.

CAPITAL AVAILABILITY

The practical applications of molecular biology, especially recombinant DNA techniques, took the form of dedicated biotechnology companies with explicitly commercial aspirations. This was particularly true in the United States. The geographic origins in the United States are best explained by a combination of three factors: availability of capital to form new companies, public funding for biomedical research, and strong university–industry ties.

The availability of capital to form companies to exploit technological opportunities has proved crucial in most high-technology sectors. After World War II several methods of raising capital were developed. The first venture capital firm was established in conjunction with MIT scientists and Boston bankers (4). Eventually the San Francisco Bay Area became an even more active center for venture capital (5). By the 1970s and the dawn of the new biotechology, venture capital firms that had grown up to fund computers, software, and telecommunications were open to help finance the launch of commercial biotechnology. The impetus to found Genentech, for example, came from venture capitalist Robert Swanson

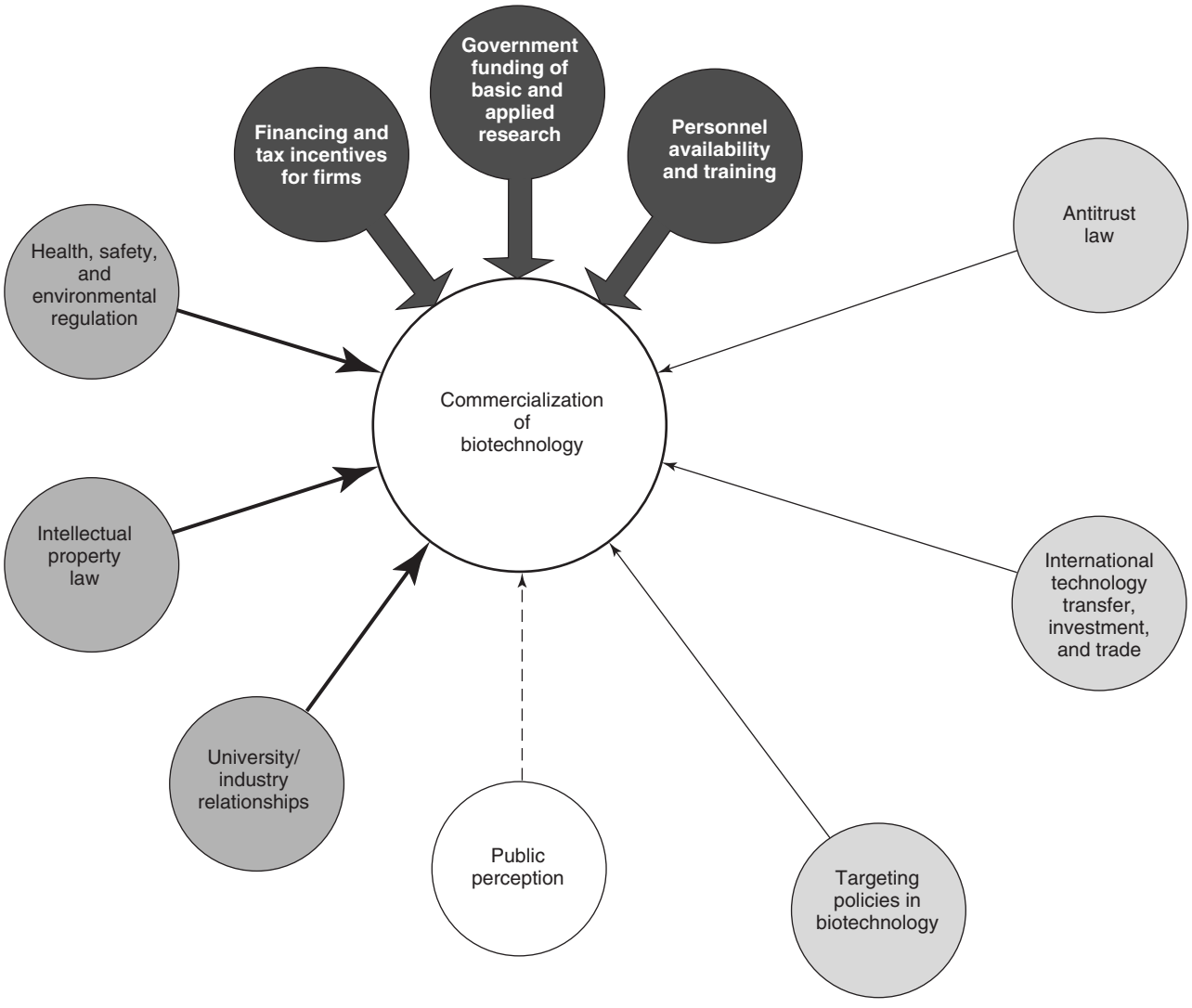

Figure 1. Relative importance of factors affecting the commercialization of biotechnology. *Source: Office of Technology Assessment*.

who approached scientist Howard Boyer of the University of California, San Francisco (6). The financier approached the scientist, and not vice versa. Once the seed was planted and companies had formed, some scientists and university administrators began to approach venture capitalists, rather then the reverse, but the origin of the modern biotechnology sector clearly arose from a member of the investment community. The national and regional environment for capital formation and access to investors willing to fund new technological ventures were critical to the emergence of biotechnology.

The need for early investment to found new companies has attracted attention in Europe and Asia, leading to government incentives and private efforts to create markets along the lines of those started in postwar Boston. Venture capital firms have become a worldwide phenomenon.

In addition to formal venture capital firms, there is another larger but less conspicuous and less easily characterized capital market. Individuals or small groups of wealthy individuals who invest several hundred thousand or millions of dollars in new ventures are often called "angels." They function as an informal market to finance new ventures and often fill gaps in the venture capital markets because they can move more swiftly. They typically close deals with a handshake. The angel market is harder to study, but is comparable in size and at least as important as the formal venture capital market (7).

The angel and venture capital markets enticed university and government scientists to found companies that were privately held, with most equity shared among the investors, scientists, and founding managers. Privately held firms often later become corporations with publicly traded stock. Once a firm's stock was publicly traded, investors could more readily trade their equity for cash.

The availability of venture capital and angel markets depended on a diverse set of government policies, including antitrust, tax (local, state, and national), and other domains of public policy. As a general rule, however, public policy governing individual investments and venture capital was far more subject to private sector actions than deliberate government policy fostering innovation. Indeed,

formal government policies intended to encourage startups and availability of risk capital has tended to be local and late in the game, and sometimes even been impediments, rather than a coherent national policy helpful early in the process of spawning a new industry (4,5). The policies that most influence the availability of startup capital have been in financial institutions and among individual investors rather than the product of deliberate government action. In contrast, government policy has been absolutely crucial in the other most important factor influencing the genesis of medical biotechnology: publicly funded biomedical research.

FUNDING FOR RESEARCH AND DEVELOPMENT

Biotechnology companies were founded to exploit a technological base that grew from substantial and sustained public investment in biomedical research. The term molecular biology first referred to a grants program funded by the private Rockefeller Foundation in the 1930s (6). As Rockefeller Foundation administrator Warren Weaver first used the term, molecular biology addressed scientific problems in the life sciences by importing techniques and scientists from the physical sciences, especially physics and chemistry. Molecular biology, and its spin-offs into commercial biotechnology, is the child of federally funded research but the grandchild of policies first developed in private philanthropy.

Before World War II, government funding for biomedical research was relatively sparse throughout the world. Academic medicine, of which research was a component, had a strong tradition in Germany, France, Great Britain, and other countries in Europe and Asia. Most research was conducted "on the side" in hospitals, or funded through private philanthropy. In the United States, the federal government funded less than private sources — the Rockefeller Foundation, the Foundation for Infantile Paralysis (later the March of Dimes), the Carnegie Corporation, universities and hospitals, and other private philanthropies.

In the years leading up to World War II, Mary Lasker and the American Cancer Foundation (later the American Cancer Society) began to pursue a new strategy for biomedical research funding that focused on inducing federal investment through the political process. After the war, this strategy caused an explosive growth of the NIH budget (9,10). The movement already underway to fund cancer research, and then heart disease research, merged with a consensus favoring federal investment in basic research most conspicuously articulated by President Roosevelt's wartime science advisor, Vannevar Bush (11). This consensus did not become embodied in the form of Bush's proposed National Research Foundation, which would have administered biomedical, military, and general science under a single roof. Instead, the Navy established the Office of Naval Research in 1946, and the other armed services then created their own research and development (R&D) organizations. The Atomic Energy Commission was created to support nuclear physics and to apply it to both military and civilian uses. Most relevant to biotechnology, NIH began to take shape as the nation's foremost patron of biomedical research.

Mary Lasker, then-Senator Claude Pepper, and a succession of NIH directors formed an "iron triangle" to expand federal funding (the vertices of the triangle were nongovernment advocates for medical research, NIH administrators, and congressional champions). They chose to focus resources on university-based research, funded through disease-oriented institutes. Through the late 1960s, Lasker forged strong ties to the chair of the House appropriations subcommittee that funded NIH, Representative John Fogarty, and to his Senate counterpart, Senator Lister Hill. Supported strongly on the inside by NIH Director James Shannon, these congressional patrons boosted biomedical research funding substantially year after year (Fig. 2) (12).

The iron triangle was reconstructed with different players after the death of Fogarty in 1967 and retirement of Hill and Shannon in 1968, and the tactics were replicated by groups wanting "their" institute (for heart disease, for neurological illnesses, for eye diseases, etc.). The direct access that disease advocacy groups had to Congress drove the rise of NIH's budget. Health research grew consistently for five decades, more consistently and more substantially than other federal science programs (Fig. 3).

NIH grew into the world's largest source of support for biomedical research. In a survey of articles from 1973 to 1980, for example, the National Cancer Institute alone accounted for 40 percent of all cancer research publications in 275 medical journals (13). Other U.S. sources — including government (mainly other NIH institutes), private nonprofit, and for-profit firms — accounted for roughly another third, with the remainder unknown or funded by an institution outside the United States. The level of public support for health research in the United States has been a critical factor in biotechnology, and it helps explain biotechnology's geographical origins. Medical applications of new technologies appear likely to remain dependent on science for the foreseeable future, so biomedical research policy will continue to be a decisive factor in the development of medical biotechnology.

The basis for this consistent and substantial growth in health research was not a desire for economic growth but for the conquest of disease. The relevance of medical

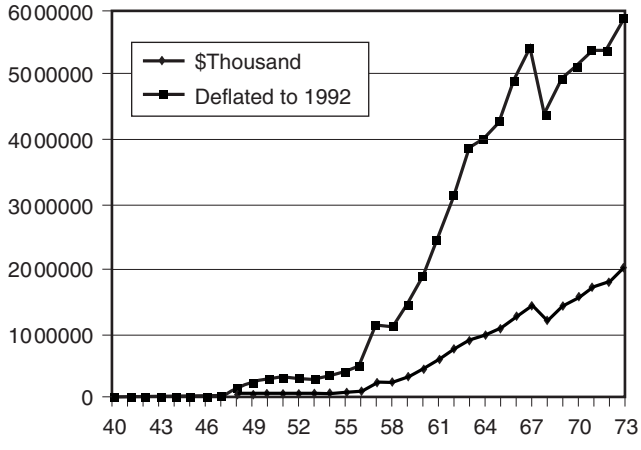

Figure 2. Postwar growth of the NIH budget.

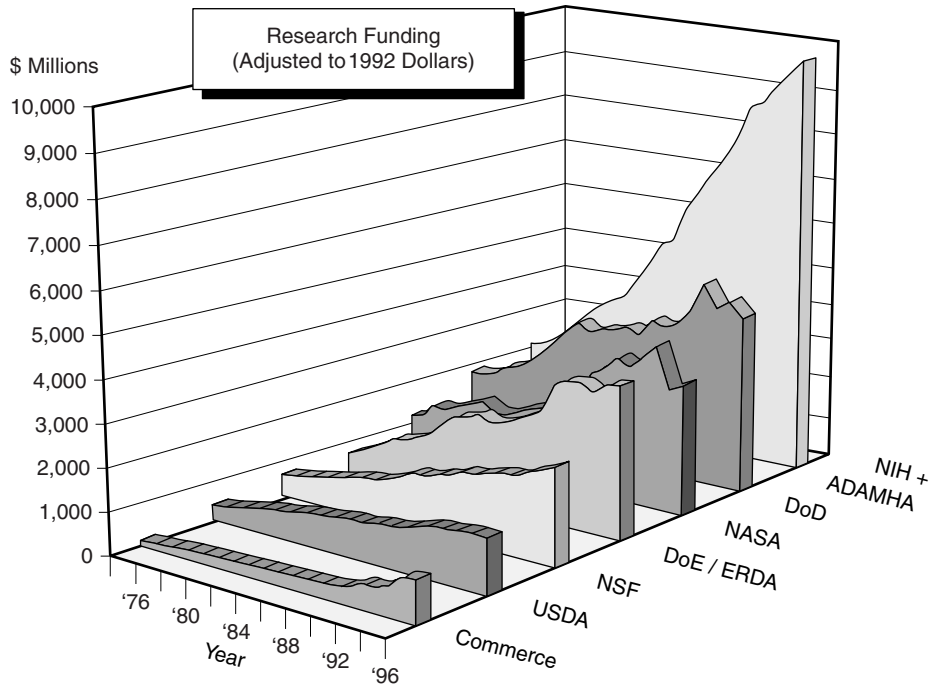

Figure 3. Rise of NIH among U.S. R&D agencies. ADAMHA = Alcohol, Drug Abuse, and Mental Health Administration, DoD = Department of Defense, NASA = National Aeronautics and Space Administration, DOE/ERDA = Department of Energy (and Energy Research and Development Administration for relevant years), NSF = National Science Foundation, USDA = U.S. Department of Agriculture, Commerce = U.S. Department of Commerce.

research to the private commercial development of drugs, devices, and services was foreseeable and foreseen, but promoting such uses was a subsidiary goal, not the primary goal.

The most important two institutions in the birth of biotechnology were universities and start-up firms, with the start-up firms seeded by university scientists. As biomedical research budgets rose and NIH institutes proliferated, the science base grew enormously. It grew most in universities and private research institutes. While Vannevar Bush's vision of one central R&D agency did not become policy, another important feature of his vision did — the emphasis on federal funding for university-based research, as opposed to government owned and operated laboratories more common in other nations and prewar U.S. R&D. (Agricultural research, the grandfather of federal R&D, was typically done by federal employees in government laboratories; military research until World War II waxed and waned with military commitments, and scientists were brought directly into the military in times of war (14).) NIH did conduct some research on its own campus, and in 1953 added a research hospital. During the Vietnam War, the intramural program of NIH grew dramatically in both size and international prestige, in part because working at NIH was a means to avoid induction into the military, but even during this era, most biomedical research dollars flowed to universities and private academic research centers affiliated with them. Those universities were in turn the birthplace of biotechnology.

Coupled with a powerful biomedical research lobby, the decentralized structure of U.S. science enabled a disproportionate growth of the life sciences, particularly medical research. One reason that health research enjoyed growth rates in excess of those in physics, chemistry, engineering, and other fields is the political popularity of health research. Improving health through research is an accepted federal responsibility in both political parties. Boosting health research budgets, moreover, does not mean having to reduce other research budgets because the United States lacks a central research ministry or ministries. With the exception of the National Science Foundation, research budgets are part of budgets for mission agencies, health and defense being the two largest, but also including space, agriculture, energy, and environment. Budgets for research are decided by congressional appropriation subcommittees that do not have to trade off a reduction in defense research that might affect computing, for example, to obtain more dollars for health research. That is, whereas most countries have a research ministry that must set priorities, no unitary science budget exists in the United States. While this is irrational to the degree that research accounts, which are intended as investments to secure future benefits, are mixed with other accounts that focus on current consumption spending, it has enabled different research fields to expand and contract independently, and health research has expanded consistently, in part because of the perceived scientific opportunities, but also because of political popularity and citizen advocacy. In some countries the science and technology ministries are independent, in others, part of a larger body with responsibility for education (e.g., Japan) or industry (e.g., the United Kingdom). The unique separation of powers and the delegation of appropriation authority to Congress under the U.S. system has led to a uniquely rapid growth of health research that would have been far less likely under Bush's unitary National Research Foundation, and it has not been replicated in other developed economies.

The United States funds both a larger total R&D budget and a far higher share devoted to health within that than other countries (see Table 1). If items not directly

Table 1. Share of R&D for Health from NSF S&E Indicators

Objective	United States (1994)	Japan (1994)	Germany (1993)	France (1993)	United Kingdom (1994)	Italy (1993)	Canada (1992)
Total R&D ($U.S. billion, 1995)	68.3	18.1	15.0	13.7	8.7	5.2	3.4
Percent R&D for health	16.5	3.0	3.3	4.5	7.2	6.1	7.8

Sources: From *Science and Engineering Indicators, 1996*, National Science Board (National Science Foundation), table 4–32; Office of Economic Cooperation and Development, Main Science and Technology Indicators database, Paris, June 1995.

related to creation of new knowledge and technology are taken out of the U.S. R&D budget, as recommended by the National Academy of Sciences (NSF), the total U.S. 'federal science and technology' budget is a smaller total, but the fraction devoted to health in 1994 rises to 28 percent (15), more than twice as concentrated on health as the next closest country. Some nations (especially Japan) categorize a substantially higher fraction of science funding as "advancement of knowledge;" there general university funds cover research that in the United States would appear in NIH "health" research accounts. Michaud and Murray have estimated health R&D as a fraction of gross domestic product (GDP), arguably the most pertinent statistic, and by their estimates the United States remains first, but by a far smaller margin (16). The real disparities are therefore not as large as the figures of the Office of Economic Cooperation and Development (OECD) figures suggest, but the U.S. research system clearly tilts much more heavily toward life sciences and health than do other countries.

Anemic public support for research is an important factor explaining Japan's relatively small role in biotechnology (2,3) for example, and the rich base of public and nonprofit support in Europe, especially the United Kingdom, goes a long way to explain Europe's large role in biotechnology. While government funding for R&D has not been generous in the UK compared to other developed countries, the R&D fraction devoted to health is high, and private philanthropies play an unusually large role, funding biomedical research at levels comparable to government (the Wellcome Trust and the Imperial Cancer Research Fund, among others). Countries with a strong pharmaceutical base, such as the U.K., Switzerland, Germany, France, and Denmark are major powers in biotechnology.

In addition to the overall level of support, the responsiveness of the R&D system is also important. Investments in promising new scientific fields and in emerging technologies must be available to sustain innovation. A system that funds specific projects case-by-case through a system of peer review is far more adaptable than one that allocates most funding through institutions (15). Examples at each end of the peer review spectrum are investigator-initiated grants funded by NIH and NSF, on one hand, and the system of institutional funding of the former Soviet Academy of Science, on the other. While institutional funding achieves many notable successes, and peer-reviewed grants can fund poor science, as a general heuristic, systems that channel a substantial fraction of their funds through competitive peer review appear to produce better work and respond better to new opportunities in the long run.

OTA's third crucial factor, the availability of trained experts, is closely tied to academic science, and thence to publicly funded health research. This is especially true in the emergent phase of a science-dependent technological sector, when the only source of training may be university laboratories. In many countries, academic training is funded through educational ministries, and the size and flexibility of those ministries, and their attentiveness to needs of an emerging field such as biotechnology, is highly dependent on policy. In the United States, most academic training in the life sciences is covered by NIH research budgets, and so the growth of NIH has created an ample supply of labor. Indeed, indicators suggest the increase of available graduate students and postdoctoral fellows has exceeded even the growth of NIH research budgets, creating a surfeit of young, trained personnel who may turn to careers in biotechnology (17).

PATENTS

Patents are more important in pharmaceuticals and biotechnology than in most other economic sectors. Patents are grounded in national laws and international treaties, and therefore heavily dependent on government policy. Governments (or organizations delegated authority by governments, e.g., the European Patent Office) convey patent rights to private parties, enabling them to exclude others from making, using, or selling an invention. The period of exclusivity generally extends from the date a patent is issued until 20 years after the patent application is filed, although patent terms can be extended in the United States under certain circumstances. Enforcement of patent rights ultimately falls to the government through administrative procedures and litigation in national court systems.

Several different kinds of patents are relevant to medical biotechnology. The most valuable are patents covering both the composition of matter (that is, the protein or chemical) and its method of manufacture. The path to current biotechnology patents leads through several seminal events. In 1980, the U.S. Supreme Court permitted patenting of a microorganism, in the case of *Diamond v. Chakrabarty* (18). That same year, the U.S.

Patent and Trademark Office issued the first of three landmark patents covering recombinant DNA to Stanford and the University of California (19–21). These Cohen-Boyer patents ultimately produced over $200 million in revenues for the two universities until the patent expired in December 1997 (22). Thus began a succession of patents for genes, DNA fragments, methods of producing useful gene products, and methods for making and characterizing DNA. Between 1980 and 1993 the U.S. Patent and Trademark Office issued over 1000 DNA-based patents, and many thousands more have been issued since (23). The rationale for patenting DNA was that while genes were found in nature, isolating and making those genes into a useful form required substantial inventive activity (24–26). A DNA or other biotechnology patent must meet the same criteria as other inventions. Three principal criteria are used worldwide, with some variation in details of interpretation. An invention must be (*1*) new (or "novel"); (*2*) inventive (or "nonobvious"), and (*3*) useful (or have "utility"). In return for the exclusion rights conferred by the patent, the inventors must describe the invention in sufficient detail to enable others to make it work ("enablement").

Therapeutic pharmaceuticals constitute the largest and most financially rewarding applications for medical biotechnology. There is considerable argument about the cost of discovering and marketing a new therapeutic pharmaceutical. The average cost is estimated in the hundreds of millions of dollars to introduce a new product (27). This average is not meaningful in any particular case because the cost varies tremendously. Through the pharmaceutical sector as whole, however, and hence for policy purposes, it is clear that R&D costs are high, and highest for therapeutics. New therapeutic products thus face the highest costs but also yield the highest payoffs. Patent protection is most important for therapeutic drugs, medical biotechnology's most profitable subsector.

Medical biotechnology encompasses many other products and services with diverse product cycle times. Some biotechnology instruments may resemble short computer or telecommunication device product cycles, measured in years rather than decades, and biotechnology informatics firms face the rapid cycle times of software, with product cycles as short as a year or two. Medical diagnostic products typically reach the market faster than therapeutics because they face a less elaborate regulatory gauntlet. Diagnostics do require FDA approval, and so face regulatory hurdles higher than for most nonmedical goods and services. The capital and time needed to secure market approval for diagnostic products, therefore, are generally lower than for therapeutics but higher than for most nonmedical goods and services. Where product cycle times are short, lead-time advantages, trade secrets, and other factors tend to be more important competitive factors than patents, although patents may still be important in negotiating licensing and technology-sharing arrangements among firms, and in raising capital for new firms. The importance of patents in therapeutic pharmaceuticals builds on a century of industrial history.

Modern biotechnology is the extension of a long trend in which pharmaceutical discovery and development has become increasingly dependent on scientific and technical advance. The modern pharmaceutical industry grew by applying science to medicine (28). Early products were often discovered initially as folk remedies or through clinical observation, but modern pharmaceutical research grew to prominence by making the hunt for new products more systematic. A century ago acetylsalicylic acid (aspirin), was identified as more potent for pain relief than the salicylic acid in bark extract, and then manufactured on a massive scale by Bayer. Companies began to synthesize and screen large numbers of compounds for medical effects. They turned to the emerging methods of organic chemistry to create the compounds and to manufacture them. Firms also used the growing power of microbiology and physiology to discover and screen potential drugs. Antibiotics were first found as products of soil fungi that used them as defenses against bacterial attack, and the industrial innovation was to identify and purify the active agents, and devise ways to produce them cheaply on a grand scale.

In the pharmaceutical business, drug discovery is only one determinant of success. Efficient conduct of clinical testing and manufacture are just as important. Pharmaceutical firms must also manage large and complex distribution systems, manufacturing plants, and they devote enormous sums to marketing. The end markets in health care are financially complex and heavily regulated, requiring expertise and management. At root, however, success follows introduction of new drugs, and innovation through drug discovery where patent protection is paramount.

The time horizon for investment in pharmaceuticals is quite long compared to other industries (29). More than a decade typically passes between discovery of a lead compound and market introduction of a therapeutic pharmaceutical (27). The level of R&D investment is also unusually high in pharmaceuticals, rivaled only by software and a few other high-technology sectors. The pharmaceutical industry invests an estimated 19 percent of revenues in R&D (30). Firms invest heavily in R&D because drug discovery is a major basis for competition. Since the early 1980s private pharmaceutical R&D investment has grown even faster than the NIH budget.

The strength of patent protection in pharmaceuticals is one major reason such substantial, long-term private R&D investments are possible. Patents increase the price that can be charged on drugs that make it all the way through the pipeline, producing revenues to fund drug discovery and development of future products. Gambardella's econometric monograph on pharmaceutical innovation concluded that managing the process of drug discovery was a powerful predictor of financial growth (31). In several other studies of pharmaceutical innovation, economists have used patents as indicators of success among pharmaceutical competitors (32,33), empirically corroborating this message.

The general criteria for patents are shared worldwide, but the criteria can be applied differently in different

countries. The interpretation of patent claims is left to litigation when one party sues another for infringing its patents. Such litigation is costly and takes place long after the R&D results have been disclosed in the patent application. Patent litigation over recombinant DNA-derived insulin cost over $30 million and consumed half a decade (34,35), and the case was decided in 1997, two decades after the crucial events disclosed in the relevant patents. Battles raged over most first generation protein therapeutics — insulin, growth hormone, tissue plasminogen activator, interferons, interleukins, and others. Such litigation is not only costly and slow, but its outcome is uncertain and may differ from country to country. The same drug may be patented in one jurisdiction and not another (e.g., Genentech's tissue plasminogen activator patent is vadid in most countries, but not the United Kingdom). This can occur with traditional drugs as well, but uncertainty is higher in biotechnology because the interpretation of patent criteria is less well settled. Patent policies may differ among countries in three areas relevant to biotechnology: animal patents, gene patents, and research tools.

If patent policies diverge, it is the patent policies in the largest markets or the first markets that will most influence the development of biotechnology. An invention made in a country can be patented in any other country, and the patents that matter most are those that cover inventions in the most lucrative markets, where prices are highest or more units can be sold. While an inventor tends to file the first patent application in his or her home country, this is not necessarily the case, and many biotechnology patents have been filed first in the largest expected markets, rather than in home countries. The largest national pharmaceutical market is the United States, followed by Japan, which has higher per-capita consumption but generally lower prices. Europe is, in aggregate, roughly comparable in size to the United States. Asian and other markets are growing in size, and recent international harmonization of patent law should lead to policies in developing economies coming to resemble those in the developed economies. Because the United States is among the few countries with no pharmaceutical price controls, its drug prices are higher, and for many drugs, the U.S. market alone accounts for most of the global profit.

UNIVERSITY–INDUSTRY RELATIONS

The pharmaceutical sector in general, and biotechnology in particular, is uniquely dependent on academic research. The late Edwin Mansfield queried industrial executives about the degree to which their product stream depended on academic science (36,37). Pharmaceuticals stood out. Executives estimated that twice as many products either would not have been developed at all or would have been substantially delayed without academic research, compared even to other high-tech sectors. The survey results are corroborated by patent data in this unusually patent-dependent sector. Pharmaceutical and biotechnology patents are much more likely to cite academic research than others (38,39). In most patent classes, academic institutions hold only a few percent of patents, but the fraction is much higher for pharmaceuticals and rises to almost one-third of all DNA-based patents in the period 1980 to 1993 (23). The number of academically owned patents has risen steeply in pharmaceuticals, and unlike a drop-off of citations to academic patents in other fields, citation of academic patents in pharmaceutical and medical patent classes has risen over the past two decades (40).

The patent story is but a small part of a larger story of academic-industrial mutualism in biotechnology. The impetus to create biotechnology firms did come first from private investors, as noted above, but it was the commercial potential of academic science that they recognized. To an even greater extent than the already science-dependent pharmaceutical business, biotechnology has emerged as a hybrid academic-industrial enterprise.

The national policies most influencing academic science are public funding decisions for health research, covered above. The policies of universities and private firms, both individually and in aggregate, also influence the development of biotechnology. Biotechnology companies develop around scientists who first use the technologies and take action to apply them commercially (41). Biotechnology companies first appeared close to universities and centers of biomedical research excellence, especially near (*1*) Stanford and the University of California campuses in San Francisco, Berkeley, Los Angeles, and San Diego; (*2*) MIT and Harvard; and (*3*) Johns Hopkins University and the NIH campus. The highest concentrations of biotechnology firms thus arose in California, Massachusetts, and Maryland, not because of a national targeting policy but because that is where the science was based.

The academic research base is necessary, but not sufficient. If federal funding for academic science were the sole determinant, then biotechnology investment would more closely parallel federal spending for health research in specific regions. While California does lead the nation in NIH grant funding, if the intramural program in Maryland is taken into account, California receives less total NIH funding than Maryland. New York is third in NIH funding and Massachusetts fourth, but Massachusetts is second in number of firms, ahead of Maryland and New York. Moreover the fraction of firms based in California and their capitalization is far more heavily weighted to California than the funding would predict. A combination of academic policies conductive to industrial collaboration at Stanford and the University of California, the availability of venture capital, and the history of specific technologies, especially recombinant DNA techniques, led to biotechnology taking root around San Francisco Bay in California.

Policies on the industrial side of the equation are clearly important in addition to factors on the academic side. The formation of small biotechnology companies is in part a signal that established firms in the relevant markets are not fully exploiting an emerging technology. Established firms have been important in introducing many of the first protein therapeutic products, but the initial discoveries generally took place either in academic laboratories or in dedicated biotechnology firms collaborating closely with university scientists. Pharmaceutical firms that adopt the R&D ethos of academic science — encouraging

open publication, forming numerous collaborations with academic groups, and allowing scientists to influence R&D rather than exclusively dictating objectives from the top down — tend to be more innovative (31,32). While it has not been similarly corroborated by empirical studies of biotechnology firms, the "academic ethos" in dedicated biotechnology firms is widely perceived to be even more important than among established pharmaceutical firms.

The importance of federally funded academic research is not unique to biotechnology, but it is more important than for most other sectors. An historical study of computing and software research notes the critical importance of federal funding and academic research in the commercial development of those fields (42). Publicly funded research at academic institutions appears to be important for several reasons. First, it is directly linked to training those who will further develop the science and technology, whether in academic science, industrial R&D, or in management. Second, the results of academic science are generally published openly, and are thus in theory available to all potential beneficiaries. This enables information to flow readily not only among academic scientists, but also among disparate industrial users. Third, the federal government can invest in research that is expected to produce broad social benefit over the long term. Private firms are not similarly motivated, except in unusual cases of monopoly or extensive market dominance (e.g., AT&T in telecommunications until the 1970s or IBM in computing until the 1980s), because no one firm can expect to capture the benefits of its R&D investment. Federally funded academic research is thus a tide that raises all ships in the biotechnology and pharmaceutical sector.

The pattern of geographic clustering is also shared with other high-technology sectors, such as computer manufacture, software engineering, and telecommunications. The region along Route 128 outside Boston and the San Francisco Bay Area have been particularly well studied (4,5,43). The reliance of industrial innovation on academe is similar to other industries whose early history depended critically on science and technical innovation (44), and the academic–industrial nexus is particularly salient in medicine (45).

REGULATION OF PRODUCTS AND SERVICES

The modern pharmaceutical industry started by purifying and manufacturing agents using chemical methods. A century ago, "patent medicines" were as likely to be peddled using grandiose, unsubstantiated medical claims as they were to be legitimate and effective medications. A system of regulation grew up to combat charlatans and quacks, and that system of regulation became most explicit and elaborate where the most money changed hands and where the health stakes of misuse were highest, that is, for products marketed as therapeutic pharmaceuticals.

For therapeutic pharmaceuticals, that is, chemical compounds that are claimed to alter a body function, firms developing the drug must test the compound for safety and efficacy. Clinical testing is the most expensive and prolonged component in the pharmaceutical product development cycle. It entails several phases of clinical trials. The usual development process entails testing the compound for toxicity in a small number of health individual volunteers (phase I), based on evidence from laboratory and animal experiments that suggest the compound might be clinically useful. If the compound proves safe to administer to humans, phase I is followed by use in a larger number of individuals (usually tens or hundreds) to establish dosage and preliminary evidence of efficacy (phase II), and then larger trials (involving hundreds or thousands of individuals) to establish efficacy and to monitor adverse outcomes and side effects (phase III). If a compound proves safe and effective, it is approved for market. A phase IV of clinical trials for new indications or for other reasons (e.g., a requirement for further data by the regulator) may take place once the drug is approved. Approval to market may be withdrawn if concerns about safety come to light.

Drugs and devices are regulated somewhat differently and details of such regulation vary among countries. In general, however, devices that affect essential body functions or raise potential safety concerns, such as heart valves or respirators, face standards of evidence about safety and efficacy similar to therapeutic pharmaceuticals, entailing a series of clinical trials. Products that pose less direct threats to safety or that merely modify well-studied devices, including most diagnostics, face lower regulatory thresholds but still require premarket approval. Some products of biotechnology, such as analytical software or research tools and medical procedures that do not entail introduction of a drug or other product into the body, may not require regulatory approval before entering the market.

The main impact of product regulation on the emergence of medical biotechnology has been to increase the barriers to entry into pharmaceutical markets by substantially increasing both the R&D costs and the time from discovery to market. The need to raise tens to hundreds of millions of dollars to cover clinical testing has deeply affected the development of biotechnology. A few dedicated biotechnology firms that quickly discovered protein therapeutics of substantial value have grown to rival the smaller established pharmaceutical firms. Most biotechnology firms, however, have found the cost of independent development prohibitive, and have forged strategic alliances with larger firms or have been wholly or partially purchased by established firms. Established pharmaceutical firms have created in-house R&D efforts that do molecular biological research quite similar to that done by dedicated biotechnology firms. The relationship between the established firms and dedicated biotechnology firms has become highly complex, and varies so much case-by-case that generalization is perilous.

PUBLIC PERCEPTION AND POLITICAL PROCESS

The economic, R&D, patent, and regulatory policies analyzed above fail to capture some factors that have influenced the speed and direction of biotechnology development in different nations. OTA refers to these, for want of a better term, as "public perception," lumping

together disparate factors that influence industrial development in different ways and to different degrees. Historical and political factors were the most conspicuous elements OTA addressed in this category. Religious views have also played a major role in national debates about select areas of biotechnology policy, such as reproductive technologies, gene transfer, stem cell and embryo research and their applications. Religious demographics of different countries vary so much that the U.S. case may not be suitable grounds for generalization.

The influence of such factors is not surprising or unusual, and is far from unique to biotechnology. Indeed, all areas of policy are influenced by religious, moral, historical, and political factors. It is impossible to capture all the factors that influence public perceptions, but many of them share a common element — they increase or mitigate public fears about biotechnology and its applications. A few factors that have been particularly important in the development of medical biotechnology deserve special mention. Two historical events, the eugenics movement in the first half of the twentieth century and the recombinant DNA debate of the 1970s, had a particularly strong influence on biotechnology policy.

The rise of national bioethics commissions is a new feature of political process in many countries and international organizations. Such commissions have developed as an attempt to grapple with moral and religious pluralism as well as the desire for explicit analysis of ethical implications of rapid scientific and technical advances. Such commissions typically cover more than biotechnology, often addressing topics in health care, end-of-life issues, reproduction, and other areas. The focus here is on bioethics and genetics, where "public bioethics" has an especially rich history.

Eugenics in the first half of the century cast a dark shadow over genetics in the second half. Eugenics as a term refers to many different notions with varying levels of political coercion, having in common directed inheritance intended to improve human populations. Eugenics first became a political movement in England and the United States (46). The ideology of eugenics spread worldwide and took its most extreme expression in association with racial hygiene, part of the ideology underlying the National Socialist Holocaust (47,48). Eugenics and racial hygiene as political movements became associated with senior academics in anthropology and human genetics (49), and this association carried over into postwar German genetics. Applications from molecular genetics were regarded as suspect in the German Green political party, and in German-speaking Europe, public distrust of biotechnology emerged in both agriculture and medicine. The first applications of biotechnology were controversial in many countries, but the controversies were more protracted and pervasive in Europe. The causal association between the history of eugenics and different perceptions of biotechnology is weak and circumstantial. It is historically plausible, but far from demonstrated. At the least, however, the resurgence of scholarship about eugenics, and particularly about the coercive social policies associated with its most florid expressions on both sides of the Atlantic, has contributed to greater vigilance about untoward consequences of social policies that intrude on choices about marriage, immigration, and reproduction, and also about discrimination based on disability or illness.

The recombinant DNA controversy of the mid-1970s also left a legacy for several decades. The recombinant DNA debate began with scientific concerns about biohazards from gene-splicing experiments. The central concern was that bacteria carrying spliced genes could spread uncontrollably and cause harm to individuals subsequently infected. Molecular biologists declared a moratorium on experiments involving recombinant DNA, which was lifted after a famous meeting at Asilomar in California, when the moratorium gave way to federal guidelines for such experiments devised and effected by NIH (50). The debate activated several social activists, who turned their energies to opposing applications that emerged from recombinant DNA and other techniques of molecular biology, and hence biotechnology. Jeremy Rifkin and his Foundation on Economic Trends became the most prominent antibiotechnology activist in the United States; and similar sentiments entered social movements commingled with Green politics in Europe and Asia. At the root of public distrust of biotechnology lay concern that the speedy advance and power of the new technologies were outstripping government and other social mechanisms to ensure they were applied fairly and safely. As with patent policy, capital formation, and public support for academic science, biotechnology emerged first in the most permissive environment, the United States. While this was where the recombinant DNA controversy first arose, it was also where guidelines governing the research were relaxed first and most extensively among developed economies.

National bioethics commissions and biotechnology share a common origin in policies governing the application of molecular biology. The first national bioethics commission, the U.S. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (National Commission), was the result of almost a decade of congressional debate. The first hearings on such a commission were convened by Senator Walter Mondale in the late 1960s, out of concern that technical advances in biology were proceeding in advance of policies to ensure wise use of the technologies in genetics, reproduction, and medicine. In 1971 attention shifted to human cloning, but there was insufficient congressional support for a bioethics commission until a series of scandals about the conduct of medical research involving human participants led to hearings by Senator Edward Kennedy. The National Commission was established in 1975 and operated until 1978. The legacy of its origins in concern about advances in biomedical research was a single report (51). As the National Commission faded out of existence, it recommended that Congress establish a new national bioethics commission with a mandate beyond protection of human participants in research. Congress did so in the form of the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research

(President's Commission), whose title was admirably self-explanatory if lamentably long. The President's Commission operated from 1980 through March 1983. Two of its reports dealt directly with genetics, on genetic testing and screening and on human uses of recombinant DNA, particularly human gene therapy (52,53). These reports came out just as the term biotechnology was coming into general use. As it neared the end of statutory term, the President's Commission recommended that Congress establish a more permanent body to deliberate and make recommendations about ethical issues in medicine, in particular about rapidly advancing technologies that touched areas of controversy and moral uncertainty, including reproduction and genetics (54). In 1985 Congress acted on this recommendation by establishing a new congressional advisory agency modeled on OTA, the Biomedical Ethics Advisory Commission. That Commission took three years to get underway, largely because of delays in congressional process, and in the end never issued a report. In effect, there was a hiatus from 1983 until 1995, when President Clinton established the National Bioethics Advisory Commission by executive order (55).

During the interregnum of bioethics commissions in the United States, many countries established national bioethics commissions and several international organizations also did so. A November 1992 OTA survey of international bioethics bodies found 31 countries and 5 international organizations had some form of advisory apparatus for bioethics (56). Many countries had more than one body, such as Denmark, France, Australia, and Canada.

The growth of national bioethics initiatives and international organizations with bioethics consultative capacity grew from an effort to link explicit analysis of ethical, social, and legal implications of medical science and technology. In many cases they were an attempt to clarify values and to sample disparate religious and social perspectives on medical practices and emerging medical technologies. Most bioethics bodies included attention to advances in genetics or biotechnology or both within their remit. Several of these bodies, particularly those in Europe and Canada, have proven influential in government policy about biotechnology. In Denmark, Sweden, and Germany, for example, commissions recommended proscribing germ line genetic alterations (introduction of DNA into a person so that genetic changes are inherited), and the national parliaments did so. In many cases bioethics commissions are serving to find common ethical arguments that resonate throughout their respective national cultures, grappling with moral pluralism in the face of advancing technologies and pressing problems arising from practical applications, both immediate and in prospect.

Medical biotechnology differs from other high technologies in that its development directly depends on the participation of people in research, as its applications are intended to influence human physiology. This means that the advance of biotechnology is more immediately relevant to the relief of human suffering than most other technologies, but it also more directly threatens fundamental social values connected to reproduction and inheritance. One response has been an effort to anticipate the social implications of biotechnology through bioethics commissions and other mechanisms. It remains too early to judge the success or failure of these responses.

BIBLIOGRAPHY

1. R. Teitelman, *Gene Dreams: Wall Street, Academia, and the Rise of Biotechnology*, 1989, New York: Basic Books.
2. U.S. Congress, *Commercial Biotechnology: An International Analysis*, Office of Technology Assessment, Washington, DC, 1984.
3. U.S. Congress, *Biotechnology in a Global Economy*, Office of Technology Assessment, Washington, DC, 1991.
4. S. Rosegrant and D.R. Lampe, *Route 128: Lessons from Boston's High-Tech Community*, Basic Books, New York, 1992.
5. A. Saxenian, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*, Harvard University Press, Cambridge, MA, 1994.
6. R.E. Kohler, *Partners in Science: Foundations and Natural Scientists 1900–1945*, University of Chicago Press, Chicago, 1991, pp. 299–300.
7. S.S. Hall, *Invisible Frontiers: The Race to Synthesize a Human Gene*, Atlantic Monthly Books, New York, 1987.
8. U.S. Congress, *Innovation and Commercialization of Emerging Technologies*, Office of Technology Assessment: Washington, DC, 1995.
9. J.T. Patterson, *The Dread Disease: Cancer and Modern American Culture*, Harvard University Press, Cambridge, MA, 1987.
10. S.P. Strickland, *Politics, Science, and Dread Disease: A Short History of United States Medical Research Policy*, Harvard University Press, Cambridge, MA, 1972.
11. V. Bush, *Science: The Endless Frontier*, Office of Scientific Research and Development, Washington, DC, 1945.
12. R.E. Miles, *The Department of Health, Education, and Welfare*, Praeger, New York, 1974.
13. H.H. Gee and F. Narin, *NIH Program Evaluation Report: An Analysis of Research Publications Supported by NIH 1973–1976 and 1977–1980.* NIH & Computer Horizons, Inc., Washington, DC, 1986.
14. A.H. Dupree, *Science and the Federal Government*, Johns Hopkins University Press, Baltimore, MD, 1985.
15. National Academy of Sciences, *Allocating Federal Funds for Science and Technology*, National Academy Press, Washington, DC, 1995, Available at: *http://www.nap.edu/readingroom/books/fedfunds/*
16. C. Michaud and C.J.L. Murray, *Resources for Health Research and Development in 1992: A Global Overview*, World Health Organization, Geneva, Switzerland, 1996.
17. National Research Council, *Trends in the Early Career of Life Scientists*, National Academy Press, Washington, DC, 1998.
18. *Diamond, Commission of Patents and Trademarks, v. Chakrabarty*, Supreme Court of the United States, Washington, DC, 1980.
19. S.N. Cohen and H.W. Boyer, U.S. Pat. 4,740,470 (1988), (to University of California and Stanford University).
20. U.S. Pat. 4, 237,224 (1980), (to S.N. Cohen and H.W. Boyer, University of California and Stanford University).
21. U.S. Pat. 4, 468,464 (1984), (to S.N. Cohen and H.W. Boyer, University of California and Stanford University).

22. National Research Council, *Intellectual Property Rights and Research Tools in Molecular Biology*. National Academy Press, Washington, DC, 1997.
23. S. McCormack, *Personal Communication, DNA Patent Database*, Kennedy Institute of Ethics, Georgetown University, & Foundation for Genetic Medicine, Manassas, VA, 1998, Available at: *http://www.geneticmedicine.org/*
24. R.S. Eisenberg, *Emory Law J.* **39**, 721–745 (1990).
25. R.S. Eisenberg, *Science* **257**, 903–908 (1992).
26. R.S. Eisenberg, *Nature genet.* **15**, 125–130 (1997).
27. U.S. Congress, *Pharmaceutical R&D: Costs, Risks, and Rewards*, Office of Technology Assessment, Washington, DC, 1993.
28. J.P. Swann, *Academic Scientists and the Pharmaceutical Industry*, Johns Hopkins University Press, Baltimore, MD, 1988.
29. National Academy of Engineering, *Time Horizons and Technology Investments*, National Academy Press, Washington, DC, 1992.
30. Pharmaceutical Research and Manufacturers Association, *Annual Survey*, Pharmaceutical Research and Manufacturers Association, Washington, DC, 1997.
31. A. Gambardella, *Science and Innovation: The U.S. Pharmaceutical Industry during the 1980s*, Cambridge University Press, New York, 1995.
32. I. Cockburn and R. Henderson, *Proc. Nat. Acad. Sci.* **93**, 12725–12730 (1996).
33. R. Henderson, A. Jaffe, and M. Trajtenberg, *Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965–1988*. National Bureau of Economic Research, Cambridge, MA, 1994.
34. E. Marshall, *Science* **277**, 1029 (1997).
35. E. Marshall, *Science* **277**, 1028–1030 (1997).
36. E. Mansfield, *Res. Policy* **20**, 1–12 (1991).
37. E. Mansfield, *Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing. Rev. Econ. Statist.* **77**, 55–65 (1995).
38. F. Narin and R.P. Rozek, *Res. Policy* **17**, 139–154 (1988).
39. F. Narin and D. Olivastro, *Res. Policy* **21**, 237–249 (1992).
40. R. Henderson, A.B. Jaffe, and M. Trajtenberg, *Review of Economics and Statistics* **80**, 119–127 (1998).
41. L.G. Zucker and M.R. Darby, *Proc. Nat. Acad. Sci.* **93**, 12709–12716 (1996).
42. National Research Council, *Funding a Revolution: Government Support for Computing Research*, National Academy Press, Washington, DC, 1999.
43. E. Moscovitch et al., *MIT: The Impact of Innovation*, Bank-Boston, Boston, MA, 1997.
44. R.R. Nelson, ed., *National Innovation Systems: A Comparative Analysis*, Oxford University Press, New York, 1993.
45. N. Rosenberg, A.C. Gelijns, and H. Dawkins, *Sources of Medical Technology: Universities and Industry*, National Academy Press, Washington, DC, 1995.
46. D.J. Kevles, *In the Name of Eugenics*, University of California Press, Berkeley, CA, 1985.
47. R.N. Proctor, *Racial Hygiene: Medicine under the Nazis*, Harvard University Press, Cambridge, MA, 1988.
48. M.B. Adams, ed., *The Wellborn Science: Eugenics in Germany, France, Brazil, and Russia*, Oxford University Press, New York, 1990.
49. B. Müller-Hill, *Murderous Science*, Oxford University Press, New York, 1988.
50. S. Krimsky, *Genetic Alchemy: The Social History of the Recombinant DNA Controversy*, MIT Press, Cambridge, MA, 1982.
51. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health, Education and Welfare, *Special Study: Implications of Advances in Biomedical and Behavioral Research*, U.S. Government Printing Office, Washington, DC, 1978.
52. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Screening and Counseling for Genetic Conditions*, U.S. Government Printing Office, Washington, DC, 1983.
53. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life: The Social and Ethical Issues of Genetic Engineering with Human Beings*, U.S. Government Printing Office, Washington, DC, 1982.
54. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Summing Up*, U.S. Government Printing Office, Washington, DC, 1983.
55. W.J. Clinton, *Executive Order 12975: Protection of Human Research Subjects and Establishment of a National Bioethics Advisory Commission*, Executive Office of the President, Washington, DC, 1995.
56. U.S. Congress, *Biomedical Ethics in U.S. Public Policy*, Office of Technology Assessment, Washington, DC, 1993.

See other entries FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE; FEDERAL REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE, FDA, ORPHAN DRUG ACT; see also PATENTS AND LICENSING entries; STRATEGIC ALLIANCES AND TECHNOLOGY LICENSING IN BIOTECHNOLOGY; TRANSFERRING INNOVATIONS FROM ACADEMIC RESEARCH INSTITUTIONS TO INDUSTRY: OVERVIEW; UNIVERSITY-INDUSTRY RESEARCH RELATIONSHIPS, ETHICS, CONFLICT OF INTEREST.

O

OWNERSHIP OF HUMAN BIOLOGICAL MATERIAL

LISA M. GEHRKE
Nilles & Nilles, S.C.
Milwaukee, Wisconsin

OUTLINE

INTRODUCTION

In recent years biological materials have become increasingly important in biomedical research. These materials, which were once considered waste materials, have become essential to research in many promising areas. As a result the demand and potential commercial value for these materials continues to rise. The legal status of these materials depends primarily on the source of origin of the materials. If the material is obtained from a nonhuman source, the law will generally consider it as property and therefore, a commodity that may be freely bought and sold in the market place. However, if the material is obtained from a human source, the law generally fails to recognize it as property.

OVERVIEW OF THE USE OF HUMAN BIOLOGICAL MATERIALS IN BIOTECHNOLOGICAL RESEARCH

The study of the human body, body parts, and biological materials has traditionally played an important role in the advancement of the biomedical and pharmaceutical sciences. The observation and study of postmortem organs, tumors, and other biological materials obtained during surgery has provided physicians and researchers with an understanding of the nature and function of the human body and human disease processes.

The introduction and refinement of biotechnological tools such as recombinant DNA, cell fusion, cloning, and bioprocessing techniques in the early 1970s dramatically changed the role of human biological materials in biomedical research and development. These techniques enabled researchers to go beyond merely observing tissues and cultivating cell lines to actually incorporating the material into new organisms to create life forms not previously known in nature. Recognizing the tremendous economic potential of these products, researchers and venture capitalists began investing in the development of biotechnologically based products and the private commercial biotech industry was born. As a result, the demand for the raw materials, namely biological materials, dramatically increased.

In 1980 the United States Supreme Court significantly changed the U.S. patent law by expanding the scope of patentable subject matter to include to artificial living organism in the landmark decision *Diamond v. Chakrabarty* (1). Prior to *Chakrabarty*, U.S. patent law considered all plants, animals and microorganisms products of nature and not products of invention. Therefore they did not qualify as patentable subject matter. *Chakrabarty* recognized that a living organism not found in nature and created by the intervention of humans, is a product of invention and not a product of nature. Consequentially artificial living organisms, qualified as patentable subject matter under the U.S. patent law.

In 1988 the United States government allocated three billion dollars to fund the human genome mapping project commonly known as the Human Genome Project (HGP). Concurrently, the Human Genome Organization (HUGO) was formed in Europe to coordinate research and foster collaboration between scientists (2). The information obtained from HGP has been a tremendous source in developing understanding of the genetic basis of human function and the genetic roots of disease. This information serves as a foundation for research and development in biopharmaceuticals and therapeutics around the world.

The biotech industry continues to grow as investors recognize the economic incentive created by the availability of patent protection for artificial living organisms combined with the value of the information obtained from HGP. This growth and increased interest in gene-based therapies has also created a corresponding increase in the demand for both human and nonhuman biological materials.

Human biological materials such as human tissues, cells, and other materials previously viewed as biological waste products are now seen as having significant economic value. As a result some groups support legally recognizing human biological materials as property. Other groups, however, opposed any legal recognition of property rights in human biological material because it commodifies the human body, which is believed to be contrary to the U.S. legal tradition, public policy, and human dignity.

PROPERTY, OWNERSHIP, AND PROPERTY RIGHTS

Ownership is a term arising out of common law that refers to an individual's right to exercise control or dominion over property. Ownership allows an individual to use, control and transfer the owned property to the exclusion of others.

Property is something having recognized value that may be owned by an individual. Property is either real or personal. Real property, such as land and items affixed to the land, is immovable. Personal property, however, is movable property. Personal property may be tangible items, such as cars, jewelry, and furniture, or intangible concepts, such as stocks, annuities, and intellectual property. For example, nonhuman biological materials such as microorganisms are considered personal property because they are tangible materials. A patent claiming a process of making the microorganism, however, is considered intangible property.

A property owner has a bundle of rights known as property rights that are created and recognized by the state. The definition of property and the nature and extent of an owner's rights in that property are determined by the state through either legislation or common law. In addition the state defines the remedies available to a property owner if the property owner's rights are violated.

Many states currently recognize nonhuman biological material as property. For example, the statutes of Illinois (3) and Maryland (4) recognize "samples, cultures, microorganisms and specimens" as property. If it is misused, the owner of the property may sue under a theory such as trespass to chattel, to recover damages for any injury to the property resulting from unauthorized use or misuse of the property. For example, an individual may own a specific bacterial culture. If a culture is misused, harmed, or destroyed by another, the owner has legal standing to sue the third party to obtain restitution for the damage to the property.

Human biological material, unlike biological material from a nonhuman source, is not currently recognized as property by any state within the United States. Several legal theories, however, have been suggested as a basis for recognizing human biological materials as property. The legal recognition of human biological material as property would create a unique situation in which the property owner is also the property subject to ownership. In this situation an injury to the person would not only create standing to sue the harming individual for injury to person but would create standing to sue for injury to property as well. Legal scholars have looked to various theories to support the recognition of human biological materials as property.

HUMAN BIOLOGICAL MATERIAL

The term "human biological material" is used to broadly refer to any replenishable or nonreplenishable substance obtained from a human body. Replenishable or regenerative materials include blood, skin, bone marrow, hair, urine, perspiration, saliva, milk, semen, and tears. Nonreplenishable or nonregenerative substances however, include body parts such as oocytes and organs, whether vital or nonvital (5).

Research facilities generally obtain human biological materials from patients, research subjects, paid donors or repository collections. The individual providing the material may be either identified or unidentified and the material sample may be used directly in research or saved in a repository collection for later use. The laws dealing with the use and transfer of these materials vary greatly depending upon the specific type of human biological material, and the source from which it was obtained.

HUMAN BIOLOGICAL MATERIAL AS PROPERTY

The recognition of human biological material as property would allow an individual to sell or transfer his or her biological material for valuable consideration such as cash. This change would impact society and the tissue recipients and as well as the tissue donors. Therefore all of their interests should be considered in this decision.

Society's Interest

U.S. public policy generally opposes recognizing a property right vested in the donor of human biological material. This position arises out of society's responsibility to ensure the safety and welfare of its citizens, and to promote scientific advancement.

Over the past 40 years, the U.S. government and the global research communities have made great advances in enacting legislation and developing ethical standards to ensure the protection of human subjects in biomedical research. A central principle common to all the laws, codes, and standards is respect for the human dignity of each person. In other words, society is generally prohibited from permitting any treatment or activity involving a patient, research subject or tissue donor that would be contrary to the dignity and integrity of that individual. Opponents of recognizing human biological material as property argue that the commodification of the human body or any of its parts is equivalent to the commodification of the human person. Since society has determined that it is a violation of human dignity to treat human beings as commodities, it is also a violation to treat the human body or any of its parts as commodities. In addition, opponents argue that recognizing a property interest in human biological materials will encourage individuals to take unnecessary risks with their health for a short-term economic gain. For example, an individual in need of money may feel compelled to sell tissues or an organ to obtain compensation even though the donation may seriously compromise his or her health.

Proponents of recognizing human biological material as property argue that any imposition of regulations

restricting the sale and purchase of body parts, tissues, or materials is a paternalistic infringement upon an individual's freedom. Under this approach, if an individual is provided all of the facts surrounding the risks and benefits of the donation, he or she should be able to sell his or her biological material for cash or other valuable consideration. Any law or regulation restricting such a transaction is considered an abuse of the state's authority. Furthermore the sale of human biological material is not equivalent to the sale of a human being.

Material Donor's Interest

It is clear that the value of human biological material has increased within the past 30 years. Research involving the use of human biological materials has produced products with great economic potential. In some cases the discovery or the invention is directly related to a specific trait found only in the biological material of a specific individual or group of individuals. Consequentially proponents of legally recognizing human biological material as property assert that the individual or group of individuals who provided the essential biological material should share in the financial gain generated by the product arising out of their unique trait. For example, if an individual exhibits a specific genetic trait that is essential to the formation of a new pharmaceutical composition, he or she should be entitled to a portion of the profits realized by the sale of that product. Alternatively, since human biological material has a recognized value, proponents argue that an individual should be entitled to sell his or her biological material in the same manner as a farmer sells a crop.

Opponents to recognizing human biological material as property assert that the treatment of human biological material as property is contrary to human dignity and integrity in the same manner as discussed above.

Material Recipient's Interest

Researcher are generally opposed to the recognition of a property interest in human biological material for several reason including fear of increased research costs, fear of future liability for the use of biological material obtained in a wrongful manner. Since a large portion of the research involving human biological materials is directed toward the production of biopharmaceuticals and biotherapeutics, researchers argue that a creating any property interest in human biological materials would only serve to stifle research and should therefore be contrary to public policy. Alternatively, some researchers advocate creating a property interest in human biological materials because they believe the tying an economic incentive to material donation would increase the overall supply of biological materials available for research.

POTENTIAL SOURCES OF PROPERTY RIGHTS

In 1987 the U.S. Office of Technology Assessment issued a report entitled *New Developments in Biotechnology; Ownership of Human Biological Tissues and Cells-Special Report* (5), which addressed the various ethical legal and public policy concerns regarding legally recognizing human biological material as property. The report identified several areas of law that may support creating a property interest in human biological material. The most relevant of these areas are discussed below.

Patent Law

A patent is broadly defined as a grant made by a government to an inventor which conveys and secures in him the exclusive right to exclude others from making, using, selling or offering for sale his invention for a term of years (6) Governments grant patent rights to their citizens as economic and social incentives to publicly disclose scientific and technological innovations. Patent protection is intended to provide economic and social incentives for inventive efforts made in the advancement of sciences and for the benefit of all. The promise of high financial returns provides the economic incentive for inventors to publicly disclose the fruits of their research, rather than concealing it from competitors. Therefore society benefits from the public disclosure of scientific advancements.

Under U.S. patent law, a patentable invention must qualify as patentable subject matter and be novel, useful and nonobvious to one skill in the technology of interest. As discussed in a previous section, in 1980 the United States Supreme Court expanded the scope of patentable subject matter to include artificial living organisms in the case of *Diamond v. Chakrabarty* (1). Since that time, inventions directed to biotechnologically modified organisms are commonly the subject matter of U.S. European, and Japanese patents.

Although biological materials, such as human biological materials, are commonly used in the development of the invention claimed in the patent, the patent is directed only to the invention formed of the human biological materials and not the biological materials themselves.

For example, an inventor may use a tissue expressing a particular genetic trait as raw material for new pharmaceutical. Although the tissue may have been essential to the development of the new pharmaceutical, the actual invention is the new pharmaceutical and not the tissue or raw material used in its creation. Since patentable subject matter must be the product of invention, only the new pharmaceutical and not the tissue used to create the new pharmaceutical would qualify for patent protection.

Although patent law may provide protection for inventions formed out of human biological materials, it does not provide a basis for creating a property right in human biological material.

Laws Relating to the Donation and Transfer of Organs Tissues and Other Biological Materials

The law relating to the donation and transfer of human biological materials looks to see if the biological material is replenishable or nonreplenishable. No federal or state law prohibits the sale of replenishable human biological materials such as blood, plasma, and semen. As a practical matter, however, state laws categorize this type of transfer as a service rather than a sale of a commodity to avoid any potential liability, such as product liability, created by the sale of an imperfect or harmful good. If the transfer

of human biological material were treated as a sale of goods, the implied warranties would be subject to implied warranty of merchantability and the implied warranty of fitness under the Uniform Commercial Code (UCC) (7).

The Implied Warranty of Merchantability would require the human biological material to be of "fair and average quality" as described by the seller. It is unclear how this would apply to the transfer of human biological materials. If the materials, such as blood or semen, were infected with a contagion, it would create significant liability for the merchants who distribute the material obtained from donors (8).

Similarly the Implied Warranty of Fitness requires the human biological material described to be suitable for the buyer's purpose (9). The law dealing with the transfer of nonreplenishable materials generally falls within the Uniform Anatomical Gift Act (UAGA) (10) and the National Organ Transplant Act (NOTA) (11).

The UAGA was finally approved in 1968 and has been adopted by all 50 states and the District of Columbia. The Act allows any competent adult to donate any or all of his or her body to medical education, research and transplantation at the time of death.

The NOTA was enacted by Congress in 1984. It prohibits the sale of any human organs such as kidneys, hearts, livers, eyes, and bone marrow. It does not, however, prohibit the sale of human tissues and cells for research, nontransplantation or commercial purposes.

In summary, neither federal nor state law explicitly prohibits the sale of either replenishable or nonreplenishable human biological material if the sale is for research, commercial or nontransplantation related purposes.

Law Trespass to Chattel and Conversion

Under civil law, any intentional interference with the personal property of another is considered "trespass to chattel." Trespass is a civil action that allows an injured party to seek restitution for any injury done to his or her person or property. A trespass to chattel action refers more specifically to any injury done to one's personal property.

If the personal property has simply been taken, the usual remedy is to have the trespassing party return the personal property and to provide monetary compensation for any diminution of the value of the property. If, however, the personal property is converted into another form rendering it useless to the owner, the law allows the injured party to seek monetary compensation under a conversion action, equivalent to the value of the property that was converted. In determining whether a conversion has taken place, the court considers the following factors: the extent and duration of the actor's exercise of dominion or control over a chattel, the actor's intent to assert a right inconsistent with the owner's right of control, the actor's good faith, the extent and duration of the resulting interference to the owner's right of control, the harm done to the chattel, and the inconvenience and expense caused to the owner.

If human biological material is recognized as property, a material donor could assert a property interest in that material. Therefore any unauthorized use of that material would constitute a misappropriation of the donated material and the donor could bring a conversion action for compensation for the unauthorized use.

The Case of John Moore's Spleen

In 1990 the Supreme Court of California considered the legal theories discussed above in the case of *John Moore v. The Regents of the University of California* (12). In this case, the Court determined whether the plaintiff, John Moore, a patient with hairy cell leukemia, stated a cause of action against his treating physician, David Golde, for using his cells in potentially lucrative medical research without his permission. Moore alleged that Golde failed to disclose his preexisting research and economic interests in the cells before obtaining his consent to extract the cells. The Court determined that Moore did state a cause of action for breach of a physician's disclosure obligations, but the Court did not for a conversion action.

Statement of the Facts. On October 5, 1976, Moore visited the UCLA Medical Center after having learned that he had hairy cell leukemia. Moore was hospitalized and extensive amounts of blood, bone marrow aspirate, and other substances were removed. Golde confirmed the diagnosis. At that time the defendant was allegedly aware that certain blood products and blood components were of great commercial value and that these substances, which were found in Moore's biological materials, would provide certain competitive, commercial, and scientific advantages.

On October 8, 1976, Golde recommended the Moore's spleen be removed. Golde informed Moore that the splenectomy was necessary to slow down the progress of his disease. Moore signed a written consent form authorizing the surgery.

Before the surgery, Golde had made arrangements to obtain a portion of Moore's spleen for a special research project following its removal. Moore was never informed of Golde's plan.

Between November 1976 and September 1983, Moore traveled from his home in Seattle to the UCLA Medical Center for treatment at Golde's direction and based on the representations that the visits were necessary for his health and well-being. On each visit, Golde withdrew samples of blood, blood serum, skin bone marrow aspirate, and sperm. Moore was told that these procedures could only be performed there and under Golde's direction.

During the course of these treatments, Golde was actively involved in research on Moore's cells and planned to benefit financially and competitively from the exclusive access to these cells through the physician–patient relationship.

Prior to August 1979, Golde established a cell line from Moore's T-lymphocytes (Mo Cell line). On January 30, 1981, the Regents applied for a patent on the cell line listing Golde as one of the inventors. The patent, U.S. patent number 4,438,032, was issued on March 20, 1984. The Regents and Golde entered into several agreements intended to commercialize the Mo cell line.

Based on the allegations discussed above, Moore filed suit stating 13 causes of action including conversion, lack of informed consent, and breach of fiduciary duty. The

superior court only recognized the cause of action for conversion.

Majority Opinion

The California Supreme Court addressed the issues of Golde's alleged breach of fiduciary duty to his patient Moore, and Moore's action for conversion of his biological materials. The Court ruled that Moore stated a cause of action for breach of fiduciary duty but did not state a cause of action for conversion.

Breach of Fiduciary Duty. Under the law of California and all U.S. jurisdictions, a physician has a duty to inform a patient of all the information material to a proposed treatment or procedure before the patient consents to the procedure. This allows the patient to make an informed decision about undergoing or proceeding with a treatment or procedure. For example, a physician has a duty to inform a patient of the risks that are believed to be relevant to the patient's decision such as the mortality rate and success rate of the treatment or procedure, and the possible side effects associated it.

The Court ruled that a physician's economic interest in a patient's treatment should be disclosed to the patient because it may be relevant to that patient in his or her decision to proceed with the treatment. Golde therefore had a duty to disclose his economic interest and the research involving Moore's biological materials to Moore before Moore then underwent the splenectomy. In addition Golde has a continuing duty to disclose his financial interest in Moore's follow-up care. Since Golde did not inform Moore of this information, the Court rule that Moore did state a cause of action against Golde for breach of fiduciary duty to a patient for failing to obtain informed consent. The suggestion is that if Moore would have been aware of Golde's financial interest in Moore's biological materials, Moore may not have undergone the treatment. This situation raises a question of fact that would be answered by a jury if the case proceeded to trial.

Conversion. The California Supreme Court did not allow Moore's claim for conversion. To state a cause of action for conversion, a party must show that he or she has a property interest in the material converted. A property interest is only present if the material is legally recognized as property. California law, however, does not recognize human biological materials as property. Therefore Moore cannot have a property interest in his biological materials and his action for conversion could not stand.

The Court discussed the public policy concerns in favor of and in opposition to recognizing human biological materials as property. Ultimately the court determined that human biological material is necessary to the advancement of biomedical research, and any impediments to the advancement of this research, such as any uncertainty of title, would be against public policy. The Court determined that a legal recognition of biological material as property should be done through legislation.

Justice Arabian's Concurrence

Justice Arabian believes the majority's opinion correctly concludes that a tissue donor does not have a property interest in his or her donated tissue. Arabian's opinion, however, stems from moral rather than economic concerns. It is feared that a recognition and enforcement of a property interest in human body parts would have a highly negative effect on human dignity. Arabian notes that the plaintiff is not left without a remedy. Moore may pursue an action under the breach of fiduciary duty theory.

Justice Broussard's Dissent

Contrary to the majority opinion, Justice Broussand believes that Moore does, in fact state a conversion cause of action because he alleged that Golde wrongfully interfered with Moore's right to determine how his spleen would be used. Unlike the majority, justice Broussard asserts that conversion is not only applicable when personal property is taken but also when it is used without authorization. This position may be supported by the Uniform Anatomical Gifts Act.

The UAGA allows a donor to control the express manner in which the biological materials will be used after his or her death. Accordingly Moore arguably had a similar right to determine how his biological materials would or would not have been used after their removal.

Justice Mosk's Dissent

Justice Mosk believes the majority's opinion is incorrect because it fails to recognize a difference between the use of materials used for nonprofit research purposes and commercial use. Mosk believes that Golde's failure to disclose his economic interest in Moore's biological materials amounted to commercial exploitation of Moore. Furthermore Mosk concluded that Moore should have a property interest in his tissues that would allow him to enter into contracts for the exploitation of his tissues if he so desires.

Impact of the Moore Decision

Following the Moore case, both public and private parties having interest in the use of human biological materials began to reevaluate their methods and policies regarding the procurement of these materials. In view of the majority opinion the focus of the discussion shifted from concern over the creation of a property interest to reviewing and revising the standards of informed used in obtaining human biological materials from patients and/ or research subjects.

ROLE OF INFORMED CONSENT IN THE PROCUREMENT OF HUMAN BIOLOGICAL MATERIALS

Since the Moore decision, society has developed a deeper understanding of the value connected to human biological material and the potential harms and benefits that may arise out of its commercialization in research. As a result the ethical principle of respect for persons has expanded to require disclosure of any potential harm or benefit resulting from the intended harvesting or use of the material, to the potential donor.

The ethical principle of *respect for persons* recognizes an individaul as an autonomous person capable of deliberation about personal goals who acts under that deliberation. Therefore, to respect the autonomy of the individual is to give weight to his considered opinions and choices while refraining from obstructing his actions unless they are harmful to others. To show lack of respect for an autonomous person is to repudiate the person's considered judgement when there is no compelling reason to do so (13).

Respect for persons requires that each research subject be given the opportunity to choose what will or will not happen to his or her body or biological materials. This opportunity is provided when adequate standards for informed consent are satisfied (13). To accomplish this, an individual must be given all the information regarding the nature of the procedure, the potential risks and benefits arising out of it and all other factors, such as economic considerations, that may bias the situation.

The requirement of informed consent was originally applied to a traditional biomedical research setting in which the subject received an experimental treatment. The requirement was intended to ensure that the research subject knew the risks involved in the procedure to be performed. Traditionally the doctrine of informed consent was not extended to require a disclosure of any future use of the human biological materials taken from a person. In recent years the informed consent requirement has been extended to apply to the procurement and use of any human biological materials obtained from a known donor as in the Moore case.

On December 13, 1995, the *Journal of the American Medical Association* published the results of an NIH workshop charged with formulating informed consent requirements and their application to the gathering of human biological materials that may be used in genetic studies, and with identifying when further consent for the use of material samples already in the possession of researchers should be obtained (14). The participants concluded informed consent is required for all genetic research using samples that may be linked to an identifiable person (14). The participants further concluded that the disclosure of genetic information could have medical, psychological, and economic implications on the material donor. Therefore individuals have the right to put limits on the use of their biological materials. For example, they may specify that their tissues shall only be used by a noncommercial entity.

In August 1999 the U.S. National Bioethics Advisory Commission (NBAC) issued its report entitled Research Involving *Human Biological Materials: Ethical Issues and Policy Guidance* (15). As in the 1995 NIH report, NBAC discusses the importance of human biological materials in the advancement of medical research and recommend obtaining informed consent from all identifiable tissue donors including disclosure of any economic interests of the parties and all potential physical, social, and psychological harms they may arise from its procurement or use.

PRACTICAL APPLICATIONS RECOMMENDATION

As a practical matter, it is essential that any individual, company or organization that is involved in the procurement and use of human biological materials, obtain full informed consent from all potential material donors. More specifically, the economic interests of the researchers, or procuring organization, the proposed use of the materials, and any physical, social, psychological or economic harms that may arises from the use of the materials, must be disclosed to the potential donor prior to the donation. In addition, it is advisable to have the potential donor explicitly transfer any title or property interest to the procuring individual or organization before the donation occurs. Although a donor's property interests in their human biological materials are not currently recognized by any state in the United States, it is important that the potential donor is fully informed about the implications and potential economic ramifications of the transfer of title before it is granted. The biological material may be transferred to a jurisdiction that recognizes such a property interest in the future. Therefore, it is important to take every precaution to ensure the title to the human biological material in question is clear.

BIBLIOGRAPHY

1. *Diamond V. Chakrabarty*, 444 U.S. 397 (1980).
2. Comment, *Am. Univ. Law Rev.* **44**, 2433, 2435 (1995).
3. 38 Illinois Annotated Statute 15-1 (1979).
4. 3A Maryland Code, Article 27, §340(h)(10).
5. U.S. Congress, Office of Technology Assessment, *New Developments in Biotechnology: Ownership of Human Biological Tissues and Cells — Special Report*, U.S. Congress, Office of technology Assessment, Washington, DC, 1987, p. 24.
6. 35 U.S.C. §271 (1996).
7. American Law Institute and National Conference on Uniform State Laws, Uniform Commercial Code (1978).
8. Uniform Commercial Code §2-314.
9. Uniform Commercial Code §2-315.
10. Uniform Anatomical Giftt Act, 8 A.U.L.A. 15 (1985).
11. National Organ Transplantation Act: Public Law 98-507 (1984).
12. *Moore v. Regents of the University of California*, 51 Cal. 3d. 120 (1990).
13. Office of the Secretary, The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, (April 18, 1979).
14. E.W. Clayton et al., *J. Am. Med. Assoc.* **274**, 1786 (1995).
15. National Bioethics Advisory Commission; *Report and Recommendations: Report Involving Human Biological Materials: Ethical Issues and Policy Guidance*, vol. 1, NBAC, Washington, DC, August 1999.

See also PATENTS AND LICENSING entries.

P

PATENTS AND LICENSING, ETHICS, INTERNATIONAL CONTROVERSIES

GÖRAN HERMERÉN
Lund University
Lund, Sweden

OUTLINE

BACKGROUND

The view, in general, on what biotechnological inventions may be patented is more restrictive in Europe than in the United States, although the differences have diminished over the last few decades. Moreover there have been important variations in the United States, Canada, and the European Union (EU) positions, as well as among their peoples over time, as a brief glimpse at their historical backgrounds will indicate.

The 1961 UPOV Convention (Union internationale pour la protection des obtensions végétables) has continued to be endorsed by many European countries and also by the United States and Japan. It provides protection for new plant varieties in a different way than the patent system does. It means that a plant breeder and plant improver can obtain exclusive right to use a particular brand of plant according to certain conditions that are specified in the convention. This convention was revised in 1978 and later in 1991. The new UPOV convention provides a somewhat stronger protection of the rights of the plant breeders.

A new instrument of International patent cooperation, the Eurasian Patent Convention, was signed in Moscow on September 9, 1994, and entered into force on August 12, 1995.

United States

In biotechnology a number of difficulties have been encountered in the attempts to patent living materials. Products of nature are not patentable under U.S. law. Nevertheless, before the rapid expansion of modern commercial biotechnology, protection by patents had been granted on materials derived from natural sources through human intervention, such as substances obtained by purification of naturally occurring products.

In 1930 the Plant Patent Act made it possible to extend protection to new and distinct asexually propagated plant varieties. To be sure, the protection was subject to certain limitations: a research exemption that allows use of protected varieties to develop new varieties, and a farmer's exemption that allows farmers to save protected seed for use on their farms or for sale to other farmers.

Before 1980 patents could be obtained on processes using bacterial strains for commercially valuable purposes such as in producing antibiotics. But bacterial strains per se could not be patented in the United States. However, the United States Supreme Court decided in 1980 that a genetically modified living single-celled bacterium, transformed to give it the capacity to break down multiple components of crude oil, could be patented. The Patent and Trademark Office (PTO) soon extended the categories eligible for patent protection, to such products as corn plants (1985), polyploid oysters (1987), as well as to "nonnaturally occurring nonhuman, multicellular living organisms, including animals."

The first patent on a transgenic animal was issued in April 1988 to Harvard University for the development of a mouse with a human onco-gene that makes it susceptible to cancer, the OncoMouse. This extension to animals has generated considerable public controversy, however. Restrictive legislation, including a moratorium on animal patenting, has been proposed (1–3).

The present American patent law is the Utility Patent Act, in Title 35 of United States Code, which among many other things states that "whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor" (35 US Code 101).

Canada

In Canada, patent legislation governing drugs evolved through a series of amendments to the Patent Act. From 1923 to 1993 Canada operated a system of "compulsory

licensing," described in some detail by Mathews (4), allowing generic copies of patented medicines to be manufactured within Canada and, by 1969, to be imported.

In 1987 the Act was amended. A compulsory license could be issued on patented medicines only after a fixed period of market protection. Moreover a price review board was created to monitor and control prices charged. Brand-name drug companies promised to invest a growing percentage of sales revenue in research and development in Canada, in exchange for patent protection.

A further amendment to the Patent Act in 1993 fundamentally changed the legislation by abolishing the system of compulsory licensing and applying general patent regulations to medicines. In that way Canadian law was brought into line with that of its trading partners, particularly the United States. It is now illegal to sell a copy of the drug until the patent has expired.

Europe

The European Patent Convention was signed in Munich on October 5, 1973. The European policy on patentability is in general stricter than the one in the United States, even though there have been variations among the national patent offices and the European Patent Office (EPO) in Munich. Before 1998 the existing legal framework did not allow the patentability of living organisms in the European Community countries, notwithstanding some decisions of the European Patent Office (which still are contested).

This changed with the Directive 98/44/EC of the European Parliament and the Council issued on July 6, 1998. This directive grants free circulation of patented biotechnological products, guaranteeing compliance with the European Patent Convention from 1973, the trade-related aspects of Intellectual Property Rights agreement of April 15, 1994, and the Rio de Janeiro Convention on Biological Diversity of June 5, 1992.

The European Patent Convention also states that European patents will not be issued for plant or animal varieties and essentially biological processes for the production of plants or animals, with the exception of microbiological processes or the products thereof. Nevertheless, EPO in 1991 issued a European patent to Harvard University on its transgenic mouse, after having refused this application a few years earlier.

According to Schatz (1998) in Europe, only some 15,000 European patent applications have been filed in biotechnology and over 2000 patents have been filed for DNA sequences isolated from the human genome used to develop therapies and medicines.

CONTROVERSIES

There are controversies within and among countries as to what can be patented, the conditions of patentability, how these conditions have been interpreted and applied, as well as concerning the exceptions to these conditions. There are also conflicting views on the entire system of patenting forms of life, in particular animals and plants. By implication, these controversies have a bearing on the reasons accepted for and against the system of patenting different forms of life. Even though there are variations in views between the developing and developed countries, the latter on the whole take a more negative view of the present system, which they feel has been exploited.

The competition over transgenic technology rights has financial implications. Those individuals who have the intellectual property rights can protect their discoveries by patents and earn money by getting royalties from others or by exploiting the intention themselves. It is then hardly surprising that there are objections to such rights and that they have sometimes resulted in court cases (5–7). But the basic issues concern what can be patented and on what conditions.

Conditions of Patentability

The conditions of patentability include utility, novelty, nonobviousness. Some of the clauses in patent laws and patent conventions contain difficult but important expressions, like "utility," "novelty," and "inventive step" as well as "living organism." Each of these conditions can be interpreted and applied in different ways (8,9). Controversies as to how they are to be interpreted and applied in practice are to be expected.

What Is Excluded? And Why?

What is excluded from patenting, and how are these exclusion clauses to be interpreted? In the European Patent Convention (EPC) a broad concept of invention is being used. There are three express exceptions:

1. Methods for treatment or therapy.
2. Essentially biological processes for producing plants and animals (and the results of these processes).
3. Inventions contrary to public order and morality (in Europe at least).

The terms "plant variety," "animal variety," "essentially biological processes," "invention," and "public order and morality" can be interpreted in different ways, since the praxis of interpretation has evolved over time. The novelty requirement presupposes that there is a distinction between earlier existing plant varieties and a new one. In this context the definition and criteria of identity of "plant variety" becomes relevant.

Klett (10), Moser (11), and Thomsen (12) have discussed the interpretation of various clauses of exemption in the legislation and practice of several different countries (Germany, Switzerland, Denmark). Other legal scholars Kern (13) and Schatz (14) have elucidated the views on patentability and the patenting practice in the United Kingdom and EPO. The various changes in the Belgian attitude toward the patenting of biological material have been described and analyzed by van Overwalle (15).

Lugagnani (16) proposes that in setting practical limits to patentability via ethical considerations, the moral judgment should move from exploitation of the invention to the nature and/or objectives of the research and development (R&D) projects that produced it. In his view, unethical R&D activities should not be rewarded by granting intellectual property rights. The crucial question

then becomes: What is the basis of these moral judgments? And who has the privilege of interpretation?

Lugagnani suggests that ethical guidance be derived from the 1996 Council of Europe Convention on human rights and biomedicine as well as from the Directive 98/44/EC of the European Parliament. According to Article 7 in the directive, the European Commission (EC) states that its European Group on Ethics in Science evaluates all ethical aspects of biotechnology. The group has already discussed several aspects in earlier reports, and several others are on their way.

Several tasks ought to be separated in the discussion of these clauses of conditions and exceptions. Clearly, there is the empirical task of finding out how, in fact, they have been interpreted by particular patent offices. In the literature, this has been discussed extensively. But, since the clauses, like a musical score, are in general open to many interpretations, there is also a question with normative implications: How are they to be interpreted? In a way the question can be understood as a means-end question: How are the clauses to be interpreted if one wants to achieve (or avoid) certain ends? Obviously an author can at the same time be interested in discussing and answering all these questions.

What Can Be Patented?

Today's candidates are microorganisms, plants, animals, and genetic material from humans as well as the methods or processes used in biotechnology to modify them. While attempts to manipulate and patent human life, also before birth (embryo, fetus), is controversial and forbidden in some countries, the situation is different when it comes to microorganisms.

If gene technology and patents only served good purposes, for example, in the production of medicines, vaccines, diagnostic methods, or protection of environment (e.g., producing bacteria that break down crude oil), objections would be more difficult to obtain and sustain. A great deal of criticism has to do with the use of human growth hormone to get salmon to grow quicker, for example, and a range of similar enhancement purposes. With regard to animals, the public generally considers it unethical to produce and patent animals with properties that cause them to suffer. It has been argued that plants, animals, and large organisms should not ever be patentable.

The general idea in the July 1998 Directive 98/44/EC of the European Parliament and the Council is that discoveries per se are not considered patentable, nor are plants and animal varieties per se, nor essentially biological procedures for the production of plants and animals. Also methods of surgical and therapeutic treatment and diagnostic methods applied to animal bodies are not considered to be inventions of the sort that may be patented. But plants and animals with newly introduced genetic traits, like the famous OncoMouse, may be protected by patents. The same goes for biological materials and material isolated from its natural environment and isolated elements of the human body with technical processes.

In Europe, inventions contrary to law and order or public morality may not be patented. This includes processes for human cloning for reproductive purposes — banned by a protocol to the Council of Europe Convention on human rights and biomedicine (17). The same holds for processes for modifying the germ-line identity of human beings and the use of human embryos. Processes for modifying the genetic identity of animals with substantial medical benefit for humans, because they are useful for treating serious diseases such as cancer or AIDS, may be protected by patents. If this is not the case, they are not considered to be patentable.

Computer-aided molecular design techniques have recently been used along with more traditional methods to design new peptides possessing specific properties, such as bactericidal activity. Patel and Colleagues (18) have argued that the ability to protect the intellectual property rights associated with the discovery of the new molecules is a key issue for commercial utilization of such peptides, and the authors have described and proposed an extension of established patenting practice.

Moreover Green (19) argues that combinatorial libraries are patentable as long as the library meets the statutory criteria of utility, novelty, and nonobviousness, and the application meets the standards of enablement, best mode, and written description. Licensing and alternatives to patenting are also considered, along with potential problems unique to combinatorial chemistry agreements. A particularly interesting illustration of the diversity of views on the ethical aspects of gene patents, reflecting deep disagreements as to both conclusions and arguments, concerns patenting of disease genes.

Patenting of Disease Genes?

The possibility and desirability of patenting disease genes is a much-debated area where heated controversies have attracted a great deal of attention around the turn of the twenty-first century. For example, McGee (20) has argued that gene patents can be ethical. He examines some arguments against patents in this area and concludes that patenting on methods for detecting the presence of a genetic correlation with disease-related (and other) phenotypes can be appropriate.

The main reason offered for this position by McGee is that there is a subtle distinction to be made between "observing DNA" and constructing a DNA-based product for diagnosis of some disease or phenotype. He examines two arguments that oppose his view: (*1*) genetic information is part of nature and (*2*) gene patents create a "toll bridge" barring research using patented genes.

Taking the first point, McGee argues that disease gene patents are more an innovation of scientists than a discovery. "Finding DNA is discovery. Correlating it with human life for the purpose of creating a diagnostic process is innovation" (20).

The second position poses a problem: How is one to avoid the issue that genetic research will be hindered, that genetic tests will not be developed, if researchers in the early stages of work are not able to pay the tolls necessary to start the research? McGee claims that this can be prevented by "correctly framed and issued gene patents."

He argues that infringement of the patent occurs not when methods are used for further research but only when the actual innovation is put to use in the same sense that the method of detecting the gene's or allele's correlation with a particular phenotype is applied to a patient.

Kimbrell (21) is an important early work that has played a central role in the debates on engineering and marketing of life. Various positions in the subsequent debate have been advocated by several researchers (22–27).

Caplan and Merz (23) have instructively compared the identification of disease genes to the land rush. They agree that the debate is timely, for a number of reasons, but contend that McGee's position is untenable. Caplan (26) argues laying claim to own what we find in the biological world cannot be done without an arrogance or hubris about our creativity, that patenting genes is profane — it is "commercializing something that ought to remain outside the realm of commerce," and that "allowing patents ignores that it is our common genetic glue that holds us together as a species, and that is something all of humanity has as its common legacy." These metaphysical concerns are especially at the heart of religious-based objections to patents, according to Caplan.

Merz and Cho (27) argue that McGee is wrong and that methods of screening for Alzheimer's disease should not be patentable. Obviously they criticize that a U.S. patent on such methods has been granted by PTO. More specifically, they argue that there are two false premises in McGee's argument:

1. The difficulty and effort involved in making a discovery makes the discovery patentable.
2. Patentability is based on the usefulness of the discovery.

They argue that market potential is not a necessary, much less a satisfactory, condition to determining whether something comprises patentable subject matter. They point out that gold and diamonds, though valuable, are not patentable, regardless of who first discovered gold and diamonds and how difficult this discovery was.

Referring, for example, to Eichenwald (28), they conclude by arguing that broad diagnostic patents may led to the abandonment of methods of prenatal or genetic testing that long have been the standard of care: "This is simply unacceptable. These patents are contrary to good medical practice, and must be prohibited" (27).

Also Tribble (29) argues against exclusive licenses with a single commercial developer for all uses of the invention: "This limits both the potential benefits for human health and the potential for greater financial returns to the academic institution that could stem from their research" (29). Therefore he proposes a kind of compromise position. He argues that scientists could protect their intellectual property rights by granting access to the research tools for research use via nonexclusive licenses to scientists in both academic and industrial laboratories, and simultaneously reap financial returns and realize the commercial potential of the patented technology. He points out that "refusal to grant non-exclusive licenses ... could have a chilling effect on biomedical research and would be contrary to public interest" (29).

Magnus (30) criticizes the approach of PTO in the United States, in particular, the constructivist interpretation of genetic testing to avoid the "product of nature" doctrine. (If something is a product of nature, it does not satisfy the earlier mentioned conditions of novelty and inventive step.) He argues that the strategy of the defenders like McGee will result in an intolerable dilemma for physicians.

ALTERNATIVES AND CONSEQUENCES

Some important questions in this debate are as follows:

1. Which are the available alternatives: protection by patents, with or without specified exemptions, trade secrets, protection of intellectual property rights in other ways, such as by the UPOV-convention, compulsory licensing or no protection at all?
2. What is the range of application, if any, of these methods of protection? Each method could have a different area of application (one for microorganisms, another for plants, still another for animals, etc.), they could be partly overlapping, or totally overlapping.
3. What are the actual or likely consequences of these alternatives for the various stakeholders, including the pharmaceutical and agrobiotechnological industry, farmers and animal breeders, the researchers, the universities, the developing countries, and the entire population of the earth, expected to double in less than 50 years?
4. What ethical considerations should be used in evaluating these consequences: (which) deontological considerations, utilitarian theories, contractualist ethics, and the like? For obvious reasons, it will be impossible to take a stand on the latter issue here. What is important is only to make clear that the choice of ethical framework will have practical consequences; hence transparency and explicitness is important.

The first of these questions deserves some elaboration. Patents provide protection of intellectual property rights. Patents also expire after a number of years and are limited to a geographic area. A patent granted by a national patent office is valid in that country, a patent granted by EPO is valid only in Europe, and so forth. Moreover it takes time and is costly to obtain a patent, and the patentee has to check that no infringements on his or her rights are made. If infringements are made, the patentee has to prove that he or she has suffered damage. Besides, it is always somewhat uncertain what the rulings of the courts will be. What is certain, however, is that the system provides patent lawyers with a great deal of work; it is good for them.

Therefore alternatives to the patenting system have been considered, and the consequences of each alternative, within each particular area (microorganisms, plants, animals, etc.) ought to be evaluated in some detail. An

alternative is to keep the invention secret in order to use and develop trade secrets. This may work (as it has for Coca Cola) if the product is difficult to copy. Further alternatives include inventing other forms of protection, like the UPOV convention for plants, or to carry on without protection of intellectual property rights at all.

REASONS FOR THE PATENTING SYSTEM

The controversies over protection of intellectual property rights via patents also concern the foundations of this protective system and the reasons offered for and against patenting. The main arguments in favor of patenting biotechnological inventions center on justice, economic growth, and openness.

Justice

The standard reasons for patenting include considerations of justice. The idea is simply that the inventor should not be without financial reward while others make a fortune on the basis of his or her invention. The inventor is entitled to economic compensation. The underlying conception of justice in this context is justice according to desert. This means, for instance, that universities can get royalties on biotechnological inventions made by their researchers. For big universities like Harvard or Stanford in the United States these royalties amount to very large sums every year. If these inventions had not been protected by patents, big companies could use these inventions free of charge to increase their own profits.

Economic Growth

The second main line of argument is concern for economic growth — the desire to stimulate the biotechnology industry by the possibility of financial rewards. Society needs research and development, and patent laws are designed to encourage this. Protection of intellectual property rights encourages investments by the private sector in agrobiotechnology simply because it increases the chances of financial reward. For instance, Dompe (31) argues that patents are a useful legal instrument to protect intellectual property rights, allowing society to promote and to support the innovation and development of useful inventions.

Openness

The disclosure requirement in patent laws prevents secrecy and attempts by researchers to preserve exclusivity in access to their discoveries. Thus patent laws work in favor of openness. To obtain a patent, the applicant has to send in a detailed description of what the applicant wants to protect. This description should make it possible for anyone in the field to repeat the process or to manufacture the product in the future. These descriptions can be bought at a very reasonable price from the patent office.

It could be objected that the patent system slow's down scientific publications. But in many countries this problem is taken care of by a clause in the patent laws. In order not to discourage early publication of research results because the inventor is considering to apply for a patent on his discoveries, patent laws in the United States and Canada contain a clause about a "grace period," during which time the author may wait to apply for a patent. In fact this encourages early publication, since the author who publishes first also has the right to (apply for) a patent. But to say this is not to say or suggest that the U.S. patent law requires filing before publication.

REASONS AGAINST THE PATENTING SYSTEM

The standard arguments offered against protection of biotechnological inventions by patenting could be classed into two broad categories: (*1*) objections to the underlying technology itself, rather than to its protection under patent laws and (*2*) objections to the actual or anticipated effects of patenting itself on research and health care, for example (32). Here focus will be on the latter ones. But on a more general note, Hoffmaster (33) has argued that patenting of life forms promotes an irreverent materialistic conception of life. It may affect our attitudes toward life negatively when life forms are treated as commodities to be bought and sold in the market (34), particularly when the organisms to be patented have been modified by gene technology so that they possess human genes (2).

Effects on Research

It has been feared that the patent system will distort the research agenda in favor of potentially lucrative projects rather than the traditional mission of expanding knowledge regardless of potential financial gain. It is clear that there is at least a potential tension between commercial imperatives and traditional scientific norms, such as is evident in Merton's oft-quoted CUDOS norms (35):

- Communism, according to which there is no private ownership of knowledge
- Universalism: hypotheses, objections, and criticism should be treated equally, regardless of the age, sex, race, and nationality of the author
- Disinterestedness: impartiality is essential and researchers should not be influenced by their own economic, religious, political and other ideological interests
- Organized Scepticism: everything may be questioned, and this ought to be done systematically

Heller and Eisenberg (36) take the metaphor of the "tragedy of the commons" as their starting point. This metaphor helps to explain why people overuse shared resources, like water and air. The authors argue that the recent proliferation of intellectual property rights in biomedical research may lead to a different tragedy, a tragedy of the "anticommons," in which people underuse scarce resources because too many owners can block one another. The authors conclude that privatization of biomedical research must be more carefully deployed to sustain both research and product development. Otherwise, more intellectual property rights may lead to

fewer useful products for improving human health. This article has provoked several comments and responses (37).

Along similar lines Merz (38) has argued that the rapidly growing number of disease gene patents — patents protecting all methods for diagnosing a particular genetic condition — will threaten the ability of physicians to provide optimal medical care to their patients. Disease gene tests are being monopolized by a small number of providers. This, according to Mertz, threatens to restrict research activities, creates unacceptable conflicts of interest, may reduce patient access to testing, and may lead to inequitable extensions of patent terms on tests and related discoveries. It will make it possible for patentees to dictate the standard of care for testing, and in that way to interfere with the practice of medicine. In view of these risks, Merz suggests an amendment of the patent law to require compulsory licensing of physicians providing medical services. In the United States, a National Institutes of Health (NIH) proposal, discussed by Flores (39), will restrict the licensing of federally funded biomedical research tools for commercial gain.

Effects on Developing Countries

The development of genetically modified plants, fruits, and animals may undermine the export of products (plants, fruits, sugar, raw material, etc.) from developing countries; the products will no longer be competitive. Since the biotechnological industry is situated in the developed countries, this could increase the gap between rich and poor countries. The rich will be made richer, while the poor will become poorer.

It has been feared that patents can be used as an instrument to exploit the third world. A double exploitation has been evoked. Researchers from rich countries, where advanced biotechnology industries are located would harvest and refine desired biological material from the developing countries. The derived products now protected by patents or the UPOV convention, would be sold back to farmers in the countries from which they originate, and the seeds made sterile so that they can only be used for one crop.

Concern has also been expressed in the *Bulletin* of the World Health Organization (WHO) that recent global developments in the regulation of trade and intellectual property rights threaten to hinder the access of populations in developing countries to essential drugs. Velasquez and Boulet (40) have argued for state intervention in the health and pharmaceutical markets in order to guarantee equitable access to these essential products.

Serageldin (41) has argued that biotechnology can contribute to food security if it benefits sustainable small-farm agriculture in developing countries. The agrobiotechnology focus has been on ethical, safety, and intellectual property rights issues. An effective system of protection of intellectual property rights would encourage investments by the private sector in agrobiotechnology by increasing the chances of financial gain. But, as Serageldin argues, in developing countries the needs of small farmers and environmental conservation are unlikely to attract private funds. Public investment will be needed, and new and imaginative public–private collaboration could make the gene revolution beneficial to developing countries as well.

Thus there is particular cause of concern for the effects of licensing and patents on farming and agriculture in the developing countries. Traditionally these countries have had two important sources of income. The first is cheap labor, which is less important today than it was, because of the rapid development and widespread use of industrial robots and information technology. The second important source of income is their raw material and agricultural products, which the developed countries do not need today to the same extent as before, because of the rapid progress of biotechnology.

Effects on Environment

The major part of the world's collected genetic resources are to be found in tropic or subtropic areas, that is to say, mainly in the developing countries. Will genetically modified plants yielding better crops in the future replace traditional domestic plants? Will this lead to a reduction of biological diversity? And is this to be deplored? For what reasons? Here empirical and value questions are intertwined. There is clearly a risk that biological diversity is threatened by present developments in plant breeding.

Will patenting increase the risk for environmental problems? Will patenting of animals make the situations of animals worse? To view the problems in a proper perspective, however, it is important to realize that environments can be endangered also by release of products that are not protected by patents. This is usually the case in air and water pollution. Thus the pressing problems of preservation of the environment need to be addressed independently of the problems discussed in this article.

A precautionary principle has been launched in the international debate, in the wake of the Bovine Spongiform Encephalopathy (BSE) ("mad cow disease") and various threats to our environment: Irreversible processes in nature should not be started, such as by using genetically modified and patented seed (or salmons) that may spread in nature in uncontrollable ways. Ethical arguments in this area had better be regarded as preliminary and provisional, and if mistaken judgments are made, such mistakes should preferably have been done for a good purpose. This, of course, presupposes that ethical analyses are made both of the research that eventually leads to a patent, in the application for a patent, and in the use of a patented product.

Effects on Industry, Farming, and Agriculture in the Developing Countries

The impact of patents on demographical structure of society may alter the settlement patterns of populations. Sparsely populated regions may become more sparsely populated due to migrations. The direct and indirect impacts of migration on various economic interests is another cause of concern and another reason for criticism of the patent system.

It has been argued, for example, that the higher cost of patented livestock or seed will restructure farming. Small farms will be put out of business, and there will be even more concentration of corporate agriculture. Agricultural patenting also undervalues the contributions of less developed countries to biological diversity (42).

Further, with the increasing globalization of the world economy, the biotechnology industry may be restructured worldwide. Knowledge and economic resources may be concentrated in a few big corporations. Already large corporations have merged, and small companies are finding it difficult to survive in the market place, particularly those in the developing countries.

VALUE CONFLICTS AND FURTHER ETHICAL ASPECTS

What are the possible benefits? Which are the risks? To make this explicit, we must look to identify some of the value conflicts. The possible economic benefits are as follows:

- Consumer values, in terms of better-tasting grapes, firmer tomatoes, bigger oranges
- Producer values, in terms of increased crops for farmers, pigs with faster or increased muscle growth
- Company values, increasing revenue, such as by developing and marketing plants resistant to pests afflicting specific species
- Environmental values, such as in developing plants resistant to herbicides so that fewer herbicides are used by farmers

The monopolies achieved by patents will stimulate technological inventions, but they may clash with ideals of free competition. A patent will give an air of legitimacy to an invention, but this is sometimes what is challenged. Free competition is a value, and so is reward for beneficial work.

These values may clash or be combined, according to each particular case. They have to be compared to other positive and negative values wherever the values clash: increasing unemployment among the population, immiserizing growth in the developing countries, health risks for patients (if patents and royalties change the standard practice of medicine so that tests cannot be performed because of expensive royalties), health risks for large segments of the population (if genetically modified products are not safe to consume), environmental hazards (if genetically modified organisms are released into nature and cause ecological problems).

So what economic, social, psychological, environmental, and medical values are at risk, and whose values are these? Further, which values are favored or promoted by the various alternative methods of protecting, or not protecting, intellectual property rights?

Is it the prosperity of one group against the prosperity of another? The prosperity of one group against the health of another? The prosperity of one group against the welfare or quality of life of another? If two families of values clash, and both are equally legitimate, we are facing an optimization problem. If one is arguably more legitimate than the other, the situation is, of course, different.

A Strategy of Analysis

The questions are difficult, strong economic and other interests are at stake, and the terminology invites talking at crosspurposes. Therefore, it is essential to be explicit and open, try to separate different kinds of disagreements, make clear what we know as well as in what areas there are considerable gaps in our knowledge, and to identify the alternatives, the stakeholders and their interests, and the principles used in assessing these interests.

Consequences for human and animal welfare, broadly defined, are clearly important considerations in this debate. Consequentialists argue that they are the only important reasons. Moral philosophers of other persuasions want to supplement consequentialist arguments with deontological, contractual, egalitarian, and other considerations, to use appeals to human rights as a basis for claims as to what human beings and animals are entitled to, or as negative restrictions for what others may do to them.

Many other ethical problems than those discussed here are raised by recent developments in biotechnology, such as issues of informed consent and remuneration to those from whom genetic material has been collected. The celebrated case Moore case (*Moore vs. University of California*) illustrates this. The court ruled that Moore was not entitled to a share in the profit made by the company that had used tissues he donated, but the court criticized the information Moore had received as inadequate (43). The special problems raised by human tissue banks (collection, preservation, access, purpose, etc) are also relevant but will not be discussed here.

Underlying Assumptions

The controversies described here are important and interesting because they bring to the foreground a number of underlying and culture-dependent assumptions on what the human being is, what the place of humans is in nature, whether there is a radical and sharp distinction between humans and other living organisms and what constitutes this difference. That is to say, what differences, if any, exist between humans and the so-called undignified biomass, between what is considered natural or unnatural? To what extent is everything living an end in itself, or are humans entitled to exploit animals and plants as well as human DNA for their own purposes, and so forth? Views on these matters may vary somewhat among cultures, and also within a culture over time. There are also indications of interesting recent value changes in the relations between humans and nature. Here is an interesting area for research.

Who owns the genetic resources, in particular the human genome? Among the alternative views are the following:

1. They are the common heritage of humankind. This is a position undermining the whole system of protecting intellectual property rights in this area

via patents: It implies no patents on disease genes, no royalties on biotechnological inventions involving the human genome.

2. The genetic resources belong to the country where it is found. This could be taken to mean that royalties for patents of biotechnological inventions based on genetic material found between the borders of a country should go to the country of origin. This is a more general point, and it is applicable outside the field of biotechnology and genetic resources as well.
3. The invention belongs to the company or the researcher who developed it, applied for, and was granted a patent — the inventor is entitled to royalties. As already mentioned, at the base of this position is an argument of justice, but a different one from that used to support the previous position.

Since a large share of the world's collected genetic resources is to be found in tropic or subtropic areas, consisting mainly of the developing countries, the second alternative might favor developing countries, and this has been their line of argument for the most part. Alternatives discussed at the Rio Convention have also been tried, calling for the transfer of technology from developed to developing countries and royalties on patents developed from genetic material found in the country, in return for the right to harvest and refine genetically the desired material from a developing country.

These basic views could be backed up by religious, social, economic, and legal reasons. Like the underlying assumptions mentioned above, they point to the cultural context in which this debate is taking place, and where there are considerable differences as to views and traditions both between and within developing as well as developed countries.

Appel (44) argues in his analysis of the human body and patent law that human life must not be allowed to be reduced to "undignified biomass" and that protection should be guaranteed from the beginning of life, from embryonic stage. Can patent offices be in a position of assessing the risk factor of a genetic engineering invention? Is there justification for excluding therapeutic treatment from European patent law? Appel argues that this has an extremely inhibiting effect on research.

As is often true in difficult cases, there are good arguments on both sides and many difficulties in assessing the relevance and tenability of the reasons given. Often one has to take a close look at each separate case before making a judgment.

CONCLUDING REMARKS

In view of the strong commercial interests, and the rapid development of biotechnology, future trends are hard to predict. But it is essential that the ethical aspects of patenting and licensing are not neglected. Different constructions of licensing and protection of intellectual property rights, whether by patents or in other ways, will create winners and losers. Who will be the winners, and who the losers, in the short and the long run? And who should decide this? For careful consideration of these issues, the many underlying conflicts of interests and value clashes need to be made explicit, as well as the principles used in weighing the conflicting interests.

The ethical issues underlying licensing and patents are relevant from the start of the research that eventually leads to the patent, to the examination of the patent application and the decision whether or not to grant a patent, and to the actual use of the patent. This means that all biotechnological research project should be decided under the aegis of an ethical review board, and not just certain medical projects involving patients.

The various patent offices sometimes express frustration when they have to deal with ethical issues. This is not something they are prepared for. It may not be the role of EPO to engage in ethical debates and take a position (45). Other national and international fora exist for the discussion and analysis of the social, ethical, and other extrajuridical aspects of patenting and patentability. One solution may be to suggest different fora and to make a clear division in the discussion and analysis between legal and ethical aspects of patent applications.

Ideally, then, a patent will not be granted if the application offends public order; the issue of morality will be relevant, and controversial applications will be contested. Lawyers will review each application and provide examples of what they consider to be compatible or not compatible with this or that related clause. An attempt could then be made to interpret the clauses in the light of this praxis or to describe the changing praxis of various patent offices.

Since patent applications are certain to come in an increasing extent from different countries, international cooperation and harmonization of patent laws will be important in the future. Comparative research in the present situation of globalization is obviously essential. What are the similarities and differences between the approaches in Europe and the United States concerning plant protection? Concerning protection of biotechnological inventions in general? With focus both on the current framework and future developments (8,46)? The studies should include countries not mentioned expressly here, such as Japan and China. By noting similarities and differences, problems as well as possibilities can be identified.

BIBLIOGRAPHY

1. R.G. Adler, 1988. *Harvard J. Law Techn.* **1**, 1–61 (1988).
2. R. Dresser, *Science* **243**, 1002–1003 (1988).
3. T.T. Moga, *J. Pat. Trademark Office Soc.* **76**, 511–545 (1994).
4. J.H. Matthews, *Clin. Invest. Med.* **19**(6), 470–478 (1996).
5. S. Hensley, *Mod. Healthcare* **27**(23), 52 (1997).
6. J.L. Fox, *Nat. Biotechnol.* **15**(13), 1332 (1997).
7. R. Dalton, *Nature* **400**, 601 (1999).
8. G. Van Overwalle, *The Legal Protection of Biotechnological Inventions in Europe and in the United States. Current Framework and Future Developments*, Leuven University Press, Leuven, Belgium, 1997.
9. R. Rogge, *Int. Rev. Indust. Property Copyright Law* **28**(4), 443–466 (1997).
10. K. Klett, *Int. Rev. Indust. Property Copyright Law* **28**(6), 845–850 (1997).

11. W. Moser, *Int. Rev. Indust. Property Copyright Law* **28**(6), 845–850 (1997).
12. H.C. Thomsen, *Int. Rev. Indust. Property Copyright Law* **28**(6), 845–850 (1997).
13. M. Kern, *Int. Rev. Indust. Property Copyright Law* **29**(3), 247–282 (1998).
14. U. Schatz, *Int. Rev. Indust. Property Copyright Law* **1**, 2–16 (1998).
15. G. Van Overwalle, *Int. Rev. Indust. Property Copyright Law* 1999 (in press).
16. V. Lugagnani, *Forum (Geneva)* **9**(3, Suppl. 3), 93–98 (1999).
17. Council of Europe, Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine, Strasbourg, France, 1996.
18. S. Patel et al., *J. Comput. Aided Mol. Des.* **12**(6), 543–556.
19. G.D. Green, *Mol. Divers.* **3**(4), 247–252 (1998).
20. G. McGee, *Camb. Q. Healthcare Ethics* **7**(4), 417–421 (1998).
21. A. Kimbrell, *The Human Body Shop: The Engineering and Marketing of Life*, Free Press, New York, 1993.
22. G. Poste, *Nature* **378**, 534–536 (1995).
23. A.L. Caplan and J. Merz, 1996. Patenting Gene Sequences. *Br. Med. J.* **312**, 926 (1996).
24. J.F. Merz et al., *Mol. Diagnosis* **2**(4), 299–304 (1997).
25. E. Marshall, *Science* **275**, 780–781 (1997).
26. A. Caplan, *Camb. Q. Healthcare Ethics* **7**(4), 422–424 (1998).
27. J. Merz and M.K. Cho, *Camb. Q. Healthcare Ethics* **7**(4), 425–428 (1998).
28. K. Eichenwald, *N.Y. Times*, May 23, A1 (1997).
29. J.L. Tribble, *Camb. Q. Healthcare Ethics* **7**(4), 429–432 (1998).
30. D. Magnus, *Camb. Q. Healthcare Ethics* **7**(4), 433–435 (1998).
31. S. Dompe, *Forum (Geneva)* **9**(3, Suppl. 3), 113–117 (1999).
32. R.S. Eisenberg, *Yale Law J.* **97**, 177–231 (1987).
33. B. Hoffmaster, *Intellect. Property J.* **4**, 1–24 (1989).
34. T.H. Murray, *Univ. Michigan J. Law Ref.* **20**(4), 1055–1088 (1987).
35. R. Merton, in *The Sociology of Science*, University of Chicago Press, Chicago, IL, 1973.
36. M.A. Heller and R.S. Eisenberg, *Science* **280**, 698–701 (1998).
37. A.H. Fried et al., *Science* **281**, 517 (1998).
38. J.F. Merz, *Clin. Chem. Mar.* **45**(3), 324–330 (1999).
39. M.A. Flores, *Nat. Biotechnol.* **17**(8), 819–820 (1999).
40. G. Velasquez et al., *Bull. World Health Organ.* **77**(3), 288–292 (1999).
41. I. Serageldin, *Science* **285**, 387–389 (1999).
42. J. McGowan, *Cult. Survival Q.* **15**(3), 20 (1991).
43. B.M. Dickens, *Transpl. Proc.* **24**(5), 2118–2119 (1992).
44. B. Appel, *Der menschliche Körper im Patentrecht*, Carl Heymann, Munich, 1995.
45. G. Van Overwalle, in *Biotechnology, Patents and Morality*, Aldershot, Ashgate, UK, 1997, pp. 139–148.
46. G. Van Overwalle, *IDEA-J. Law Technol. (USA)* **39**, 143–194 (1997).

See other entries FEDERAL REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE, FDA, ORPHAN DRUG ACT; HUMAN GENOME DIVERSITY PROJECT; OWNERSHIP OF HUMAN BIOLOGICAL MATERIAL; see also PATENTS AND LICENSING entries.

PATENTS AND LICENSING, ETHICS, MORAL STATUS OF HUMAN TISSUES: SALE, ABANDONMENT OR GIFT

COURTNEY S. CAMPBELL
Oregon State University
Corvallis, Oregon

OUTLINE

INTRODUCTION

Almost every person who enters a doctor's office for a routine examination can be sure to leave some of their body tissue behind. Typically requests for a blood or a urine sample are a minimally invasive part of the standard routine of clinical practice, enabling physician diagnoses of ailments, and more commonly, of good health. In nonclinical settings, such as research studies, samples of human body tissues may be collected for purposes of assessing disease prevalence or susceptibility. DNA analysis of waste tissue, such as saliva or hair, can reveal very detailed genetic information about a person. Persons may also part with their body tissue with the intent to benefit others: Gametic tissue, for example, may be donated or sold to enhance the prospects of fertility treatment of another person in otherwise good health. In acute crisis settings, donation of an organ or tissue, such as bone marrow, may be a life-saving therapy.

There are, in short, many purposes in clinical medicine and biomedical research that involve retrieval and sampling of human body tissues. The procurement and use of human tissue in diagnostic, therapeutic, research, and educational purposes illustrates how contemporary scientific study understands the human body as a source of medical information. The National Bioethics Advisory Commission recently estimated there are over 282 million body tissue specimens collected throughout the United States, a figure that is increasing by an additional 20 million specimens each year (1). This practice has been

accelerated significantly by research on the human genome and by the advent of genetic technologies in medicine, which can provide valuable diagnostic and prognostic information about a person, family or ethnic community.

This increasing use of body tissue as an information resource in the grand projects of contemporary biomedicine presupposes views about the moral status of the human body, of its parts and tissues individually and in aggregate, and of conceptions of self-identity. These biomedical assumptions may be compatible with, or in conflict with, philosophical and religious perspectives, as well as cultural and personal attitudes. This article will reflect on some major issues regarding the moral status of body tissues, beginning with an examination of the relation between the self, the body, and excised body tissue. Analysis and ethical reflection on the status of body tissue requires attention to prominent metaphorical understandings in scientific and research discourse, and provides a basis for ethical transfers of body tissue from patient to researcher. However, important questions must be addressed regarding the limits of use of body tissue in order to avoid abuse.

BODY AND SELF

Attitudes Toward the Body

Moral reflection on the status of body tissue must first begin by considering tissue in relation to the organic bodily whole and to the self. Bioethicist William F. May has distinguished five philosophical attitudes on the status of the body and its parts that are useful to review in the context of retrieval of human tissue (2):

Dualism — Gnosticism — Holism — Materialism — Idealism

May differentiates these attitudes based on questions regarding whether the body has a phenomenal reality, is ontologically good, and is intrinsic or incidental to personal identity.

Dualism. Dualism affirms the phenomenal reality of the body but denies its goodness because of the body's association with flesh and matter. Dualism portrays the body as "at war" with the self, its literal mortal enemy. Parts of the body, disposal of the body, and persons who come into contact with the corpse are denigrated, stigmatized, and considered taboo and sources of pollution or uncleanness.

Idealism. At the opposite end of the continuum, idealism denies that bodily life has any ultimate significance. The body, disease, and death are constructs of the mind that can be transcended through identification with a separate realm of the spiritual. The self seeks its true home in this noncorporeal realm of spirit. Idealism supports practices of spiritual healing rather than medical ministrations in response to disease. Body parts, or study of the body, are viewed as existentially indifferent within the idealistic worldview.

Gnosticism. The gnostic attitude seeks liberation of the true self from the body through knowledge. The true self resides in a disembodied mind or consciousness, while the body is a prison of the soul. Liberation entails overcoming the burdens of mortality, including finitude, disease and death. Body parts have no significant value, and may even have disvalue, insofar as a preoccupation with bodily life can hinder the quest for liberation.

In contemporary Western thought, May argues that the gnostic attitude has assumed the form of philosophical Cartesianism. Philosophical and moral emphasis is placed on the mind or consciousness, distinct from the body. The body (and body tissue) is a possession and instrument of the rational will, useful as a tool to achieve the self's own goals and aims, but assumes no independent value of its own. Except in those circumstances where the body may impinge on and frustrate the plans and desires of the rational self (e.g., in illness), the body and constituent body tissues can be assessed as morally indifferent.

The Cartesian separation of self (mind) and body (matter) is embedded within the ideology of biomedical research; the body may be perceived merely as a resource for obtaining raw biological materials from which information and knowledge can be retrieved. In turn, this knowledge can be converted into technologies, therapies, and pharmaceutics that enable mastery over nature (3). The separation of self and body would also seem to imply there is no moral necessity for informed consent to the removal of bodily tissue, or subsequent research and manipulation. Even though bodily space is violated by such procedures, since a body is incidental to personal identity, the retrieval of an organ, tissue, or cell cannot be said to violate the person and their integrity. In this regard informed consent appears to be a rule without a rationale. However, the increasing importance of informed consent procedures in the procurement of human tissue for research (4) is a signal of a fundamental inadequacy of the Cartesian/gnostic perspective.

Materialism. In contrast to a knowledge-based quest for philosophic liberation and biomedical mastery of nature, the materialistic attitude to the body is informed by an ideology that human life is subject to, and at the mercy of, powers in nature that are arbitrary, abusive, and destructive. This can lead to two different and conflicting perspectives, avoidance and denial (of aging, death, etc.), or resistance, both of which resonate in contemporary Western cultures, including biomedicine. The perspective of resistance is enacted primarily through the practice of medicine and its war on death and human disease (5). In the medical effort to resist the inevitable natural course of bodily morbidity and mortality, the body assumes the role of primary "battleground." Patient consent to invasive procedures (an "invasion" not only of the body, but against nature) still leaves the patient as principally a passive observer to the battle plan carried out by physicians. A successful war often requires excision of body parts and removal of tissue. This body tissue is not only surplus but also is a valuable diagnostic and prognostic information source.

Holism. Within the typology constructed by May, the dualistic, gnostic, and materialistic attitudes are compatible with the interests of biomedical research in human tissue. The idealistic attitude, by contrast, finds in medical research a misguided attempt at medicalization of the metaphysical. This implies that scientific proposals for using human tissue samples are an *ethical* issue primarily in those traditions — philosophical, religious, and medical — that affirm a holistic attitude of human embodiment. Holism expresses an intrinsic relationship of body and self, in contrast to the instrumental value embedded in Cartesian and materialist thought. Holistic traditions claim for the body an ontological reality of the body (in contrast to idealism) and an intrinsic moral goodness (in contrast to dualism). Moreover the traditions of holism affirm the intrinsic nature of the body to personal identity: Human beings are embodied selves, not simply a soul or a mind housed within a corporeal prison.

The holistic attitude thereby supports a presumption in favor of bodily integrity and intactness (even beyond death in some cultural and religious traditions). However, this presumption can be overridden for purposes of preserving the health of the person or for providing direct therapeutic benefits to others, as in transplantation. The reality and goodness of the body entail the use of medical procedures to restore and heal. The fundamental connection of body and self makes consent of the person a moral mandate with respect to invasive medical procedures and removal of bodily tissue, healthy or diseased. Given the integration of body and self, holism places significant emphasis on accountability for uses of the body and an orientation of uses of excised body tissue toward the common good. Ethical positions and liturgical rituals, for example, justify sharing of the body and its constituent organs and tissues as a form of altruistic service to others. Thus an understanding of body tissues as "gifts" is prominent in discourse with respect to organ transplants or transfusions of vital tissues, such as blood or bone marrow (6,7).

A more circumspect attitude is characteristic of holism toward research or educational uses of body tissue. Until the onset of the era of genetic and molecular medicine in the mid-1980s, holistic discourse largely followed the lead of technological capacities and examined the status of human organs and tissues with respect to donation for therapeutic goals. Despite ongoing collection and analysis of pathological tissue specimens, very little ethical discussion was devoted to nontherapeutic research uses of human tissue, other than questions pertaining to autopsy procedures.

Conflicts: Bodily Integrity and Disintegration

Although these five attitudes about the body are compelling as basic or ideal types, none can fully accommodate the complete range of bodily experience nor the new requests for body tissue precipitated by emerging biomedical technologies. Holism requires modification to accommodate scientific interest in body tissue, particularly those tissues that are not core to the conception of self. Similarly attitudes congruent with scientific interests — dualism, gnosticism, materialism — must be adapted to account for human experience that displays that the self is more than the mind, and that the body, and at least some bodily tissues, are not mere instruments but are morally and symbolically significant.

For example, empirical studies suggest that it is not only the body as a whole but also body parts, organs, and tissues that can be formative of self-identity (8). Visible parts of the body, such as skin, genitals, fingers, hands, legs, and eyes, as well as the heart, have a strong correlation with a sense of self. With the exception of the heart, nonvisible organs and tissues are not as strongly incorporated into a sense of self. Thus not all body parts possess equal ontic status, or are equally important to self-identity. It is possible, as Belk (8,9) suggests, that the less an organ or tissue is connected with a sense of self-identity, the more willing a person will be to donate it for use by others or to have it retrieved for purposes of a scientific research study.

Despite this possible convergence of scientific and nonscientific attitudes on the body regarding the specific question of the status of body tissue, an important question about the starting point for inquiry is still unresolved. Philosophical or religious claims about the ethical significance of bodily integrity can conflict with scientific and research interest in parts of the body, that is, the body as a disintegrated entity.

While the human body as an organic totality has long been the subject of philosophical reflection and a symbol of political (Plato's "Republic") or religious ("the body of Christ") communities, the interest of medical science often begins with a micro unit as its point of inquiry, the "building blocks of life," or DNA. DNA constitutes the fundamental components of life, with stem cells and somatic cells, tissues, organs, bodily systems, and the like, viewed as more complex, functional expressions of basic genetic materials. The scientific value of the body as a totality is thereby a means to the goal of deciphering the codes, messages, and functions of the genetic building blocks that contain valuable information.

In some circumstances these different starting points will generate quite different ethical conclusions about appropriate interventions, manipulations, and excisions of the body and its tissues. Legal philosopher E. Richard Gold cites the "disparate claims of scientific investigation and religious belief" as the exemplary case of incommensurate values regarding the body. According to Gold, "The body, from a scientific viewpoint, is a source of knowledge of physical development, aging, and disease. From a religious perspective, the body is understood as a sacred object, being created in the image of God. . . .The scientist values the body instrumentally, as a means to acquire knowledge; the believer values the body intrinsically, for being an image of God" (10).

The Mechanistic Body

Perhaps the most prominent discursive scientific understanding of the body, which has its roots in the intellectual Enlightenment, the industrial revolution, and Cartesian science, is that of a "machine." Among others of his generation, Thomas Jefferson clearly expressed this understanding toward the end of his life: "But our machines have now been running seventy or eighty years, and we must expect

that, worn as they are, here a pivot, there a wheel, now a pinion, next a spring, will be giving way; and however we may tinker them up for a while, all will at length surcease motion" (10). The role of biomedicine becomes very clear in the mechanistic understanding of the body: to "fix" the worn-out parts through technical acumen and knowledge, perhaps to replace others through transplantation, and to improve on still others through research and engineering. Research on body tissue becomes an important gauge by which to determine the various developmental and functional processes of cells and tissues, as well as providing a means to identify what has gone wrong in case of malfunction.

Clearly, a valuation of the body that is mechanistic and instrumentalist will permit more research interventions and manipulations than will a position that affirms the intrinsic value of the body and its inextricable intertwining with the self. Moreover tissues retrieved from a mechanized body may elicit sentiments of revulsion rather than respect. The distinction between the status of organs, tissues, and fluids when they are incorporated within the body rather than outside and separate from the body is displayed in a memorable illustration by psychologist Gordon Allport: "Think first of swallowing the saliva in your mouth, or do so. Then imagine expectorating it into a tumbler and drinking it! What seemed natural and 'mine' suddenly becomes disgusting and alien. . . .What I perceive as separate from my body becomes, in the twinkling of an eye, cold and foreign" (12).

The emphasis on bodily integrity in the western religious faiths has culminated in the development of stigmas and taboos regarding certain bodily tissues when they are external rather than internal to the body (12). Characteristically the dis-incorporation of bodily tissues is assessed by religious thought with reference to issues of purity and cleanliness. A very prominent historical illustration of purity, which has permeated secular culture and has not been entirely overcome in contemporary religious communities, are stigmas and taboos surrounding menstruation (13).

There are then different ways of assessing the moral status of the body and of body organs and tissues depending on the "place" of the organs or tissues, that is (*1*) intrinsic to self-identity (e.g., heart) or incidental, (*2*) visible (eyes, skin) or hidden (kidney), and (*3*) integrated (circulating blood) or dis-incorporated (bodily excretions). In general, it may be claimed that the more an organ, tissue, or fluid possesses intrinsic connection, visibility, and integration, the more its retrieval and use for biomedical research purposes may present ethical questions. Put another way, philosophical and religious thought on the body begins with a strong presumption that the status of the body as a whole is greater than the sum of its parts. Body organs and tissues contain potent symbolic significance when considered as part of the bodily whole. Yet, when considered in isolation from the rest of the body, organs and tissue may generate revulsion and stigmas. This ambivalence can be heightened or ameliorated by scientific research that seeks to use body tissue for any number of research, therapeutic, or commercial purposes.

These perspectives and conflicts must be faced directly in the face of, and in anticipation of, technological advances in the realms of genetics and reproduction. Such advances have vastly expanded the biomedical gaze and the scope of scientific and medical interest in the body and its tissues. There is now virtually no part of the body that may not be the subject of scientific scrutiny. Genetic analysis in particular offers the prospect of gaining information about human character traits and behaviors, including susceptibilities to illness and bodily responses to disease, through study, analysis, and understanding DNA. Such information is not otherwise possible through an examination of the bodily totality. An assessment of body tissue as it is related to, and separated from, the body is inescapable in the context of biomedicine and biotechnology.

UNDERSTANDINGS OF BODY TISSUE

Metaphors of Body Tissue

Assessing the moral status of disincorporated body tissue can be illuminated by analysis of important metaphors that emerge in scientific and ethical discourse. These metaphors function to provide an understanding of something relatively unfamiliar, such as the significance of DNA or body cells, in terms of something more familiar and concrete.

Library. A central metaphor in discourse about the human genome project, for example, is that of a "book" or "library." This metaphor highlights the "information" stored in genes and cells, which scientists seek to retrieve and interpret through genetic analysis. Although information itself is often said to be value free, the process of determining what counts as information and what is "junk DNA" is clearly value laden. Indeed, as Rosner and Johnson observe, "By the choices [the genome project] makes — the choices of what books to include in the library and in what condition — the Project will determine what is 'correct,' what is 'real.' It will necessarily set standards, defining and cataloging what it means to be human, limiting what range of diversity is acceptable" (14).

The library metaphor highlights the extraction of information as the primary feature in research on body tissue. This implies that bodily tissue has at best instrumental moral status. What is concealed by the metaphor, however, is that unlike the information contained in a library, which is generally publicly accessible, the information that is retrieved from genomic studies is highly personal and sensitive. The information stored in any DNA library can be interpreted in ways that hold out implications for personal identity, including comparisons to a collective group standard or prospects of personal risks of disease susceptibility.

The information itself is of moral significance and should thus be treated in a manner equivalent to other forms of medical information. As with any information that is personally identifiable, there are legitimate concerns about the uses and potential for abuses of genetic information. Thus, in some jurisdictions, regulations

and laws have been passed to ensure privacy and confidentiality of genomic information and to protect persons from discrimination.

For example, the Genetic Privacy Act passed in 1995 in Oregon affirms: "The DNA molecule contains information about an individual's probable medical future. . . .Genetic information is uniquely private and personal information that should not be collected, retained, or disclosed without the individual's authorization." While employers or insurance companies may, with consent, have access to a person's genetic information, the law prohibits the use of such information for purposes of hiring or setting insurance rates (15).

Resource. A second prominent metaphorical understanding of human body tissue is that of a "resource," most commonly, a resource that has been removed from its "natural" integrated state in the body and is thus accessible for a variety of scientific manipulations and purposes. More specifically, body tissue is portrayed in terms of a valuable mineral resource: Genetic analysis is described as a "gold rush," umbilical cords are referred to as "clinical gold," a patient's body as a "gold mine," and the researchers themselves assume the role of "gene prospectors" (16).

Ultimately the image behind these metaphors is that of the body as the "new frontier" encountered by cutting-edge science, a wilderness that biomedical science seeks to map and mine and ultimately to master and domesticate. Exploring the frontier requires and rewards individual initiative, thus it is not surprising that aspects of the resource metaphor move in the direction of commercial and entrepreneurial exploitation of body tissue. As Andrews and Nelkin express it, "Body parts are extracted like a mineral, harvested like a crop, or mined like a resource" (16).

The frontier or wilderness is also a realm beyond moral control; thus, as the frontier begins to be explored, and resources extracted, rules must be developed to facilitate a coordinated discovery effort. If the metaphor follows historical patterns, however, then there is ethical reason for concern: If the human body of the early twenty-first century is treated in terms analogous to the land and mineral rushes of the nineteenth century, short-term benefits may be obtained at the expense of harms that befall future generations. When the bodily frontier is viewed as an exploitable natural resource of immense scientific interest, a question arises about the extent to which researchers will take steps that respect and ensure the integrity of the whole.

The resource metaphor of body tissue is distinct from but certainly compatible with the book or library metaphor. Indeed, body tissue is a resource *precisely* because it contains information. The purpose of procuring tissues for research is to generate generalizable knowledge, for example, advancing researchers' understanding of the genetic dimensions of human diseases. However, this research has yet to yield significant promise of therapeutic benefits to persons afflicted with those diseases. Similarly investors in a biotech company engaged in genetic research may have to stay with their investment over the long haul until the research bears fruit in a commercially marketable product. This distinction between basic research that provides information or enhanced diagnostic capabilities in the short term but defers therapeutic (and entrepreneurial) benefits to the distant future is one point of criticism in ethical assessments of genetics research (17).

Property. Both the library and the resource metaphor are inadequate in at least one morally essential respect: They overlook authorization from the person from whom the body tissues are retrieved. The genetic library is not a public library, but contains very intimate, private information, that can accessed only by a select few, and only then upon consent. Similarly a person's body tissue is not a resource commons. Thus, lest scientific explorations on the frontier of the body become unnecessarily invasive of privacy and trespass integrity, or body tissue be subject to the research equivalent of a public land grab, a third metaphor of body tissue has been promoted to address issues of control and authorization. This is the model of the body and of body tissues as "property" (18,19). The property understanding stems from a claim of self-ownership, and it seeks to authorize the individual person, patient, or research subject, rather than the scientific prospector, with control over the use and disposition of the body and of body parts.

Not unlike some of the scientific attitudes about the body discussed previously, the property metaphor tends to treat the body as incidental rather than intrinsic to personal identity. The philosophical assumption is that the body even as a totality is distinct from the self, and body organs and tissues can therefore be transferred or alienated to others without compromising the integrity of the self. Thus, with the proviso that informed consent is obtained from the person, these assumptions make the property perspective very conducive to scientific interest in body tissue. However, conflict can arise when, for example, a patient and a researcher assert competing claims or "property rights" to excised body tissues, or to cell lines that have been isolated from retrieved tissue.

The Moore Case

The most prominent example of such conflict involved the case of John Moore, who sought treatment for a rare disease, hairy-cell leukemia (20). Moore's physician recommended a splenectomy, which was performed with Moore's consent. However, the physician recognized early into the course of treatment that Moore's body was overproducing lymphokines, an important component of the human immune system. Samples of Moore's blood, skin, bone marrow, and sperm were then retrieved in follow-up visits. Cells from these tissues were cultured and altered to create an immortal cell line, which was patented and commercially developed, without Moore's knowledge. Upon learning of this activity, Moore brought a lawsuit, alleging that he retained ownership of his body tissues and cells and that the research had wrongfully used his "property" to benefit others. Moore's proprietary interests in the removed body tissue and cells were rejected by the Supreme Court of California, although

the court did find he had a compensatory claim if he had not been informed by the physician of the prospective research use and commercial development of the tissue. The majority opinion rejected Moore's claim on the basis that granting Moore property rights would impede scientific progress. A decision in favor of Moore, it was claimed, would open the door to actionable claims by patients any time body tissue was retrieved for research purposes. Moreover pharmaceutical companies would be deterred from investing capital and energy in new product development for fear of litigation.

The Moore case throws into sharp relief many questions about ownership and control embedded in any situation where body tissue is removed from a patient. There is no dispute that, when the tissues and cells are functioning as integrated parts of the intact bodily organism, they are under the authority and control of the patient. However, does the patient relinquish any and all claims and interests to the tissue once it is excised from the body? Do physicians or researchers gain rights of ownership to the tissues that are removed by their labor? Such questions necessitate attention to various modes of transfer and acquisition of body tissue.

TRANSFER AND ACQUISITION

The property, library, and resource metaphors assume the validity of the use of body tissue, so long as consent is voluntary and informed. Consent provides authorization for the transfer of body tissue from a patient or subject to a researcher. However, consent is only a necessary, rather than a sufficient, condition of morally justifiable transfers. Questions must still be addressed concerning the limits to be placed on such transfers and on avoiding abuse in research on acquired tissue.

Modes of Transfer

An analogy can usefully illustrate the moral context of these issues. I will adapt Belk's observation that "the house is a symbolic body for the family" (8,9), and construct an analogy wherein household goods take the place of human tissue samples. The relevant question is the manner of transferring household goods. I want to focus on four modes of transfer, all of which have correlates in methods of transfer of bodily tissue (21).

Donation. One transfer method is to donate certain goods such as clothing to a community goodwill program. This presents an example of a gift or an altruistic action designed to benefit others and to enhance a recipient's quality of life. Similarly various life-saving organs and tissues, such as blood, have been referred to as "gifts of life" in popular discourse.

Contribution. A second form of transfer concerns those household goods whose designed use has been depleted by members of the household, for example, food products that come in plastic or cardboard containers or newspapers. These containers and other materials could still be used for some household functions, such as for storage or for starting a fire; they are not without household utility. It is also possible, however, to contribute these materials for a greater good, the benefit of the community through recycling. An important difference between donation and a contribution is that the donated goods are typically transferred and used in their original form, while a contributed good, be it a household container or body tissue, undergoes some alteration before re-use.

The recycling approach requires communal support of organizations with the knowledge and expertise to convert the recycled materials into valued products. Similarly a biomedical research institution may retrieve certain tissues contributed for the greater good of the advancement of scientific knowledge. Researchers can then subject the tissue to a research process in which cells are isolated to yield a valuable immortal cell line or genes analyzed to provide information.

Abandonment. A different set of household materials are those goods that have been consumed completely, and whose benefits have been exhausted by household members. The packaging protecting these materials is now "waste" and is commonly discarded through a community utility such as a trash collection service. Household refuse has no personal meaning to the discarder, who is typically quite willing to abandon the refuse and even pay a fee to have the items removed. This does not, of course, preclude the possibility that this refuse might have value to someone else who is willing to take the time to sort through the materials. Similarly at least some bodily tissues are referred to as "waste" or "garbage" in biomedical literature (22,23), even though research processes may transform the tissues into a beneficial and profitable product.

Property Sale. A fourth form involves transferring the household goods to others through an informal market transaction, the "garage" or "yard" sale (or through classified advertising). In these situations the original owner hopes to obtain some financial gain through the transfer, although the expectation is that neither the purchase price, nor the quality of the goods, will be as high as if the purchaser had obtained a similar item through an established business. These informal transactions occur parallel to, but outside of, formal and regulated market mechanisms and controls; once the transaction has taken place, there is no expectation of a continued buyer–seller relationship. Procurement of some body tissue does occur through buying and selling, although typically the seller remains anonymous.

It may be the case that all of these modes of transfer, and their biomedical analogies, have a role in considering ethical issues in research acquisition of body tissues. As suggested, there are precedents already for tissue as a donated gift, as a research recyclable, as abandoned refuse, and as a commodity. However, certain tissue may be considered more or less appropriate to a particular form of transfer.

This differentiation may, in part, rely on a sense of connectedness to self-identity delineated above. Some body parts, such as the heart, eyes, or blood, may have such

symbolic significance and connection to personal identity that their donation is the moral equivalent of a gift of self. Donation or contribution imagery may seem more fitting with the transfer of such tissues than abandonment or selling. Other body tissues, for example, urine or hair clippings, may have such minimal value to a sense of self that they are considered "waste" and routinely abandoned. Still other organs and tissues, such as a pancreas, liver, spleen, or marrow, may fall in between these examples, not as central to personal identity as the heart or eyes but not as incidental as urine either. Still other tissue, such as gametes or blood, may be considered so vital to the creation of life or to the maintenance of life that society is willing to permit a limited market to function to procure these tissues.

Of the above modes of transfer, gifts or donations have a prominence in discourse relative to contribution, abandonment, or sales. The discussion below will explain this prominence through examining the features of what may be designated the "donation paradigm," and it then will assess the extent to which methods of contribution, abandonment, and selling retain or violate these features.

The Donation Paradigm

Ethical thought on the body and its use within medicine has presupposed a context within which organs and tissues are donated for therapeutic purposes of healing, restoring, or saving life. This moral ideal is emphasized through the language of "gift," "altruism," "sacrifice," and so on, on the part of the donor and that of "benefits" for recipients. The donation paradigm seems constructed by four principal features:

Altruistic Intent. The intent of the donor of an organ or tissue is structured by gift-giving to beneficiaries or recipients, such as persons on a waiting list for a transplant (although the identity of such persons may be veiled from the donor).

Therapeutic Expectation. The expected outcome of a gift of the body is that it will be used in such a way as to offer a pronounced therapeutic prospect for the recipient. These prospects encompass both enhancing quality of life and preserving life.

Generative Re-incorporation. Body tissue that has been retrieved from the donor, or dis-incorporated, should in most circumstances be re-incorporated within the body of the recipient. As noted above, tissue that is dis-incorporated may evoke sentiments of revulsion and practices of stigma and taboo. Some religious practices and rituals require burial of removed body parts, or re-incorporation in the earth. This is particularly the case with body parts that have an identifiable human form: In Jewish thought body parts such as limbs, composed of "flesh, sinew, and bones," should under most circumstances be buried. Roman Catholic tradition distinguishes major from minor parts of the body in a similar manner. Major parts of the body are those, such as a limb, that retain their "human quality" following excision and should be buried (24). Such concerns reflect in part the importance of these visible body parts for self-identity.

Re-incorporation of body organs or tissues in a human recipient has a generative power in that it offers the prospect of new or renewed life to the recipient. In general, then, the donation paradigm prioritizes practices in which body tissue remains with a body (even if transferred and transplanted to the body of another person) and thus symbolizes the significance of bodily integrity and holism.

Recipient Responsibilities. The gift of body tissue carries with it certain responsibilities on the part of the recipient, which are embedded in everyday practices of sharing and gift-giving (25,26). These responsibilities include a sentiment of gratitude toward the gift-giver, or toward the institutional structure that mediates the gift transfer. Gratitude may also be enacted in the actions and conduct of the recipient by which he or she makes grateful use of the gift. In addition a gift induces a responsibility of reciprocity. Reciprocity does not necessarily mean the continuation of the gift relationship between the initial giver and recipient; rather, a recipient of donated blood, for example, may at some time in the future become a donor for other strangers.

The donation paradigm provides a moral justification for medical use of human body tissue. It is especially prominent in bioethics discourse because it highlights simultaneously moral virtue and altruism, and medical success, the restoration of life to the nearly dead. The paradigm is limited, however, for the most part to medical practices of transplantation or transfusion, that is, those practices that promise some form of therapeutic outcome from the gift. However, different models seem necessary to accommodate biomedical retrieval and use of body tissue in circumstances where nontherapeutic uses of body tissue are contemplated (e.g., tissue contributed for research and educational purposes) or where altruistic intent is absent (e.g., abandonment and sales).

Contributions

The idea of body tissue as a contribution shares with the donation paradigm the intention of gift-giving or altruism. However, the anticipated outcomes are rather different. The goal in research or educational uses of human tissue is to advance scientific learning and to generate knowledge that is generalizable, not to furnish a prospective therapy for an individual recipient (although certainly some research progresses in the direction of therapeutic outcomes). This is not an incidental issue, for some cultural and religious traditions that are generally supportive of donation for therapeutic goals find it much more difficult to provide a justification for biomedical research undertaken without therapeutic intent. In examining the ethical status of tissue contributions, we will rely on the fourfold features of the donation paradigm to illuminate distinctions and comparisons.

Intent. A person who places their recyclable household materials at the curbside is engaging in nonobligatory gift-giving in an important respect because they could well retain the materials for their own purposes, such as

storage, collections, or fire-starters. While they are not making a personalized gift to a specific individual, they do contribute to a cause that is larger than themselves and the benefits they might provide in a direct or mediated relationship with another person in need. The cause in the domestic case may be "environmental preservation"; in the case of body tissue, it may be designated as "scientific discovery" or "medical progress." The contribution in both cases is one of nonspecific generosity; it is "nonspecific" in that a "cause," rather than a person, is the intended beneficiary. The contribution is also a generous act in that a person is participating in the advancement of the larger cause when they could just as easily place the recyclable material in the refuse bin without moral blame, or request abandonment or return of the body tissue.

Contributor Expectation. Unlike the therapeutic expectation embedded in the donation paradigm, a contribution of body tissue for research or educational purposes does not bring direct and immediate benefits to a specific or designated individual. To be sure, there is an expectation of benefits on the part of the contributor but these are much more diffuse in space and remote over time. The recipient of the tissue contribution, a science researcher or medical educator, may develop protocols or pedagogies that over time will accrue benefits to the larger good of society, or at least to those persons with a stake in biomedical research or education. Perhaps the tissue will turn out to have "immortal" features, either through its cell progeny or through laboratory learning that is transferred to clinical practice. Still the objective in contributions is advancing knowledge, not therapy; learning, not curing. The good of the greater cause, not an individual in need, predominates.

This difference in expectation marks a distinction between donations and contributions of body tissue. There are also marked differences between contribution and abandonment or sales. A willingness to contribute does not imply that the contribution has minimal or no value to the contributor. A plastic milk container that can be recycled is equally serviceable as a water jug. Similarly a person may attribute value to many tissue specimens, including blood, reproductive tissues, skin or hair that have been retrieved or excised from the body. The prime difference of contributions lies in the fact that something of value is contributed to a person or organization through whose work the society realizes a benefit that is greater than would have been the case had the contributor decided to retain the material. A contribution thereby intends a benefit for the common good of all.

Symbolic Re-incorporation. The donation paradigm involves re-incorporating removed tissue or organs into another body, whether an organ transplant or blood transfusion into a person, or as is common among some religious rituals, burial in the earth. Re-incorporation practices are not literally possible in contributions, because the research and educational purposes necessitate analysis or study of body tissue in isolation from the body totality. However, a form of symbolic re-incorporation is possible with contributions of body tissue. For example, just as recycling contributes to the good of the larger communal whole, the contribution of body tissues for research can provide information that can then be integrated within, or require revisions to, a larger, symbolic whole, designated as the "body of scientific knowledge."

Recipient Responsibilities. Contributions in general are acknowledged in some form by the recipient, but implementing an acknowledgment is more difficult with respect to contributions of bodily tissue. It is important that those who seek informed consent for nontherapeutic uses of body tissue not presume that a patient or subject will simply "sign off" to any and all uses made of retrieved body tissues.

However, since the moral point of the contribution is to facilitate the achievement of some cause greater than the interests of either tissue source or recipient, the researcher should assume a role of trustee and steward for the community. This entails a research responsibility to use the contribution of body tissue for the common good. At a minimum this requires treating the information generated by tissue research with safeguards that ensure protection of the contributor against discrimination or harm.

The appeal to the "common good" does not preclude recourse to the private sector to carry out research; in some cases, as with domestic recyclables, the good of all can be more efficiently and effectively achieved through private sector initiatives. However, in a contribution approach, profit interests must be subordinated to the common good and the greater cause that the contribution is designed to advance. In short, retrieved body tissue is a source of good, and not merely a resource for financial gain. Body tissue that is contributed for purposes of the general cause of advancing science should not be viewed as merely an economic asset. Such limitations may not apply, however, in situations of acquisition of body tissue through abandonment or sale.

Abandonment

Many of the 282 million specimens of human biological materials stored in labs in the United States have been acquired following abandonment by the patient or research subject. A person who has undergone a surgical procedure during which certain body tissue has been removed is characteristically not interested in retaining the tissue themselves; the removed tissue assumes the status of "surplus" or bodily "refuse" that the patient is quite willing to discard, especially if the tissue is implicated in a disease condition. In some cases (though not all) the person may sign a general consent form permitting research or educational uses of their tissue. This tissue is a resource for scientific study and exploitation, but this permission is an attenuated form of donative or contributory intent.

In other circumstances, as with the placenta or umbilical cord, the removed tissue may not have been a source of disease but a sustainer of life that is no longer biologically needed. Nonetheless, for some persons, as well as cultural traditions such as Native American, these tissues are profound symbols of life and relationships, and the patient may retain them. For many others, however, they are surplus membranes that can be discarded following birth; even if research consent is given, the primary intent is to be rid of the tissue.

This suggests departures in several respects from the features of the moral model that inform the donation paradigm, and with some modifications, contributions of tissue. If a donative intent is present at all, it is relatively insignificant and possesses secondary importance. There is also likely to be no expectation of a therapeutic outcome and minimal interest in any research or educational uses. Are efforts at re-incorporation attempted? As Murray writes, in most cases of the removal of a diseased organ, "the tissue is disposed of, usually by incineration." Yet this does not relieve pathologists or researchers of recipient responsibilities, including respect and dignity appropriate to "the fact that the tissue is from a human" (7). Moreover, since research analysis could establish linkage to the tissue source, considerations of privacy and confidentiality continue to apply.

Sales

Certain body tissues, though not solid organs, are allowed to be bought and sold. These include blood, skin, and reproductive tissue. Financial transactions for body tissue stretch the language of "donor" beyond all reasonable usage. The language of tissue "source," "seller," or "vendor" is more appropriate. In sales of body tissue the altruistic intent of the donation paradigm has been replaced by the commercial prospects of property (27). The property understanding of the body, and its correlate, an understanding that excised or retrieved body tissue is, potentially, a "commodity," recognizes the ownership rights of the person to his or her own body and body tissue. As with any other form of property, these rights can be transferred to, and acquired by others, including researchers or tissue banks through contractual agreement.

A contract that transfers property rights in body tissue means that a person's interests in their body tissue are effectively ended with the transaction. The tissue source may harbor expectations for a successful fertility outcome for the buyer, or some therapeutic developments by researchers, but that is not the primary rationale for the commodity exchange of tissue for money. If it were, then the tissue source should be equally satisfied with a donation or contribution of body tissue. Rather, the point is to use a shared procedure, the market mechanisms that function to transfer goods and services in various social realms, to facilitate mutual self-interest. Once the transaction is completed, the new property holder possesses rights to use and dispose of the commodity at his or her discretion.

Within the transaction, the social identity of the researcher-recipient of body tissue marks a striking contrast to that found in the donation paradigm or in contribution. The latter approaches present the recipient as a trustee and steward of what are communal resources, with a responsibility to use the resources as consonant with the common good. No such strictures apply in a commodity exchange of body tissue. The researcher functions more along the lines of a private entrepreneur, seeking to protect personal property as both a scientific and an economic investment. In some cases this protection is secured through patenting of an innovation, such as an immortal cell line derived from body tissue. For example, in the case of John Moore, the physician entered into contracts with pharmaceutical companies to enable successful exploitation of the estimated $3 billion potential of the cell line. A patent on the cell line and its derived products was secured and transferred to the research university (9).

The property paradigm thus has a very attenuated view of recipient responsibilities; those responsibilities are primarily directed toward advancing the interests of the recipient rather than shared societal interests. The researcher is not bound by responsibilities of gratitude or reciprocity because the work of the scientific community has made possible the retrieval of the tissue in the first place. Were it not for the initial labor of physicians and researchers, there would be neither donations nor sales of body tissue. The main responsibility of the researcher-recipient thus appears to be ensuring that informed consent takes place.

While the donation paradigm has few critics, other than those who may fault it for providing an inadequate supply of organs and body tissue, the property paradigm of the body, and the idea of tissue as a purchasable commodity, has received criticism similar to that directed at (so far, unsuccessful) proposals for a commercial market in solid organs. These criticisms include questions about the impact of financial incentives on voluntary, informed consent; "quality assurance" mechanisms to ensure safety; and the potential for the property paradigm to compete with, and ultimately erode, communal support for donation. Underlying these procedural and consequentialist concerns is a substantive criticism that commodification of body tissue, and understanding the body as personal property, ultimately is an affront to human dignity. Proponents of the property paradigm have argued in response that it provides assurance of personal control over the disposal of body tissue and freedom of choice for the tissue source. A commercial market, it is claimed, can and should supplement gifts of the body; after all, property can be sold as well as donated. Moreover a commercial market could increase supplies of scarce body tissue.

One common concern is that the effort to increase supplies of tissue for either therapy and research through permitting sales risks undue influence and coercion of the decision maker. Thus laws prohibit women from selling tissue from their aborted fetuses, and regulations are currently being discussed that would impose similar restrictions on couples in vitro fertilization (IVF) programs who are informed about the possibility of research on stem cells derived from surplus embryos. In other circumstances, however, such as procurement of gametes, financial incentives are not necessarily deemed as compromising the voluntariness of a choice. With respect to gamete procurement, the voluntary guidelines promulgated by various professional societies that discourage excessive financial inducements seem to have gained general acceptance.

CONCLUSION

Innovative biomedical research, especially in the field of molecular genetics, has heightened scientific interest in tissue samples, and thereby posed new questions

about the moral status of body tissue in research and educational settings. In a diverse and pluralistic culture, wherein many traditions, cultures, and communities vest body tissue with tremendous symbolic significance, the moral burden of proof falls on the scientific community and researchers to justify acquisition of body tissue. Researchers seeking access to body tissue must make a compelling case for its use, ensure that their scientific design will likely generate significant results, determine that there are no other alternatives to achieve the research objectives, facilitate a process of informed consent with tissue sources, and provide guarantees of confidentiality and anonymity as appropriate. In this way body tissue can legitimately be used while minimizing abuses.

BIBLIOGRAPHY

1. National Bioethics Advisory Commission, *The Use of Human Biological Materials in Research: Ethical Issues and Policy Guidance*, Rockville, MD, 1999.
2. W.F. May, *The Patient's Ordeal*, Indiana University Press, Bloomington, IN, 1991.
3. C.S. Campbell, in L.S. Cahill and M.A. Farley, eds., *Embodiment, Morality and Medicine*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995, pp. 169–183.
4. National Bioethics Advisory Commission, *The Use of Human Biological Materials in Research: Ethical Issues and Policy Guidance*, Rockville, MD, 1999.
5. J.F. Childress, *Practical Reasoning in Bioethics*, Indiana University Press, Bloomington, IN, 1997.
6. W.F. May, *Hastings Cent. Rep.* **15**, 38–42 (1985).
7. T.H. Murray, *IRB* **8**, 1–5 (1986).
8. R.W. Belk, in J. Shanteau and R.J. Harris, eds., *Organ Donation and Transplantation: Psychological and Behavioral Factors*, American Psychological Association, Washington, DC, 1990, pp. 139–149.
9. R.W. Belk, *J. Consumer Res.* **15**, 139–168 (1988).
10. E.R. Gold, *Body Parts: Property Rights and the Ownership of Human Biological Materials*, Georgetown University Press, Washington, DC, 1996.
11. S. Nuland, *How We Die: Reflections on Life's Final Chapter*, Knopf, New York, 1994.
12. G. Allport, *Becoming*, Yale University Press, New Haven, CT, 1955.
13. J. Lee and J. Sasser-Coen, *Blood Stories: Menarche and the Politics of the Female Body in Contemporary U.S. Society*, Routledge, New York, 1996.
14. M. Rosner and T.R. Johnson, *Hypatia* **10**, 104–129 (1995).
15. Legislative Assembly, State of Oregon, *1995 Oregon Revised Statutes*, vol. 10, Salem, OR, 1995 (ORS 659.700).
16. L. Andrews and D. Nelkin, *The Lancet* **351**, 53–57 (1998).
17. H. Bouma et al., *Christian Faith, Health, and Medical Practice*, W.B. Eerdmans Publishing, Grand Rapids, MI, 1989.
18. R. Scott, *The Body as Property*, Viking Press, New York, 1981.
19. L. Andrews, *Hastings Cent. Rep.* **16**, 28–38 (1986).
20. *Moore v. Regents of the University of California* (51 Cal. 3d 120 [1990]); Cf. E.R. Gold, *Body Parts: Property Rights and the Ownership of Human Biological Materials*, Georgetown University Press, Washington, DC, 1996.
21. C.S. Campbell, *Kennedy Inst. Ethics J.* **8**, 275–305 (1998).
22. G.J. Annas, *Hastings Cent. Rep.* **18**, 37–39 (1988).
23. National Bioethics Advisory Commission, *The Use of Human Biological Materials in Research: Ethical Issues and Policy Guidance*, Rockville, MD, 1999.
24. J.F. Childress, in D.A. Knight and P.J. Paris, eds., *Justice and the Holy*, Scholars Press, Atlanta, GA, 1989, pp. 215–240.
25. P. Camenisch, *J. Relig. Ethics* **9**, 1–34 (1981).
26. T.H. Murray, *Hastings Cent. Rep.* **17**, 30–38 (1987).
27. C.S. Campbell, *Hastings Cent. Rep.* **22**, 34–42 (1992).

See other entries GENETIC INFORMATION, ETHICS, ETHICAL ISSUES IN TISSUE BANKING AND HUMAN SUBJECT RESEARCH IN STORED TISSUES; HUMAN GENOME DIVERSITY PROJECT; OWNERSHIP OF HUMAN BIOLOGICAL MATERIAL; see also PATENTS AND LICENSING entries.

PATENTS AND LICENSING, ETHICS, ORGANIZATIONS WITH PROMINENT POSITIONS ON GENE PATENTING

R. COLE-TURNER
Pittsburgh Theological Seminary
Pittsburgh, Pennsylvania

OUTLINE

INTRODUCTION

One might think that the question of the patentablity of DNA and other biological components is narrow and technical, of interest only to lawyers, directors of research, venture capitalists, and a handful of scientists. In fact it is a question of great interest and concern to a wide range of organizations from every continent that for some reason feel compelled to express an opinion on this subject. And while the technical matters of patenting law and trade policy are often lost in the public debate, it is clear from the wide range of organizations with positions and from the intense passion with which they sometimes express their opinions that deeply held convictions are at stake in this question of the ownership and commercial use of biological components. Because patent law and trade policy are created through political processes, the conflict among organizations on biological patenting may have wide significance for the biotechnology industry for years to come.

The organizations with positions on patenting range across the entire spectrum of institutional styles and continuities of identity, from religious organizations, with

thousands of years of unbroken institutional identity, to conferences that meet briefly but leave a lasting impact. What counts as an organization? For the purposes of this summary, transient organizations, such as coalitions and conferences, are included, not only because their impact often outlives their structure but because of their galvanizing effect on their complex constituencies and on public consciousness. Some of these "organizations" are virtual; that is, they exist primarily if not exclusively on the Internet. Internet technology, so important to genome research and sequence publication, and thus partly determinative of what is patentable and what is not, is also used by individuals and communities worldwide to organize their opposition to biotechnology, to commercial use of genetics, and specifically to biological patenting. The impact of such "virtual organizations" and of Internet-based community organization should not be underestimated. Furthermore, precisely because of the fluid boundaries of organizations, "virtual organizations" can be remarkably successful in swaying the position of traditional organizations, such as religious bodies or governments or the United Nations.

The positions of these organizations on biological patenting also span the whole spectrum of possible views. Some of the organizations included here (mostly the environmental and indigenous peoples groups) oppose not only patenting but biotechnology itself and see opposition to patenting as a strategy toward their broader end. Others (mostly the religious organizations) support biotechnology, at least for many purposes, but oppose many if not all forms of biological patenting. Many others (e.g., scientific organizations) support biotechnology and therefore oppose certain forms of patenting on the grounds that they impede progress in the field. Others organizations (largely the industry groups) support broad patenting of genetic and biological discoveries, including (for some) short DNA sequences of unknown function.

The organizations that have take public positions on biological patenting may be grouped into seven categories, namely religious, indigenous peoples, environmentalist, public interest science, scientific, and industrial. Each category will be considered in the sections that follow.

RELIGIOUS ORGANIZATIONS

Of all the public statements made on biological patenting, perhaps the most widely reported is the May 1995 statement, the *Joint Appeal against Human and Animal Patenting*. The text of the statement, which was endorsed by the heads of over 80 religious organizations in nearly all faith traditions, consists of three sentences: "We, the undersigned religious leaders, oppose the patenting of human and animal life forms. We are disturbed by the U.S. Patent Office's recent decision to patent human body parts and several genetically engineered animals. We believe that humans and animals are creations of God, not humans, and as such should not be patented as human inventions" (1).

The Joint Appeal was organized by the General Board of Church and Society of the United Methodist Church and promoted in cooperation with the Foundation for Economic Trends, whose leader, Jeremy Rifkin, is a well-known activist in opposition to genetic technology. In a press statement that accompanied the release of the Appeal, the Rev. Kenneth L. Carder, Resident Bishop of the Nashville, Tennessee, Area, the United Methodist Church, argued for the Appeal on religious grounds: "The issue is the commodification of life and the reduction of life to its commercial value and marketability.... Life is a sacred gift from God the Creator. As a gift from God, life has intrinsic value. The patenting of genes, the building blocks of life, tends to reduce it to its economic worth.... Patenting further identifies life mechanistically and blurs the distinction between the animate and the inanimate.... The conflict is between reverence for life and exploitation of life, life valued for its marketability and life valued as an intrinsic gift. It is a conflict between utilitarianism and transcendent meaning. It is conflict between control based on economic profit and access based upon shared need" (2).

The 1995 Appeal was not the first time that religious leaders had joined to speak publicly on this issue. In 1987 a *Statement of Religious Leaders against Animal Patenting* was signed by seven national leaders of mainstream Protestantism and of Judaism and issued in response to U.S. Patent Office decision to allow patenting of transgenic animals. The statement objected to the decision, saying: "No decisions about the patenting of genetically altered animals should be made without the most careful examination of all possible consequences." The statement urges "support for congressional measures to halt such actions by the Patent Office, and to enable a careful consideration of these questions by the Congress and the public. The gift of life from God, in all its forms and species, should not be regard solely as if it were a chemical product subject to genetic alteration and patentable for economic benefit. Moral, social, and spiritual issues deserve far more serious consideration before binding decision are made in this area" (3). It should be noted that most of the religious leaders who signed the 1987 or the 1995 statements did so without their religious body having discussed a policy on patenting, much less having taken an official stance.

In these two statements, theological opposition to biological patenting is grounded on the belief that life is a gift from God, its creator and (one might say) inventor, who alone should possess any intellectual property rights in biological phenomena. The sheer act of patenting, the claim to possess a property right in the design of a biological organism, is an affront to this theological conviction. Furthermore patenting is an act that reduces an organism to mere matter and denies the special status that living things enjoy in relation to their creator. These two theological principles are the twin pillars upon which religious opposition to patenting is based. The first, that God is the creator, giver, and rightful source and definition of the value of the creation, leads in many of these statements immediately into the second principle, that God has so ordered creation that living things enjoy a special status that must not be violated or ignored. The first of these principles warns against the danger of anthropocentrism or the view that we (or any other creature) can define the value or purpose of creatures. The

second principle warns against reductionism, the view that all creation is nothing but matter, valueless apart from its usefulness to us.

It is the second of these twin principles that comes to the fore in positions taken by the Church of Scotland and found in a report of the European Ecumenical Commission for Church and Society (EECCS), a commission created by the Protestant, Anglican, and some Orthodox Churches to relate to the European Union and the Council of Europe. (It should be noted that the Church of Scotland has been especially strong in supporting the development of the positions taken by the EECCS.) The 1998 EECCS report, entitled "EECCS and Bioethics," holds that, "[c]oncerning patenting, it is a matter of ethical concern that commercial demands are now tending to abuse the normal distinctions between what is alive and what is not, and what is discovery and what is human invention. The mere knowledge of a gene should not be patentable in itself, nor should an entire transgenic organism — animal or plant — when it is only a tiny modified gene sequence that is novel. Moreover, there seems to be a real danger that genetically modified organisms are looked upon only as commodities in a global market place" (4, p. 5). The theological significance of the distinction between the living and the nonliving is also asserted in a major position paper of the World Council of Churches (WCC), whose participants come from every continent and nearly every tradition within Christianity. Among the concerns raised about biological patenting, the report states: "The US Supreme Court decision on patenting of life forms rested upon a specific, highly reductive conception of life, which sought to remove any distinction between living and non-living matter that could serve as an obstacle to the patenting of living but unnatural organisms. It would be easy to underestimate the philosophical, moral and ideological importance of the abolition of such a distinction, precisely because it allows a shift in accepted ideas as to what may be done to living things" (5, p. 18). It must be noted, however, that what is being opposed here is biological patenting, not biotechnology. In fact the statements are clear that they do not intend to prohibit biotechnology, even if they do raise concerns about some of its uses. But it is not clear in the statements why the living/nonliving distinction would prohibit patenting but would permit biotechnology. Nor is it clear just where, in biological phenomena, one should find the distinction between the living and the nonliving, or whether theology can offer any help in locating that distinction.

Religious opposition to some aspects of biological patenting is of course grounded in other theological and ethical concerns beyond the two central principles. Indeed, the WCC 1982 report lists "five causes for concern," of which the living/nonliving distinction is only one and which does not put "life as gift" on the list. Among the other four concerns, the report notes that "some argue that patent protection in this form will allow existing and emergent corporations to make excessive profits out of human suffering, to sustain an unhealthy worldwide dependence on pharmaceutical products, and to ignore real health needs because they offer lesser economic returns." In addition, due to inconsistencies and imprecision in patent law, there is a concern about "extremely costly patent litigation. This prospect is likely to cause major distortions in the process of innovation and to restrict the range of products made and the kinds of organizations making them" (5).

The report notes the concern that the patent system will lead "to a curtailment of communication and greater reluctance to exchange new research materials and to engage in collaborative projects." The final concern noted in the 1982 WCC report is that in view of the fact that "much of the research leading to commercially valuable knowledge was supported by public funds," patents leading to private profit will cheat the public twice. "Not only does this deplete the public exchequer but the needs of the community might be better served by arranging for the commercial exploitation of publicly funded research by publicly owned corporations." In particular, "churches and other institutions should press for alternative ways in which the labours of science are transmuted into the maximum human good." In light of these concerns, however, the report does not call for sweeping opposition to patenting but merely "[t]hat the WCC maintain a watching brief on the issues arising from the patenting of micro-organisms" (5, p. 26). The general thrust of the argument is that excessive patenting will lead to injustices in profits and to social inequities in the progress of research for the benefit of all.

The WCC 1982 report was followed in 1989 with a report from the WCC Subunit on Church and Society, entitled *Biotechnology: Its Challenges to the Churches and the World*. Like the 1982 report, the 1989 document included a major section on intellectual property, specifically on "The Patenting of Life." The report notes the economic impact of biotechnology patents on farmers and on scientific research, where patenting "is leading to a curtailment of communication and sharing of resources in the scientific community." The 1989 report notes, of course, that the alternative to patenting might be "complete research secrecy." But of all the economic impacts of patenting, "perhaps the greatest concern in the patenting of life is that the lure of patenting can cause a misappropriation of Third World genetic resources by corporations looking for patentable, genetic products. Since a significant amount of useful genetic material is found in the tropical and subtropical countries, First World patenting of life could increasingly exploit the collective resources and germplasm of Third World countries and peoples." This is not an insurmountable barrier, however, because already there are examples of "responsible sharing of information and benefits" (6, p. 22).

While the 1989 report observes that patenting may itself contribute to environmental problems, it moves quickly to the "Ethical and Theological Impacts" of patenting, where its strongest concerns are expressed. Here again, we find the living/nonliving distinction clearly articulated as the leading theological concern: "The patenting of life encodes into law a reductionist conception of life which seeks to remove any distinction between living and non-living things.... This mechanistic view directly contradicts the sacramental, interrelated view of life intrinsic to a theology of the integrity of creation. As

expressed by Arie Brouwer, former General Secretary of the National Council of Churches of the USA, 'Reverence for all life ... may be eroded by subtle economic pressures to view life as if it were an industrial product invented and manufactured by humans.... The gift of life from God, in all its forms and species, should not be regarded solely as if it were a chemical product subject to genetic alteration and patentable for economic benefit.' The churches should question all those technologies, whether traditional or modern, whose only stance toward creation is one of exploitation and profit, ignoring the biblical call to 'tend the garden and keep it'" (6, pp. 22–23). The final sentence is instructive in that the target of questioning is not patenting but technology itself; if technology is reductionist and blurs the distinction between the living and the nonliving, it should be challenged and rejected.

The report then enumerates (with cautious endorsement) a list of concerns about patenting: (*1*) First World control of Third World genetic resources by corporations looking for patentable genetic products, (*2*) encoding a reductionist conception of life by removing the distinction between living and nonliving things, (*3*) creating the loss of genetic diversity, (*4*) animal patenting creating the profit incentive for cross-species genetic transfers that leads to great animal suffering, (*5*) threatening environmental destruction by encouraging the deliberate release of genetically engineered organisms, (*6*) severe effects of patenting life on small farmers and producers, and (*7*) continued and growing confusion in the legal community over precise and acceptable definitions of patentable living matter (6). In light of these concerns, the report concludes with the recommendation from the Subunit on Church and Society that "The World Council of Churches believes that animal life-forms should not be patented and calls for further study of the profound moral and social implications of patenting life forms" (6, p. 23). In contrast to the preceding discussion, which is wide-ranging in scope, the final recommendation is limited to opposition to patenting of transgenic animals and leaves aside many other areas of biological and genetic patenting.

Somewhat in structural parallel to WCC, but operating on the national level within the United States, the National Council of the Churches of Christ (NCC) has also engaged in studies of the religious and ethical implications of biotechnology and has specifically addressed biological patents. A 1986 NCC report, entitled *Genetic Science for Human Benefit*, notes that "[s]cientists, investors and managers who provide the knowledge and capital necessary for biotechnological development and marketing deserve fair compensation for their ingenuity, work and willingness to incur economic risks." Patenting, however, can lead to "threats to science itself.... More serious still is the admonition against monopolistic ownership of genetically modified organisms or substances which are known to be essential to human life or for nourishment and health" (7, p. 14).

In contrast to this somewhat "pro-patenting" stance, more recent church positions have tended to view patenting with greater suspicion. For example, one of the largest Protestant denominations in the United States, the United Methodist Church, adopted a resolution at its highest national gathering, the 1992 General Conference. The resolution states that "[t]he position taken by the church ... is consistent with our understanding of the sanctity of God's creation and God's ownership of life. Therefore, exclusive ownership rights of genes as a means of making genetic technologies accessible raises serious theological concerns. While patents on organisms themselves are opposed, process patents, wherein the method for engineering a new organism is patented, provide a means of economic return on investment while avoiding exclusive ownership of the organism and can be supported.... We urge that genes and genetically modified organisms (human, plant, animal) be held as common resources and not be exclusively controlled, or patented. We support improvements in the procedures for granting patents on processes and techniques as a way to reward new developments in this area" (8, pp. 332–333). It was on the basis of this position that United Methodist officials went on to organize the May 1995 Joint Appeal, which was signed by many leaders of other religious bodies.

The largest U.S. Protestant denomination, the Southern Baptist Convention, adopted a resolution "On the Patenting of Animal and Human Genes" at its June 1995 national gathering, declaring that "the scriptures of both the Old and New Testaments plainly teach that God alone is creator and owner of all he has made ... [and] humans legitimately may own individual or groups of animals of a given species, but not an entire species and its progeny ... [Therefore,] we, the messengers of the Southern Baptist Convention ... do hereby affirm our conviction that God alone is Creator and owner of all creation ... we call upon the President, the Congress, the National Institutes of Health, and the United States Patent Office to place an immediate moratorium on the patenting of animal and human tissues and genetic sequences until a full and complete discussion has occurred" (9, pp. 7–8). The United Methodist and the Southern Baptist statements both stress the first theological principle noted earlier, namely, that God is the creator and owner of life. While the living/nonliving distinction is not unimportant, it is less important that the principle that God is creator and therefore owner of all living things.

The principle that God is owner of life and therefore we cannot claim ownership could be understood in a way that makes God appear possessive and petty. On the other hand, it can be interpreted as pointing to something important about the creation and its transcendent (as opposed to anthropocentric) ground of value. In this view, "as Creator, God reserves the right to determine how the knowledge of organic development is to be used. In a way, it is as if God the Creator were the first to patent genes, not to exclude us from using the knowledge, but to exclude us from excluding others.... It is as if anyone who tries to patent a gene should have the application rejected because it encroaches upon a prior patent whose holder, God, has given free license to everyone. As Creator, God has and reserves the right to define how we may use the knowledge of the creation, including the knowledge of DNA sequences. God exercises this right by saying, in effect, that this knowledge is to be used without exclusion, that is, without patent protection" (10, p. 154).

The second key principle, that of the importance of the distinction between the living and the nonliving creation, is more commonly asserted than the first in the theological statements opposing patenting. The fullest development of this distinction is found in the report noted earlier of the EECCS Bioethics Working Group. The report declares: "We object to the patenting of living organisms, to the patenting of human genetic material of any sort, and to non-human genetic material which has not undergone a major change by inventive means" (11, p. 38). This is because there is an important conceptual distinction that must be maintained between what is living and what is not. "Boundaries need to be drawn to make this distinction [between animate and inanimate matter] clear, to avoid reducing life conceptually to being merely an economic commodity, and then treating it as such" (11, p. 43).

It should be noted, however, that alongside the living/nonliving distinction, another boundary, this time between the human and the nonhuman, is also asserted in this report and elsewhere in church statements. For example, the report declares that "[p]atenting any part of the human genome is ethically abhorrent, in principle" (11, p. 44). In this context we should note that Roman Catholic concerns about patenting are not as sweeping or as urgent as those voiced by Protestants, in part because of the relatively greater emphasis Catholics place on the human/nonhuman distinction than do Protestants. For example, in one of the few Catholic statements on the subject, Pope John Paul II said to the Pontifical Academy of Sciences in 1994 that "...we rejoice that numerous researchers have refused to allow discoveries made about the [human] genome to be patented. Since the human body is not an object that can be disposed of at will, the results of research should be made available to the whole scientific community and cannot be the property of a small group" (12, p. 3). The opposition to patenting is both mild and limited to human DNA.

Generally speaking, Protestant statements rely on the living/nonliving distinction more than on the human/nonhuman distinction to define the limits of acceptable use of patenting as the means of intellectual property protection. According to the statement of EECCS, patenting is simply unacceptable for living things, in part because patenting is so firmly associated in the public mind with mechanical inventions that its application to living things conveys the wrong attitude toward life, one that is prejudicial in regarding living things as mere matter, mere mechanisms. In contrast to a mechanistic view, "[a]nimate material presents a radical discontinuity from mechanical and chemical inventions, which requires a different way of thinking about intellectual property.... Living organisms of any kind should not be patentable" (11, p. 42). For jurists or corporations to patent living things is an inappropriate use of governmental and financial power to the rule on a moral and theological question. In the strongest terms, the statement holds that "[w]e deplore the implications of various court decisions regarding patenting of living organisms.... This represents an unacceptable paradigm shift in how life forms are regarded.... This view sees nature entirely in anthropocentric terms of its utility to humans, as tools and products, and has lost the sense of respect for animal or plant as of value in itself. This perception runs contrary to Christian understanding that all of creation owes its existence to God, and its significance is first of all what it is before God, irrespective of any use to which human beings might think of putting it.... An animal, plant or microorganism owes its creation ultimately to God, not human endeavour. It cannot be interpreted as an invention or a process, in the normal sense of either word" (11, p. 43). But this does not necessarily mean that biotechnology should not be given *some form* of intellectual property protection. If "patenting" connotes "mechanism," surely some other legal category could be created that protects inventive value without demeaning natural value. And so the report suggests "that consideration should be given to developing an alternative form of intellectual property for biological material" (11, p. 38). In other words, a new form of intellectual property protection that encourages research but does not contribute to reductionism would be acceptable.

It should be expected that religious organizations, perhaps including those of other major faith traditions beside Christianity, will continue to express opinions on biological patents. The question of patenting can easily become the lightening rod for general concerns about technology, the human role in nature, and the meaning of life itself. "While patent law may seem to be a technical and arcane subject and therefore of little relevance to the religious community, patenting biotechnology raises important questions traditionally within the domain of religious ethics. These include such issues as the nature and status of life, the relationship between the Creator and the creation, and the protection of human dignity. Placing biology within the jurisdiction of intellectual property law also has important symbolic implications relating to whether humanity can or should claim ownership of forms of life" (13, p. 8).

INDIGENOUS PEOPLES AND THEIR COALITIONS

At the 1995 United Nations conference on women in Beijing, women representing indigenous peoples met for extended discussion resulting in the *Declaration of Indigenous Women*, which addressed many concerns, among them attitudes toward nature and the question of biological patenting. The Declaration affirms a view of the earth as "our mother. From her we get our life, and our ability to live. It is our responsibility to care for our mother and in caring for our mother, we care for ourselves. Women, all females, are the manifestation of Mother Earth in human form" (14). From this consciousness arose a concern about technology, global trade, and the right of corporations to patent biological components. Recent global trade agreements are criticized as "new instruments for the appropriation and privatisation of our community intellectual rights through the introduction of the trade-related intellectual property rights (TRIPs). This facilitates and legitimises the piracy of our biological, cultural, and intellectual resources and heritage by transnational corporations." These global agreements threaten traditional social values: "Our indigenous values and practice of sharing knowledge among ourselves,

and mutual exchange will become things of the past because we are being forced to play by the rules of the market." Furthermore global trade agreements legitimize the theft of traditional knowledge by incorporating it into patentable technological knowledge. "Bio-prospecting, which is nothing but the alienation of our invaluable intellectual and cultural heritage through scientific collection missions and ethnobotanical research, is another feature of recolonisation. . . . Their bid for the patenting of life forms is the ultimate colonisation and commodification of everything we hold sacred." The threat envisioned here is threefold: patenting endangers the value of traditional knowledge of plants or animals, cultural values such as sharing, and ultimately the survival of the people as a distinct identity. "It won't matter any more that we will disappear because we will be 'immortalised' as 'isolates of historic interest' by the Human Genetic Diversity Project." In highly passionate language, the Declaration urges resistance: "It is an imperative for us, as Indigenous Peoples, to stand in their way, because it means more ethnocide and genocide for us. It will lead to the disappearance of the diverse biological and cultural resources in this world which we have sustained. It will cause the further erosion and destruction of our indigenous knowledge, spirituality, and culture. It will exacerbate the conflicts occurring on our lands and communities and our displacement from our ancestral territories" (14).

The most immediate and specific form of resistance, according to the Declaration, lies in refusing to accept the universal validity of global trade agreements and their assertion of biological patenting. "We demand that the Western concept and practice of intellectual property rights as defined by the TRIPs in GATT, not be applied to indigenous peoples' communities and territories. We demand that the World Trade Organisation recognise our intellectual and cultural rights and not allow the domain of private intellectual rights and corporate monopolies to violate these. . . . We call for a stop to the patenting of all life forms. This to us, is the ultimate commodification of life which we hold sacred" (14). The Human Genetic Diversity project is particularly offensive, and the idea of any patent applications derived from these samples is especially abhorrent.

An incident that served to galvanize global opposition by indigenous peoples to biological patenting was the attempt by the United States Department of Commerce to file a patent application on a cell line derived from a Guaymi Indian woman from Panama. The Guaymi tribal president was widely quoted as saying, "I never imagined people would patent plants and animals. It's fundamentally immoral, contrary to the Guaymi view of nature, and our place in it. To patent human material . . . to take human DNA and patent its products . . . that violates the integrity of life itself, and our deepest sense of morality" (15, p. B5). Not only did this event and comment inspire opposition among indigenous people; they evoked a response far beyond indigenous peoples and served to mobilize opposition to patenting within Christianity and other religions.

So important is the issue of intellectual property to indigenous peoples that international conferences have been convened on this subject alone. For instance, 150 delegates from as many as 14 countries met in 1993 in New Zealand and issued the *Mataatua Declaration on Cultural and Intellectual Property Rights of Indigenous Peoples*. In the view of those gathered, "[w]e declare that Indigenous Peoples of the world have the right to self determination: and in exercising that right must be recognised as the exclusive owners of their cultural and intellectual property." The conference issued a call to indigenous peoples to develop protection strategies to resist corporations and governments in their unilateral pursuit of traditional knowledge or intellectual property protection. "Commercialisation of any traditional plants and medicines of Indigenous Peoples, must be managed by the indigenous peoples who have inherited such knowledge" (16) In the same year the *Intellectual Property Rights and Biodiversity* statement was developed by the Coordinating Body of Indigenous Organisations of the Amazon Basin (COICA). The COICA statement voices generalized suspicions about international intellectual property accords, which are seen as a new form of colonialist exploitation. The issue then becomes a matter of self-determination and of territorial sovereignty, especially in controlling access to traditional knowledge of plants within the traditional territory. "Biodiversity and the culture and intellectual property of a people are concepts that mean indigenous territoriality. . . . For members of indigenous people, knowledge and determination of the use of resources are collective and intergenerational. No indigenous population, whether of individuals or communities, nor the Government, can sell or transfer ownership of resources which are the property of the people and which each generation has an obligation to safeguard for the next" (17). Here again, the issues of traditional knowledge and of cultural values are both seen as threatened by global trade agreements. In particular, indigenous people are resisting the idea that they lack intellectual property or that they have no system to protect and regulate its value, and that therefore a intellectual property system can be asserted over their heads, so to speak. "Adjusting indigenous systems to the prevailing intellectual property systems (as a world-wide concept and practice) changes the indigenous regulatory systems themselves." Precisely because intellectual property systems are already in place, the assertion of alien Western-style "patents and other intellectual property rights to forms of life are unacceptable to indigenous people" (17).

In part through the assistance of the United Nations, indigenous peoples from around the world have come together to address common concerns, such as intellectual property. A gathering in Geneva in July 1999 resulted in *No to Patenting of Life! Indigenous Peoples' Statement on the Trade-Related Aspects of Intellectual Property Rights (Trips) of the WTO Agreement*, which begins with language that echoes the 1995 Beijing Women's Declaration. "*We, Indigenous Peoples* from around the world, believe that nobody can own what exists in nature except nature herself. A human being cannot own its own mother. Humankind is part of Mother Nature, we have created nothing and so we can in no way claim to be owners of

what does not belong to us" (18). The concern that is raised in the statement is that by asserting a global regime, trade agreements force global and therefore local consensus in favor of a western and modern worldview. What looks to be a conflict of law and economy is ultimately a clash of worldviews, as "western legal property regimes have been imposed on us, contradicting our own cosmologies and values" (18). The idea of owning nature "goes against the very essence of indigenous spirituality which regards all creation as sacred." This concern is remarkably similar to that expressed by EECCS, which protested that courts, by ruling on patenting, were in effect ruling on theology, on worldview, and deciding in favor of a mechanistic reductionism and against traditional theological or spiritual views of nature and life.

In the 1999 No to Patenting of Life! statement, as in the Beijing Declaration, the claim is made that traditional cultures have rich legacies of intellectual property, but that the intellectual value is commonly held because it is the result of common effort. These two systems of intellectual property are bound to clash, and indigenous peoples will inevitably lose in the ensuing struggle. This conviction is clearly asserted in the closing statement of the Consultation on the Protection and Conservation of Indigenous Knowledge, meeting in East Malaysia in 1995, which declared: "The intellectual property rights system is in favour of the industrialized countries of the North who have the resources to claim patent and copyright, resulting in the continuous exploitation and appropriation of genetic resources, indigenous knowledge and culture of the indigenous peoples for commercial purposes" (19). The problem is both structural and conceptual. As relatively weaker partners in political and economic discussions, indigenous peoples are unable to assert claims and to protect them. But the problem is also deeper in that the very concept of intellectual property advanced by the international trade agreements is at odds with the cultural foundations and worldview of many indigenous peoples. "The intellectual property rights system totally ignores the close inter-relationship between indigenous peoples, their knowledge, genetic resources and their environment" (19). Or as stated elsewhere: "The inherent conflict between these two knowledge systems and the manner in which they are protected and used will cause further disintegration of our communal values and practices. It can also lead to infighting between indigenous communities over who has ownership over a particular knowledge or innovation" (19). It must be clearly recognized that in these statements, indigenous peoples are making the claim that they possess intellectual property in the form of their traditional knowledge of nature. But their claim, they believe, is wholly incommensurate with Western legal protections and indefensible within global trade accords. If the accords win, their claims will inevitably lose. This is because the accords reflect one view of nature and ownership, one that favors technical over traditional knowledge. Furthermore it should be understood that it is not the idea of the value of knowledge of biology that is disputed in these statements. At dispute is something far more profound, namely, how *bios,* life itself, is to be regarded and how value is to be owned.

ENVIRONMENTAL GROUPS

Environmental groups have been active in voicing opposition to biological patenting, largely out of the conviction that widespread patenting will likely reinforce certain tendencies in the development of genetics and biotechnology that will have especially detrimental effects on the environment. Groups such as Greenpeace have organized and issued press releases to try to persuade the European Parliament (EP) to pass legislation that prohibits or sharply limits biological patents. In a press release in response to a 1997 vote by the EP, Greenpeace states that "[g]enes, and living organisms are not inventions and therefore they should not be patentable. Patent law was introduced to protect technical inventions. It has been enshrined in law that patents should only be granted if inventions are novel, are not discoveries and can be manufactured. Genes, plants and animals clearly do not fit with patent protection. They exist in nature so are discovered and not novel. Because they reproduce biologically, they are not man made. Allowing patents on life is highly immoral because it would give monopoly control over life to private interests for profit alone" (20). The argument here is based on Greenpeace's view that biological patents fail to meet the technical requirement of novelty. That argument expanded on a press statement issued just two weeks earlier: "Greenpeace believes it is immoral to claim genes, cells and living organisms are inventions of man to be used and controlled by commercial interests.... Greenpeace believes that by allowing patents on animals they will simply be seen as machines to be treated as people wish. Their dignity will be sacrificed for flimsy supposed benefits to humans." Of course, the core conviction for Greenpeace or any broad-based environmental group is that biological patents favor biotechnology, which threatens the environment. "Greenpeace also believes genetically engineered plants bring threats to the environment which are irreversible and unpredictable in nature" (21).

This environmental focus is echoed in a 1997 *Statement on Life and Evolution* drafted during the 1997 State of the World Forum (November 4–9) in San Francisco by environmentalists and posted on the Internet for others to sign. "Life is an intimate web of relations that evolves in its own right, interfacing and integrating its myriad of diverse elements. The complexity and interdependence of all forms of life have the consequence that the process of evolution cannot be controlled, though it can be influenced. It involves an unpredictable creative unfolding that calls for sensitive participation from all the players, particularly from the youngest, most recent arrivals, human beings." The statement makes a broader claim when it asserts that "[l]ife must not be treated as a commodity that can be owned, in whole or in part, by anyone, including those who wish to manipulate it in order to design new life forms for human convenience and profit. There should be no patents on organisms or their parts." The statement then returns to the core theme of threat to the environment, which of course leads not just to opposition to patenting but to a rejection of biotechnology: "We must also recognise the potential dangers of genetic engineering to health and biodiversity, and the ethical problems it poses for

our responsibilities to life. We propose a moratorium on commercial releases of genetically engineered products and a comprehensive public enquiry into the legitimate and safe uses of genetic engineering" (22). In contrast to the statements of religious and indigenous peoples groups, the statements of environmentalist groups about patenting are limited in the depth of their argument.

PUBLIC INTEREST SCIENCE GROUPS

Also limited in depth are the brief but influential statements of organizations like the Council for Responsible Genetics (CRG), which argues that "[n]o individual, institution or corporation should be able to claim ownership over species or varieties of living organisms. Nor should they be able to hold patents on organs, cells, genes or proteins, whether naturally occurring, genetically altered or otherwise modified." Their reasons are largely prudential or consequentialist, namely, that patent protection will restrict the practice of good science. The CRG statement adds that "[p]atenting organisms and their DNA promotes the concept that life is a commodity and the view that living beings are gene machines to be exploited for profit" (23). In an effort to build support for its position, CRG circulates a petition on the Web entitled *No Patents on Life*!

A similar strategy is found in the *World Scientists' Statement Calling for a Moratorium on GM Crops and Ban on Patents* (1999), issued by a group that is largely self-identifying and uses the Internet to recruit supporters. Its statement contains the following call: "Ban patents on living organisms, cell lines and genes." The supporting argument echoes the concerns raised by indigenous peoples: "The patenting of living organisms, cell lines and genes under the Trade Related Intellectual Property Rights agreement are sanctioning acts of piracy of intellectual and genetic resources from Third World nations, and at the same time, increasing corporate monopoly on food production and distribution. Small farmers all over the world are being marginalized, threatening long term food security for all" (24).

Another organization, the Crucible Group, met in 1993 and claims to "represent the widest cross section of sociopolitical perspectives and agricultural experience that may have ever been assembled. . . ." Perhaps because of the diverse perspectives, the gathering produced a report but did not claim that it is a "consensus document." Notably they did not achieve an agreement on biological patenting except to conclude that the matter is urgent, divisive, and requires global attention. "Sensing, on the one hand, a certain uncertainty and lack of understanding related to intellectual property regimes and, on the other hand, the opportunity to create a new covenant in support of wider innovative processes, the Crucible Group recommends that the United Nations convene an international conference on society and innovation. Now, and at this conference, policymakers must bear in mind that some people, countries, and cultures have deep ethical concerns about biotechnology and the concept of life patenting" (25).

SCIENTIFIC ORGANIZATIONS

The Human Genome Organization (HUGO), which is an international association of leading researchers in human genome research, has issued a *Statement on Patenting of DNA Sequences*. While the statement does not oppose gene patenting altogether, it notes that "[i]t would be ironic and unfortunate if the patent system were to reward the routine while discouraging the innovative. Yet that could be the result of offering broad patent rights to those who undertake massive but routine sequencing efforts — whether for ESTs or for full genes — while granting more limited rights or no rights to those who make the far more difficult and significant discoveries of underlying biological functions. A second, equally unfortunate outcome would arise if a partial sequence publication or submission to a database precluded patenting of innovative disease gene discoveries leading to improved medical diagnostics and therapeutics. This could lead to inhibition of contributions to databases and lack of investment protection for the innovative. We hope that the system will find some way to adjust to the changing realities in this field to promote and protect this important and ongoing process of discovery in the public interest" (26, p. 7).

This statement was followed by HUGO in 1997 with a one-page comment that updates and clarifies HUGO's position, namely, that "*reaffirms* . . . that HUGO does not oppose patenting of useful benefits derived from genetic information, but does explicitly oppose the patenting of short sequences from randomly isolated portions of genes encoding proteins of uncertain functions . . . *regrets* the decision of some patent offices, such as the US PTO, to grant patents on ESTs based on their utility 'as probes to identify specific DNA sequences,' urging these offices to rescind these decisions and, pending this, to strictly limit their claims to specified uses, since it would be untenable to make all subsequent innovation in which EST sequence would be involved in one way or other dependent upon such patents . . . [and] *urges* all large-scale sequencing centres and their funding agencies to adopt the policy of immediate release, without privileged access for any party, of all human genome sequence information in order to secure an optimal functioning of the international network, as well as to avoid unfair distortions of the system" (27).

HUGO's position is quite similar to the view put forward by the American Society of Human Genetics (ASHG) in its *Position Paper on Patenting of Expressed Sequence Tags* of November 1991. ASHG states that it has "taken the position that the issuing of patents for ESTs is likely to do far more harm than good. . . ." Once again, this partial objection to patenting is not to be construed as general opposition, for "ASHG has not opposed patenting of genetic information when that information had utility." The fear is that the patenting of short sequences of unknown function will inhibit research in genetic science, and therefore "[t]he ASHG does not support the concept of patenting a short sequence from a randomly isolated portion of a gene encoding a protein of unknown function" (28).

As should be expected, the opposition of science organizations to patenting will be based on and limited to the negative effects of patenting upon scientific research.

Likewise, physicians might be expected to object to patenting that would limit the prompt application of research to medicine, and in fact the American College of Medical Genetics (ACMG) has voiced this concern. In its *Position Statement on Gene Patents and Accessibility of Gene Testing*, ACMG asserts the belief that gene testing "must remain widely accessible and affordable, and that the development and improvement of safe and effective genetic tests should not be hindered. The decision of the Patent and Trademark Office (PTO) to permit the patenting of naturally occurring genes and disease-causing mutations has produced numerous difficulties. While the ACMG disagrees with the PTO over this fundamental issue, we have further concerns over current patterns of enforcement of patents on genes that are important in the diagnosis, management and risk assessment of human disease." It should be noted that the scope of ACMG's objection is broader than that of the genetic science organizations, precisely because ACMG is concerned not merely for research but for the transfer of the benefits of research to the clinical setting. With that in view, ACMG complains of "...exorbitant up-front fees and per-test fees, and licensing agreements that seek proportions of reimbursement from testing services. These limit the accessibility of competitively priced genetic testing services and hinder test-specific development of national programs for quality assurance. They also limit the number of knowledgeable individuals who can assist physicians, laboratory geneticists and counselors in the diagnosis, management and care of at-risk patients." Patent protection and excessive licensing fees unduly restrain the field of clinical genetics, and "restricting the availability of gene testing has long-term implications beyond patient care. It affects the training of the next generation of medical and laboratory geneticists, physicians, and scientists in the area enveloped by the patent or license. It also retards the usually very rapid improvement of a test that occurs through the addition of new mutations or the use of new techniques by numerous laboratories that have accumulated samples from affected individuals over many years" (29).

As a result of this analysis, which is argued entirely on grounds that patenting will inhibit clinical genetics, the statement concludes that "it is the ACMG's position that ... [g]enes and their mutations are naturally occurring substances that should not be patented ... [p]atents on genes with clinical implications must be very broadly licensed ... [and] [l]icensing agreements should not limit access through excessive royalties and other unreasonable terms" (29).

Most scientific organizations draw back from such broad opposition to patenting and limit their concerns to specific misuses of the patenting process, especially the patenting of short sequences of unknown function. In a statement issued by the British Society for Human Genetics (BSHG) in 1997, the general principle of the patentability of genetic knowledge is affirmed: "Patenting is a valuable means of protecting intellectual property and promoting investment in developing products for the diagnosis and treatment of genetic disease." However, "[t]he discovery of gene sequence has for some little time been a well understood process. There is nothing novel or inventive about this in principle, and as such new gene sequences should not be patentable, even where a straightforward utility e.g. diagnostic testing has been specified, unless there has been real progress towards the design of a specific commercial product." Patent offices must address, in greater precision, questions of usefulness or novelty, which "cannot reside in the mere description of a nucleotide sequence. It must rest in either novel methodologies for discovering the sequence or a novel use or application of the sequence. Conventional technology, conventionally applied, should not result in patents on newly isolated sequences." In a similar way, the standard for a claim for usefulness must be clear and fairly high: "A claim for utility must describe a utility specific to the sequence in question, and not simply rehearse those possible applications of any known gene sequence which are part of the general public state of the art." Specifically, claims of utility should be denied if they are based upon "use for isolating the full gene sequence ... use for detecting mutations in the gene ... [or] use for studying expression or function of the gene." Stated positively, the criterion of utility applicable to gene patenting ought to be "some meaningful indication that the sequence being patented has a reasonable prospect of being developed into a marketable product (which may be a diagnostic test) ... [or] a proposed specific use — for example, diagnosis of mutations in people with a specified clinical indication." The BSHG statement concludes with a call to patent offices to limit the scope of gene patents to the "specified applications which meet the novelty and utility criteria" (30).

It was noted earlier that the possible patenting of discoveries linked to the Human Genome Diversity Project is especially offensive in the view of the statements of indigenous peoples. To respond to this concern, the North American Regional Committee of the Human Genome Diversity Project has drafted a model protocol for collecting DNA samples. The model contains this description of the position of the HGDP regarding intellectual property rights: "The HGDP has no position on questions of patentability, although individuals participating in the HGDP hold a variety of positions. The HGDP does, however, hold clear positions about the commercial use of its samples and of the information derived from them." There are two principles to which all HGDP researchers are required to give assent: "First, it [HGDP] has resolved that it will not profit from any commercial uses of samples it gathers or knowledge derived from those samples. Second, it has vowed to ensure that, should commercial products be developed as a result of the HGDP's collections, a fair share of the financial rewards shall return to the sampled populations" (31, p. 1466). Enforcement mechanisms are not specified, nor is it clear that this will satisfy the concerns of indigenous peoples.

INDUSTRY ORGANIZATIONS

In response to the May 18, 1995, Joint Appeal of religious leaders that called for sweeping bans on biological patenting, Gerald J. Mossinghoff of the Pharmaceutical

Research and Manufacturers of America (PhRMA) stated that "PhRMA believes that it would be immoral for the pharmaceutical and biotechnology industries to walk away from new technologies that could stop pain, suffering and hunger. Because patents on animals, cell lines, genes, and their products are necessary to foster such scientific enterprise, PhRMA believes it is a moral imperative that the patenting of these types of inventions be maintained and encouraged" (32). This position is more fully developed in a PhRMA Policy Paper, *Strong Patent Protection Is Essential*, which argues that patent protection is necessary if research is to be funded and new pharmaceutical products brought to market. "Without strong patent protection, there simply would be no research-based pharmaceutical industry, which discovers and develops virtually all new medicines. As a result, few new life-saving, cost-effective medicines would be developed, and improvements in the quality of health care would be sharply curtailed" (33). The view is echoed by the major firms engaged in this area of research, such as SmithKline Beecham, whose 1997 "Patenting Statement" states: "Patent protection offers the only effective incentive for bringing to market the many commercial and industrial applications for which genetic inventions may be used.... SmithKline Beecham supports the patentability of any inventions which meet the federal patentability requirements of subject matter, utility, novelty, and non-obviousness" (34).

Any suggestion that genes or biological components are *living* and therefore inherently unpatentable is of course rejected in industry statements. The position of the Biotechnology Industry Organization (BIO) is expressed by Alan Goldhammer in these words: "BIO supports continuation of the current law to allow patenting of human genes when the applicant meets the necessary criteria for securing any patent: the invention must be novel, nonobvious and useful. When these standard criteria are met, patents should be issued irrespective of the nature of the invention. No exception should be made for patents on genes, life forms, or any other subject matter" (35). Without any doubt, BIO's position is fundamentally at odds with the views put forward by the organizations of indigenous persons and by some of the religious groups.

While the position of industry does not change significantly from country to country, it is interesting to note that regional industry groups advance the additional pro-patenting argument that national opposition to patenting will undermine a nation's biotechnology industry and therefore the nation's (or the region's) entire economy. For example, the Forum for European Bio-Industry Coordination (FEBC), in a Directive on the legal protection of biotechnological inventions, argues that a positive environment on biological patenting will encourage regional reinvestment in research and development. "Without patents there would be less investment in research. Patents are the foundation on which the development of new products like pharmaceuticals ... depend." At a time when Europe was considering a sharp limitation on biological patents, the industry warning was clear: "FEBC believes that any weakening of the draft would put Europe at a further disadvantage and will shift the emphasis of research in biotechnology further towards the USA and Japan." The statement documents the warning with evidence that appeared to show that biotechnology investment in Europe was falling behind that in the United States and Japan, and stated that "[p]art of the reason for this gap is the lack of harmonised patent practice. The message is clear: although European investment has increased, the competitive gap between Europe and the US and Japan is still increasing" (36).

Many in industry and in government recognize the urgency of the problem posed by intellectual property protection for advances in knowledge in genetics and biology. The Human Genome Project, which has established an informal agreement that raw sequence data will be posted daily on the Web rather than patented, finds itself in competition with private efforts to sequence the genome in order to gain proprietary advantage. Speaking of the differences between the publically funded National Human Genome Research Institute (NHGRI) and private efforts such as the joint Perkin-Elmer-TIGR genome project, Francis Collins, the Director of NHGRI, told Congress that the "release of sequence data from the Perkin-Elmer-TIGR effort will occur quarterly, rather than daily. The policy of daily release of DNA sequence data by publicly-funded efforts was arrived at because of the great interest in the scientific community in gaining access to this highly valuable information. Any delay can result in wasted effort in research" (37). Deep differences of opinion and philosophy exist even here between scientists and laboratories that are engaged in essentially the same research.

BIBLIOGRAPHY

1. General Board of Church and Society of the United Methodist Church, Press Conference, Washington, DC, May 18, 1995.
2. Rev. Kenneth L. Carder, Statement to the press, Washington, DC, May 18, 1995.
3. Rev. Wesley Granberg-Michaelson, Statement of Religious Leaders against Animal Patenting, August 6, 1987; appended to testimony before the Subcommittee on Courts, Civil Liberties and the Administration of Justice of the Committee on the Judiciary, U.S. House of Representatives, Concerning Animal Patenting, November 5, 1987.
4. European Ecumenical Commission for Church and Society, Working Group on Bioethics, *EECCS and Bioethics*, EECCS, Strasbourg, France, 1998.
5. World Council of Churches, *Manipulating Life: Ethical Issues in Genetic Engineering*, WCC, Geneva, 1982.
6. World Council of Churches, Subunit on Church and Society *Biotechnology: Its Challenges to the Churches and the World*, WCC, Geneva, Switzerland, August 1989.
7. National Council of the Churches of Christ in the U.S.A., *Genetic Science for Human Benefit*, NCC, New York, 1986.
8. United Methodist Church, *Book of Resolutions, 1992*, Nashville, 1992.
9. Southern Baptist Convention, *SBC Bulletin*, Nashville, 1995, pp. 7–8.
10. R. Cole-Turner, in A.R. Chapman, ed., *Perspectives on Genetic Patenting: Religion, Science, and Industry in Dialogue*, American Association for the Advancement of Science, Washington, DC, 1999, pp. 149–165.

11. Ecumenical Commission on Church and Society, Bioethics Working Group, *Critique of the Draft EU Patenting Directive, a Submission to the European Parliament and the European Commission*, EECCS, Strasbourg, France September 1996; printed *Drawing the Line — The Ethics of Biotechnology*, Occasional Paper No. 5, by the Ecumenical Association for Church and Society, 1997, pp. 37–50.
12. Pope John Paul II, *L'Osservatore Romano*, November 9, 3 (1994).
13. A.R. Chapman, in A.R. Chapman, ed., *Perspectives on Genetic Patenting: Religion, Science, and Industry in Dialogue*, American Association for the Advancement of Science, Washington, DC, 1999, pp. 7–40.
14. *Declaration of Indigenous Women*, Beijing, 1995, Available at: *http://www.southside.org.sg/souths/twn/title/indig-cn.htm*
15. P.L. Bereano, *Seattle Times, Sunday*, August 27, B5 (1995).
16. *The Mataatua Declaration on Cultural and Intellectual Property Rights of Indigenous Peoples*, Aotearoa New Zealand, June 1993, Available at: *http://users.ox.ac.uk/~wgttr/mataatua.htm*
17. Coordinating Body of Indigenous Organisations of the Amazon Basin (COICA), *The COICA Statement: Intellectual Property Rights and Biodiversity*, Santa Cruz de la Sierra, Bolivia, September 1994, Available at: *http://users.ox.ac.uk/~wgtrr/coica.htm*
18. "No to Patenting of Life! Indigenous Peoples' Statement on the Trade-Related Aspects of Intellectual Property Rights (TRIPS) of the WTO Agreement," signed at the United Nations, Geneva, Switzerland, on July 25, 1999, Available at: *http://www.alphacdc.com/ien/intellectual property.html*
19. United Nations Development Programme, *Consultation on the Protection and Conservation of Indigenous Knowledge, Sabah, East Malaysia*, February 24–27, 1995, Available at: *http://users.ox.ac.uk/~wgtrr/sabah.htm*
20. Greenpeace, *Greenpeace Deplores EP Vote on Bio Patent Directive; Leverage Buyout of Nature — Ethics under Corporate Control?* Strasbourg and Brussels, July 16, 1997.
21. Greenpeace, *Patents on Life: Immoral, Driven by Greed and Damaging to Medical Research*, Strasbourg and Brussels, July 2, 1997.
22. *Statement on Life and Evolution*, San Francisco, November 1997, Available at: *http://www.psrast.org/swf97reg.htm*
23. Council for Responsible Genetics, *No Patents on Life! DNA Patents Create Corporate Monopolies on Living Organisms*, Cambridge, MA, 1994, Available at *http://www.genewatch.org/*
24. *World Scientists' Statement Calling for a Moratorium on GM Crops and Ban on Patents* (1999), Available at: *http://www.i-sis.dircon.co.uk/home.htm*
25. Crucible Group, *Statement*, 1993, Available at: *http://www.idrc.ca/books/725/preface.html*
26. Human Genome Organization, *HUGO Statement on Patenting of DNA Sequences*, Bethesda, MD, January 1995.
27. Human Genome Organization, Bethesda, MD, 1997.
28. American Society of Human Genetics, *Position Paper on Patenting of Expressed Sequence Tags*, November 1991.
29. American College of Medical Genetics, *Position Statement on Gene Patents and Accessibility of Gene Testing*, August 2, 1999.
30. British Society for Human Genetics, *Patenting of Human Gene Sequences and the EU Draft Directive*, September 1997, Available at: *http://www.bham.ac.uk/BSHG/patent eu.htm*
31. North American Regional Committee of the Human Genome Diversity Project, Model Protocol: Proposed Model Ethical Protocol for Collecting DNA Samples, *Houston Law Rev.* **33**(5), 1431–1473 (1997).
32. Pharmaceutical Research and Manufactures of America (PhRMA), News Release, Washington, DC, May 18, 1995.
33. Pharmaceutical Research and Manufacturers of America (PhRMA), Policy Paper, 1995.
34. SmithKline Beecham, Patenting statement, 1997.
35. A. Goldhammer, Director of Technical Affairs, Biotechnology Industry Organization (BIO), Testimony before the Subcommittee on Technology of the House of Representatives, Committee on Science, September 17, 1996.
36. Forum for European Bio-Industry Coordination (FEBC), *FEBC Views on the Directive on the Legal Protection of Biotechnological Inventions*, Brussels, 1995.
37. Statement of Francis S. Collins, M.D., Ph.D., Director, National Human Genome Research Institute on the Human Genome Project, Statement before the Subcommittee on Energy and the Environment Committee on Science, United States House of Representatives, June 17, 1998.

See other entries HUMAN GENOME DIVERSITY PROJECT; OWNERSHIP OF HUMAN BIOLOGICAL MATERIAL; see also PATENTS AND LICENSING entries.

PATENTS AND LICENSING, ETHICS, OWNERSHIP OF ANIMAL, AND PLANT GENES

MICHELE SVATOS
Iowa State University
Ames, Iowa

OUTLINE

INTRODUCTION

Current biotechnology patenting practices raise major ethical, legal, and economic controversies. Patenting "life," including plants, animals, and their genes, has been criticized as unethical and inimical to social justice. The Council for Responsible Genetics circulated "The No Patents on Life! Petition," which argues:

> The plants, animals and microorganisms comprising life on earth are part of the natural world into which we are all born. The conversion of these species, their molecules or parts into corporate property through patent monopolies is counter to the interests of the peoples of this country and of the world. No individual, institution, or corporation should be able to claim ownership over species or varieties of living organisms. Nor should they be able to hold patents on organs, cells, genes or proteins, whether naturally occurring, genetically altered or otherwise modified (1).

Patent advocates retort that of course life cannot be patented — merely genes and genetically modified organisms (GMOs) (2). However, it is unlikely that patent critics actually misunderstand that point, as is evident in No Patents on Life! petition; indeed, many of its supporters are scientists. Life itself, as some élan vitale, surely cannot be patented, although one can patent living organisms and genes; that alone is bad enough according to critics. Let us consider the arguments for and against animal and plant patents. After a brief review of the legal status of life patenting and the relation of patent law to morality, we will look at arguments specifically addressed at animal patenting first, before considering arguments that apply more generally to both animal and plant patents.

PATENT LAW AND MORALITY

Patents grant the right to exclude others from making, using, or selling an invention for a period of 20 years. A patentable invention must be novel, useful, and nonobvious. Ideas, theories, mathematical algorithms, laws of nature, and the like, cannot be patented; processes (e.g., production methods, special techniques, and diagnostic methods), products (e.g., microorganisms, enzymes, plasmids, cell lines, and DNA and RNA sequences), and new uses of an existing product can be patented.

Plants were the first living organisms to receive explicit patentlike protection. The Plant Patent Act of 1930 provided protection of asexually reproduced plants, and it was mainly applied to flowers and certain fruits. The broader 1970 Plant Variety Protection Act (PVPA) provides patentlike protection for new varieties of plants, although it differs from a patent in lacking the utility requirement. Plant breeders use both plant variety protection and patents to protect genetically modified plants and plant genes.

Since the landmark 1980 Supreme Court decision *Diamond v. Chakrabarty*, living organisms — "anything under the sun that is made by man" — can be patented in the United States (3). However, the *Chakrabarty* decision focused on a microorganism, and did not explicitly address higher life forms (4). In 1987 the Commissioner of the U.S. Patent and Trademark Office (PTO) formally affirmed the patentability of "nonnaturally occurring non-human multi-cellular living organisms, including animals." It clarified that "Products found in nature will not be considered to be patentable subject matter . . . unless given a new form, quality, properties or combination not present in the original article existing in nature in accordance with existing law" (5).

In 1988 the first transgenic animal was patented, the "Harvard mouse" or "Oncomouse" (6), genetically modified for hypersensitivity to carcinogens for medical research. Animals are genetically modified for medical research, agriculture, and as pharmaceutical biofactories. Animal, plant, and human genes may be patented, including cloned, unmodified genes. Most animal patents thus far are for mice and rats.

Morality comes into play at various levels in the legal practice of patenting (7). Normally the extension of the patent system to a new technological discipline is automatic and may be assumed. However, unique features of new technologies can result in difficult questions of interpretation for patent law. In modern biotechnology, the distinction between discovery and invention is becoming blurred. Moreover genetically modified organisms (GMOs) are unique as inventions. Not only are some of them alive, but also they are able to reproduce on their own, and are not well standardized, easily described, and so on. If they are released into the environment, they will interact with it unpredictably. These are good reasons to pause, rather than extending or excluding patent protection automatically. Old definitions and criteria must be reconsidered and sometimes redrawn.

Although the legal question of the patentability of life forms has been settled in the United States and Europe, some popular sentiment against it remains in these countries. Moreover it remains a lively issue within the international community, as developed nations pressure less developed countries (LDCs) to adopt Western intellectual property laws. LDCs are rich in genetic resources, and many object to intellectual property laws and alleged "biopiracy" on grounds of morality and social justice. Thus the moral debate over life patenting has become increasingly important to discussions of international trade relations and international justice. Are there any special reasons to believe that animal and plant patents are more or less justified than other types of patents, that they should be subject to special conditions or exclusions, or that they raise special problems of ethics and social justice?

ANIMAL PATENTS

Animal patents are criticized on the grounds that they:

- Encourage increased production of transgenic animals, many of which suffer greatly
- Violate animals' rights
- Encourage the reduction of animals to mere inventions
- Engender additionally most of the criticisms that apply to plant patents (see the discussion below)

Several philosophers and legal scholars, such as Baruch Brody (9), Robert Merges (10), and Rebecca Dresser (11), have considered the arguments against animal and plant patenting in depth. They have argued that the criticisms are without merit or point, to problems that are not fundamentally about patenting. Let us consider the arguments for and against animal patenting in greater detail.

Animal Welfare

Animal welfare advocates warn that animal patents encourage more research that, on balance, tends to increase animal suffering. With animal patents, claims Michael Fox of the U.S. Humane Society, "the wholesale industrialized exploitation of the animal kingdom will be sanctioned, protected, and intensified" (8). Although genetic engineering may sometimes decrease animal suffering, critics are concerned that its record thus far is not good, and that this trend will continue. In fact genetic engineering poses special problems for animal welfare, above and beyond normal concerns about the treatment of farm and research animals. While many traditionally bred animals are routinely subjected to inhumane treatment, transgenic animals may be engineered to inevitably suffer even under the best conditions, for example, cancer-prone or hairless mice (12). The U.S. Animal Welfare Act does not apply to farm animals, birds, or rodents, making critics skeptical that genetic engineers will be subject to limits on cruelty (9). Insofar as genetic engineering is likely to cause more suffering for animals than it prevents, and insofar as patents promote the development and commercialization of genetically engineered animals, animal welfare advocates attack animal patenting as an economic incentive for much morally problematic research.

Patent advocates respond that animal welfare advocates should attack the problem of cruelty to animals directly, rather than through patenting (9,11–13). Banning such patents would discourage even research designed to alleviate animal suffering, and which may help many humans. Even if some transgenic animals suffer more, their suffering might be outweighed if it is possible to use fewer but more efficient animals. Moreover patents are perfectly compatible with strong animal welfare regulations, and in themselves are morally neutral.

A few animal welfare advocates have admitted that if sufficient animal protection regulations were in place, they would have less or no opposition to animal patenting. This makes sense for utilitarian animal welfare advocates, many inspired by utilitarian philosopher Peter Singer's book *Animal Welfare*. Singer argues that animal suffering is morally relevant but may be overridden by the likelihood of greater benefit to others (14). Thus some research on animals could be justified — if there were no adequate alternatives, if the benefit were great enough, and if the suffering were small enough. How much animal research would remain is a matter of debate; many utilitarians argue that it would be very little. Most of animal agriculture is not justified on a utilitarian animal welfare view, since vegetarianism is usually a feasible alternative. Thus utilitarian animal welfare advocates, as critics of most of animal agriculture and research, are natural critics of animal patents, but mostly because and insofar as it is likely to increase the suffering of animals. The key assumption is that patented transgenic animals are likely to suffer in order to more efficiently promote our medical and agricultural ends. While this initially may seem extremely pessimistic, the charge warrants examination. For utilitarian animal welfare advocates, criticisms of animal patenting would naturally be combined with a more direct approach toward alleviating animal suffering and strengthening animal welfare laws — and indeed that is often the case, as with Michael Fox.

Animal Rights

A more radical view of animals and morality, inspired by philosopher Tom Regan, holds that animal suffering is morally relevant but that animals also have moral *rights* and may not be used as a mere means to our medical, agricultural, or recreational ends — regardless of the benefits (15). Animal rights advocates are likely to be categorically opposed to animal patenting, on the ground that it involves the use of animals as property and inventions, mere means to our ends. They are likely to accept genetic engineering of animals only where it might be justified on human infants, done on behalf of the child or animal rather than directly or indirectly in service of adult goals. Thus disease-resistant transgenic chickens would probably not be justified even if the animals suffered less, in the animal rights view. Even where genetic engineering might be justified, animal rights advocates could say that it is unjust to patent the genetically engineered animal, since doing so amounts to treating that variety of animal as mere property rather than as a living entity with rights and its own ends.

Thus, even radical revisions in animal welfare laws are unlikely to move animal *rights* advocates (as opposed to utilitarian animal welfare advocates). Patent defenders sometimes are confused and frustrated by animal rights advocates' criticism of patenting. We treat animals as property all the time. Most people do not consider animals to have rights, so animal welfare concerns could be more directly addressed through animal welfare regulations (10). However, all this makes sense given a certain philosophical grounding for concern for animals, namely an animal rights view. Not only could animal patents work to perpetuate or promote animal suffering, the mere practice of patenting animals violates animal rights by treating them as means to our ends. Patent advocates may respond that public policy should not be molded to accord with one particular minority view of animals and morality. However, for animal rights advocates, the situation is analogous to slavery. If all humans have moral rights, then slavery should be abolished, regardless of whether the majority believes in it and benefits from it. Animal patenting could be compared to patenting of genetically engineered slaves: although slavery is the main practice that should be abolished, the logical implication is that patenting of slaves should be abolished as well, especially if it eases the commercialization of slaves. If the main practice seems unlikely to disappear in the near future, then attacking the newer, more controversial patenting practice makes

sense, particularly if there is a real possibility of banning patenting.

Extending an animal rights view further, biotechnology critics such as Jeremy Rifkin argue that animals have a right to species integrity violated by genetic engineering (16). Defenders reply that genetic material can move between species in nature, and we have been genetically manipulating plants and animals for generations through traditional breeding. Moreover the notion of species' rights is extremely controversial, even compared to the concept of animal rights. (4,10). Finally, this criticism does seem directed more toward genetic engineering than patenting.

Reductionism

Animal and sometimes plant patents are also criticized on grounds that are less directly related to animal welfare. Critics argue that patents are reductionistic and encourage and intensify the commodification and objectification of living organisms, particularly animals. Characterizing GMOs as human "inventions" attaches too much importance to the contribution of scientists and too little importance to what God or nature has given us. It demonstrates an attitude of hubris or arrogance that is also apparent in our release of these organisms into the environment when we cannot fully understand or control the risks. Let us consider these objections, which take both secular and theological forms, in more detail.

Objectification. Patent critics argue that the language used to describe transgenic animals reveals a mechanistic view of them. They are bioreactors, biofactories, disease "models," and so on. While objectification occurs without animal patents, patenting takes the objectification to a new level by reconceptualizing animals as human-made utility inventions and "compositions of matter" (17,18). Critics argue that God or nature creates life, and it is absurd and arrogant to patent living organisms as human inventions. Even GMOs usually have only one or a couple of transgenes, which are merely transferred from another living being rather than being created de novo. At most we are moving around a few parts of God's or nature's creations (19,20).

Patent advocates reply that a patent does not mean that one has created an invention from nothing; patents are allowed for improvements on preexisting things. Of course, humans did not invent the mouse, and they do not patent mice, the entire species. Instead, they patent mice that have been genetically altered in a small but significant way. This does not detract from nature or God's handiwork, it merely adds to it. Scientists would be the first to admit that they are far from being able to create an entirely new animal out of nothing. Moreover genetic engineering, patenting, and the use of terms such as "biofactory," are compatible with proper respect for life in other contexts. Although humans are technically "compositions of matter," they are also much more, and so are other living organisms (21).

Patent critics may be skeptical of this response. It might be compared to the argument that treating people as sex objects is compatible with respecting them in other contexts; feminists are skeptical that compartmentalizing really works that well. Objectifying animals helps us view their plight as morally irrelevant, just as objectifying women helps rationalize violence against them, and objectifying the enemy in war helps soldiers commit acts of violence against them (18,22). While it is not clear that this argument holds merit, it is also not clear that patent advocates are correct that viewing animals and plants as inventions has no larger implications for our treatment of them and our conception of humanity's relationship to the rest of the living world. There may be grounds for making a distinction between the ownership of particular animals, and the ownership of, say, all mammals with a certain gene. There is, one might argue, a difference in attitude (which is simply a fact of our society) between the owner of a cow or dog, and the owner of a technology/invention. This takes the mechanistic view of nature even further than it has previously been extended. We may think that we have finally refuted and rejected Descartes's view that animals are simply machines — but wait! They are, after all, but our own technologies, our own inventions. To say that we own an animal is different than saying that we invented them or that they are a technical solution to a technical problem (7). It is a mistake to completely attribute this debate to a misunderstanding of the practical, patent law distinction between discoveries (unpatentable) and inventions (patentable). Indeed, some molecular biologists agree with the public on this point, hardly out of ignorance of science or patent law. Science and technology, discovery and invention, are becoming increasingly blurred in biotechnology. If a gene, its function, and its structure are scientific discoveries, how can the purified and isolated gene become an invention (23)? The public is perplexed and accuses patent lawyers of playing word games. Critics may charge that patent law has focused too much on technical definitions at the expense of the commonsense "fact" that human, animal, and plant genes are discoveries, not inventions, and that even genetically modified plants and animals still do not qualify as inventions (24,25). We will raise this issue again in the final section.

Commodification. The Council for Responsible Genetics argues that "Patenting organisms and their DNA transforms them into commodities for profit and promotes the view that living beings are little more than 'gene machines'" (26). One may well not be opposed to the research itself but rather to certain forms of its commercialization made possible by patenting. Most people support genetic research that would lead to new forms of human gene therapy. They might even be in favor of commercializing gene therapy, to the extent of making it a private for-profit venture and allowing many aspects of the research (e.g., new processes) to be patented. Opponents of patenting genes argue that genes should and in some sense do belong to everyone and no one. Every researcher should be free to work on them and try to use them to benefit humankind, without having to get a license and pay a royalty. Of course, if companies are able to patent only the process they use but not the gene product, it will be difficult for them to control profits, but to stake a claim on genes themselves is to go too far. Opponents of animal

patents may view the primary problem as patenting itself rather than, or above and beyond, the suffering that occurs. Groups such as Global Action in the Interest of Animals (GAIA) object to the patenting of animals even where the animals do not suffer, or actually benefit, from a genetic modification (17).

Patent advocates reply that we have owned and bred animals for centuries. The fact that a genetically engineered line of mice is patented is unlikely to affect our treatment of those animals, and animal welfare concerns should be separated from patenting (4,9). Moreover patents are needed in order for companies to have an incentive to invest in and commercialize socially useful research. However, critics, including many scientists, are skeptical that this is actually necessary, and argue that it may actually hinder research (27–29).

Why object to patenting genes and GMOs, apart from concerns about genetic engineering itself? This view should not be confused with a criticism of research into a particular area. It is rather an argument for limiting the subjects of product patents, regardless of whether we wish to limit the research itself, and regardless of the fact that patenting genes probably does meet the criteria of current patent law and is explicitly supported by the *Chakrabarty* decision.

HUBRIS AND RISKS

"From an ethical perspective, the patenting of animals reflects a human arrogance toward other living creatures that is contrary to the concept of the inherent sanctity of every unique being and the recognition of the ecological and spiritual interconnectedness of all life" (30). Michael Fox argues that patenting will accelerate the "transformation of life and of the creative process to serve purely human ends, and as many see it, the end of the natural world" (8). Such critics also often question the wisdom of moving genes around. They worry that we do not adequately understand the risks, that we are transgressing natural barriers that might be best respected, and that we often harm transgenic animals. In our hubris we consider GMOs and even unmodified genes to be our inventions, and we further assume that these creations are so under our control that the risks of introducing them are minimal (31). Critics see patenting living organisms and genes as arrogant and unwise. They worry that the upshot of patenting living organisms may be the devaluation of all life, including human life (18,31).

Patent advocates respond that once again, critics should attack the real "culprit," rather than patenting. Concerns about the risks of genetic engineering should be addressed by the U.S. Environmental Protection Agency (EPA), the Food and Drug Administration (FDA), or other regulatory agencies, rather than the patent office. The mere fact that one has patented a genetically engineered organism provides no assurance that one will actually be able to use and market it; it may not receive regulatory approval for release (32). If regulatory processes for genetically engineered organisms are seen as inadequate, then they should be criticized, rather than patenting. If genetic engineering is unwise or unnatural (either in general, or in particular cases), then criticize genetic engineering itself rather than patenting (9).

Rather than being a *symptom of* worries about risks, the hubris objection to patenting "life" is sometimes more accurately classified as *parallel to* concerns about risks. Biotechnology critics hold that our arrogance and greed leads us to engage in risky behavior, and at the same time leads us to consider living organisms and genetic resources as mere inventions to be patented. Another part of the same arrogance and greed is our treatment of animals, which is poor to begin with and often only intensified by genetic engineering. We hold ourselves up as masters and creators of the natural world rather than a humble part of it. This view can take theistic or secular forms, and it may be accompanied by more or less skeptical attitudes toward the possibility of wise, safe genetic engineering. Some critics see genetic engineering as entirely unnatural or immoral, and others do not reject it out of hand but only want it to be done and applied with wisdom and justice.

PATENTING, FARMERS, AND "BIOSERFDOM"

Some farmers' advocates oppose plant and animal patenting. One major concern is that patenting will further encourage factory farms and the consolidation of agribusiness at the expense of small farms and farming communities. Critics charge that agribusiness is becoming an oligopoly. As companies merge and become vertically integrated, they control more of agriculture, and independent animal and plant breeders and dealers can scarcely compete (33). With large multinational corporations heavily invested in genetic engineering, they have to heavily promote their products to recoup research costs. Large corporations are better equipped to handle patent applications and infringement cases. If there are few competitors left, then there will eventually be few affordable, nongenetically engineered varieties available to small farmers, who are so squeezed for profits that they can scarcely afford to make more expensive or less efficient purchasing decisions than their corporate competitors. Practically speaking, they will have little alternative to purchasing expensive patented varieties, both because little else will be offered for sale and because their larger competitors will also be doing it. The genetically engineered varieties are more likely to be well suited to large corporate farms, with their high technology inputs and large economy of scale.

Patents may further mean that farmers are prohibited from breeding animals and seeds for their own use. Patent critics and even sympathizers have argued that at minimum, there must be an exemption for breeding by farmers, particularly small- and medium-sized farmers. Currently U.S. law does not have such an exemption, although European law has an exemption for "small" farmers (33). In the absence of such an exemption, patent advocate Robert Merges argues there are two practical problems. First, considerable record keeping would probably be required of farmers to prove that they did not breed any patented animals without paying the proper royalty. Second, enforcement procedures might be difficult, and might require farmers to open their

farms' animals to inspections to ensure that no patent law violations had taken place (10). Even if farmers are allowed to breed patented varieties without paying royalties, genetic drift assures that in a few generations, most will want to renew their stock from commercial breeders to retain desired traits.

These practical problems also raise moral issues, namely the reduction in the autonomy and privacy of farmers. Critics charge that even Western farmers are becoming poorly paid employees with no applicable minimum wage laws and little autonomy. "As the life industry dictates more and more of the farm-level management decisions, the farmer becomes little more then a 'renter' of proprietary germplasm and information, a step in the food/industrial manufacturing process. Farmers and consumers thus increasingly lose control over what products they grow and consume, and which food production processes they chose to support" (34, p. 5). Some critics have denounced this as bioserfdom. They worry that animal and plant patenting is merely another step on the way to bioserfdom and corporate agriculture, with poor conditions for farmers and animals alike. Let us consider the role of licensing restrictions and the so-called terminator seed in promoting this situation. As we will see, an exemption or derogation for farmers might not be sufficient to prevent it.

Seed Saving and Licensing Restrictions

Farmers traditionally have saved seeds from part of their crop for breeding, trading, and planting the next season. This still occurs to a certain extent in developed countries, particularly for non-hybrid crops such as cotton and rice. The practice is more common among poor farmers in less developed countries (LDCs). "Between 15 and 20 percent of the world's food supply is grown by poor farmers who save their seed. These farmers feed at least 1.4 billion people" (35). Seed companies view seed saving as a problem in expanding their market and protecting their intellectual property, namely proprietary seeds. An analogy may be made to computer software; farmers buy the right to plant the seeds but not to make "pirated copies" by planting second generation seeds. "From an industrial perspective, plant varieties are to be used for growing a crop. All other unauthorized uses of varieties lower returns on investment and therefore must be eliminated" (33, p. 139). Constantly improving seeds is one way to encourage farmers to buy seeds every growing season. However, seed research is expensive and most improvements are small, so this alone may not be sufficient, particularly for nonhybrid crops. So seed companies look to new ways to protect their intellectual property from what they view as the piracy of farmers.

Restrictive licensing agreements attempt to ensure that seeds will not be saved and replanted without payment of royalties. Monsanto has been criticized for its practice of requiring such contracts, which include a rather unpopular provision that Monsanto's agents — sometimes dubbed "gene police" — may enter the farm and test for Monsanto genes for a period of three years following the initial purchase of seeds (36). Monsanto even provides a toll-free number for farmers to report suspected violations and has publicized the names of violators on radio stations, in addition to suing for compensation. Of course, no one is *required* to buy seeds from Monsanto, so if farmers don't like this policy, they may take their business elsewhere, and avoid trading seed with Monsanto customers.

Such a policy of policing intellectual property is criticized as an example of bioserfdom. It is obviously not popular with many farmers and requires considerable expenditures for enforcement and litigation. In LDCs, matters are even more difficult; enforcement costs might be higher, and legal support dubious or lacking. Indeed, proprietary seeds could make it to a developing country without the company's knowledge, and many LDCs are notorious for failing to protect intellectual property. Such licensing agreements may turn out to be an unpopular and not especially effective strategy of protecting intellectual property.

Technology Protection System: "Terminator" Seeds

A further step in protecting seed intellectual property from farmers is to build a technological "fence" around proprietary seeds, comparable to software copy protection schemes (37). Indeed, such a fence is in the development stages: TPS (technology protection system) seed is being developed by the U.S. Department of Agriculture (USDA) and Delta and Pine Land Co., dubbed "the terminator seed," "suicide seeds," or "traitor technology" by Rurual Advancement Foundation International (RAFI) and others. TPS seeds are genetically modified with an extra gene that seed companies can "turn on" before selling the seeds (38) so that second generation seeds are sterile. "The Terminator provides a built-in biological 'patent,' enforced by engineered genes" (35, p. 1). Farmers would buy TPS seeds, which offer farmers no direct benefit in themselves, because the TPS trait would be paired with other valuable traits. They would then have no alternative but to purchase new seeds each growing season, or to buy whatever seed varieties do not include TPS. Of course, critics worry that agribusiness industry is shaping up in such a way that there will be few choices.

The media, developing countries, and several non-governmental organizations (NGOs) have expressed outrage over TPS. NGOs such as RAFI have called for a global ban, claiming that:

> It is a global threat to farmers, biodiversity, and food security. The seed-sterilizing technology threatens to eliminate the age-old right of farmers to save seed from their harvest and it jeopardizes the food security of 1.4 billion people — resource-poor farmers in the South — who depend on farm-saved seed. . . .If the Terminator technology is widely utilized, it will give the multinational seed and agrochemical industry an unprecedented and extremely dangerous capacity to control the world's food supply (39).

Similarly the Center of Education and Technology in Chile warns, "This is an immoral technique that robs farming communities of their age-old right to save seed and their role as plant breeders. Farmers and governments everywhere should declare the use of the technology as contrary to public order and national security. This is the neutron bomb of agriculture" (40). Influential groups such

as CGIAR (Consultive Group on International Agricultural Research) have condemned TPS, and several countries have banned it.

In response to these sorts of criticisms, the USDA released a *Fact Sheet: Why USDA's Technology Protection System (a.k.a. "Terminator") Benefits Agriculture—A Discovery to Spur New Crop Improvement*. The fact sheet claims:

> Because of [farmers'] seed-saving practice, companies are often reluctant to make research investments in many crops because they cannot recoup their multiyear investment in developing improved varieties through sales in one year (38).

The fact sheet goes on to assert that farmers as well as the environment will benefit from new seed varieties that seed companies will be inspired to develop, given the reassurance that their intellectual property will be protected. In particular, "small farmers may benefit greatly if the invention stimulates the extension of biotechnology to 'minor crops' such as tomatoes," since the seed industry will see greater potential for return on their research investment into these minor crops (38). Advocates argue that farmers in developing countries will still be able to save their traditional and public seed varieties. Rather than posing a threat to world food security, it will do the opposite, as seed companies will feel more secure in releasing genetically improved varieties in developing countries with poor intellectual property protection. This will help even the playing field between U.S. farmers who abide by intellectual property laws, and LDC farmers who often do not. In other words, it requires everyone to pay their fair share of the seed research that they benefit from. The market for terminator seeds will not extend beyond its benefits to farmers since no one is forced to buy TPS seeds. Moreover it is a safe technology that poses no risks to the environment, and that may indeed help prevent the unintentional spread of other genetically modified traits. Finally, TPS is part of a larger research program that will allow the controlled expression of many traits in plants, and that has the potential to enormously benefit agriculture.

This is scarcely the place to resolve the controversy, which has received little serious academic attention (41). However, it is clear that TPS is appropriately named Technology Protection System. The debate must be understood as part of the debate over plant patents, especially restrictive licensing agreements. Although TPS is several years away from commercial release and may never prove viable in its current form, expect further attempts at "fencing" plant—and eventually animal—intellectual property.

"Bioserfdom," Western Farmers, and LDCs

In sum, if patents tend to work more to the advantage of large corporations, and if increasing control of agribusiness by large corporations tends to work to the advantage of corporate factory farms, then small farmers view animal and plant patents as another stake being driven through the heart of the Jeffersonian ideal of the yeoman farmer. Critics believe that patents are a government incentive that works mostly in the favor of corporate agribusiness, and contrary to the purposes of preserving the independent family farm. If the government does not want to support family farms, they should at least not provide incentives for the corporate competition that already enjoys many advantages. Again, this is not to say that individual patents might not help small farmers, but on balance the practice of biotech patenting is seen as a tool for agribusiness and corporate farming. Corporate breeders want farmers to keep coming back to them every time rather than breeding their own sometimes, and patents and licensing agreements help ensure that this happens. Many farmers are skeptical that this ultimately will be to their advantage, even if it does encourage more investment in research. The agricultural technology treadmill may help early adopters of new technologies, but given current low prices and overproduction, in the end these new technologies will only drive down food prices further—which might be nice for the consumer but not for most farmers. A strong system of patents in agriculture is, critics say, likely to promote corporate dependence rather than sustainable farming.

If this is true, will patents at least benefit farmers and consumers in countries where overproduction is scarcely a problem? Patent critics argue that these countries are most in need of sustainable solutions rather than increased dependency on Western corporate agriculture. The poorest farmers in those countries are subsistence farmers who cannot afford genetically engineered animals and seeds, plus the capital-intensive farming methods that tend to go with them. Critics worry that like the Green Revolution, new genetic technologies will simply perpetuate or even worsen inequalities of wealth by benefiting large, wealthy farms in those countries (60). Moreover, even if the result is cheaper food, this does not benefit subsistence families. Furthermore many of the crops grown in those countries are actually for export, and do not help feed the poor in that country. Multinational corporations argue that patents are good for American farmers because it is unfair that American farmers pay royalties while farmers in LDCs do not. However, advocates of LDCs argue that they can hardly afford to pay such royalties; how much are U.S. farmers really disadvantaged when compared to farmers in LDCs?

Patent advocates respond that "the economic forces driving a family farm into liquidation, or an academic institution into embracing some corporate suitor, operate quite independently of patents . . . [patenting] is essentially neutral as to oligopolistic trends that may be at work in the present economy. . . .Think twice before turning to the patent system for a means to alter their course" (32, p. 9). "If the government wants to avoid any negative impact of animal patents on the family farm, the appropriate approach is to create mechanisms to enable all farmers to gain access to this new technological development through agricultural extension services and special subsidies" (4). Although other approaches may be necessary to address many of these concerns, patent critics may still be justified in their concern that extending patent law to living organisms is likely to provide incentives that help structure the market in ways that are more favorable to corporate agriculture.

PATENTING AND BIODIVERSITY

If farmers are not allowed to breed plants and animals themselves, biotechnology critics such as Vandana Shiva charge that this increases the centralization of breeding and promotes genetic uniformity rather than biodiversity (19,43). As farmers relinquish their independence and are compelled to forgo most breeding, and as independent breeders go out of business or are bought out, there are fewer options offered and less difference among the options. Monoculture is already a dangerous trend in agriculture, and patenting will just encourage the marketing of varieties with a great deal of genetic uniformity.

Patent advocates respond that monoculture is indeed a worry, however, patents do not contribute to it. Indeed, allowing the patenting of genetically engineered plants and animals may make it more profitable for researchers to innovate to avoid the dangers of genetic uniformity, and to preserve genetic resources for use as raw materials (32). Shiva and others reply that genetic engineering may alter traits or create new ones, but generally this involves only one or a few genes. Thus, even if genetic engineering provides new varieties, they may differ little from other varieties in ways other than the engineered trait (19). Moreover regardless of how new the trait is, if millions of "copies" are sold, there is considerable genetic uniformity. Insofar as patents and licensing agreements limit the ability of farmers to engage in independent breeding, these practices in themselves do contribute to monoculture and the loss of biodiversity.

PATENTING, BIOPIRACY, TRIPS, AND NATIONAL SELF-DETERMINATION

In response to concerns about patenting and bioserfdom, patent advocates assert that "the patent office is not the place to structure a morally appropriate program for the international economic order" (9). However, in the age of information technology, developed countries want and demand strong intellectual property protection worldwide, and consider the absence of such laws to be an unfair trade barrier; developed countries use the patent office to promote their own goals including fairness. Patent advocates often miss the fact that plant and animal patenting raises issues of national self-determination, at least in the era of TRIPS (Trade Related Intellectual Property Rights). "With the advent of TRIPS, virtually all the world's nations have lost their right to determine the balance of private and public benefits designed to meet national goals. Instead, they must comply with a single international standard designed to open their markets to transnational corporate interests" (44, p. 1). Although many LDCs are opposed to plant and animal patenting (if not all patenting), they are being required to adopt patent or patentlike protection for plants and animals, or else suffer serious economic consequences.

According to patent critics, TRIPS are particularly egregious, considering the fact that developed countries have taken or "pirated" the genetic resources of the South for centuries; each side accuses the other of piracy and attempts to take the moral high road. While the North is economically and technologically rich, the less developed South is genetically rich. Plant and animal varieties have been exported from LDCs for centuries, with no compensation to the countries where they were domesticated and made more valuable over many generations of work by indigenous peoples. Indeed, they are treated as and referred to as unimproved genetic resources, rather than the collective property of indigenous communities. International seed banks have been established to save these genetic resources, and many Northern researchers come looking for genetic "gold" to bring back to their laboratories, often with indigenous people as their free "tour guide" (19,45). Recently some researchers have offered compensation to the country of origin if a successful product is created, but much so-called biopiracy occurs without compensation. Moreover compensation is usually small, and often the indigenous community has no control over how the money is used. If it is devoted to preserving the country's genetic resources, this may be rather self-serving on the part of researchers who want to mine further genetic treasures (33).

Many LDCs see genetic resources as the common heritage of humankind which they have no desire to privatize, much less have privatized by others (16,46). They are outraged to discover that medicinals or crops first developed in their country are patented by multinational corporations, or even by the U.S. government. Indeed, the U.S. National Institutes of Health (NIH) has gone so far as to apply for a patent on cultured tissue samples taken by indigenous peoples themselves (which it later dropped). Critics are aghast that not only the country's genetic resources, but also the tissues from its people, are being privatized and commercialized. Patents on some of India's genetic resources have been extremely controversial. India accuses corporations of pirating the sacred neem tree, traditional medicinal uses of turmeric, and basmati rice. While some of these patents have been dropped or overturned, countries such as India continue to worry about biopiracy, and object to international pressure to adopt strong intellectual property protection, particularly over living organisms and genes.

This debate raises issues of international justice, national self-determination, and what should remain in the commons rather than being privatized (47). Many patent critics view genetic resources as a commons, to be shared for the benefit of all (46). Some patent advocates reply that if LDCs are concerned about biopiracy, they should privatize their own genetic resources. If research companies come and find these resources unclaimed, the indigenous peoples have no one to blame but themselves. Simply privatize and fence in the genetic commons, and corporations will have to buy the genetic resources they need from other countries (48). Critics view this response as impractical and beside the point. Even if the practical problems could be surmounted, the moral objections to privatizing the commons would remain. Part of the concern is that developed countries are taking unfair advantage of poor countries, but another part of the concern is that privatizing genetic resources is simply not appropriate; they should belong to no one and everyone, for the mutual good of all.

CONCLUSION

The debate over patenting the products of biotechnology is often polarized. Critics are characterized as being anti-biotechnology, or at least anti-genetic engineering, and quite possibly technophobic. Often the objections to patenting, which to a great degree are interrelated, are characterized as attacking the wrong practice, perhaps to increase public criticism of biotechnology in any way possible, or perhaps because the critics themselves are confused. However, sometimes the attack of patenting is perfectly logical, even if it may rest on the attack of other practices.

Moral Justification of the Patent System

Patenting is characterized by its advocates as morally neutral, and even most critics would agree that we do not want the patent office to act as the arbiter of technology and morality. On the other hand, patenting is supposed to advance morally significant goals. Patents were established to promote useful inventions and thereby benefit society — a form of utilitarian justification. Alternatively, they were established to protect the natural rights of inventors to the fruits of their labors (50). Thus, in a sense, the patent system depends on moral argument. What is the appropriate scope of patent law in biotechnology? It depends on what we take to be the real philosophical justification of the patent system. If patents are justified by the extent to which they encourage research, would this extend to the patenting of living organisms and even human genes? Or if the justification is the protection of the natural rights of inventors to the fruits of their labors, can this be extended to living organisms and especially human genes (27)?

If we choose the former rationale, then it is reasonable to question the extent to which plant and animal patents are likely to benefit society as a whole, particularly in an era when the Western patent system is being imposed internationally against the wishes of numerous countries. Many Westerners simply assume that patents are effective and beneficial. However, it is instructive to examine this assumption, which some argue is unwarranted (27,49). Patents historically were grants of monopoly protection for those who imported, not invented, technologies. Patents initially played a greater role in technology transfer and excluding competitors than in promoting research. They are now being used to keep developing countries from the "piracy" European countries initially designed them to promote when it was in their own interests. In nineteenth-century Europe, many philosophers and economists opposed patents on the ground that they were forms of corporate protectionism that prevented the efficient operation of the free market (50). Now most people assume the opposite, but is the assumption justified? Perhaps not. Patents have played an interesting role in the transformation of agriculture into agribusiness, and it is instructive to study the history of agribusiness opposition to and later support of patents (33,45,51). Indeed, most of the debate over "patenting life" is explicitly tied to larger questions and assumptions about the structure of agriculture and the effects of globalization.

If, on the other hand, the protection of the natural rights of inventors is the primary justification for patents, then it is perfectly reasonable to question the extent of these rights. In particular, it makes sense to consider what belongs in the genetic commons as discoveries and the natural heritage of humankind rather than industrial or government property. Patent defenders, even philosophers, often ignore arguments about the genetic commons. Yet this is the most unique and independent part of the debate. This is a very philosophical issue about the nature of property rights and the commons, and does not rest on larger concerns about animal welfare, family farms, and the like. As such, and given its role in the international debate over TRIPS, this issue deserves greater attention, especially from philosophers.

Novelty, Nonobviousness, Breadth, and Patenting Discoveries

Many responses to criticism of plant and animal patenting ignore problems of the breadth, novelty, and utility of some of the patents being issued. Such problems are not necessarily new or unique, but they may be worse when applied to GMOs. Questions about patenting living organisms have been raised by many scholars who are not necessarily biotechnology critics, such as Louis Guenin (52), Philippe Ducor (53), and Brian Cannon (54). It is instructive to see how the moral arguments against this patenting practice fits within this more academic debate. Moreover some of these scholars have specific proposals that might make patent law more consistent and logical while at the same time addressing some of the moral concerns regarding the patenting of living organisms and genes (although such proposals are beyond the scope of this article). Let us briefly consider some of these criticisms.

Charges of biopiracy may be recast as concerns about the way that the patent law criterion of novelty is being applied. Patent applications must identify any prior art that would be relevant to the patent. The fact that indigenous peoples may have cultivated and improved a variety over generations, or developed a particular use for a plant, is not reflected in prior art, which consists primarily of Western scientific writings. Thus something invented by indigenous peoples could easily be patented as a new invention — even if it had not been modified at all. On appeal, the patent might or might not be rejected — assuming that someone has the knowledge and resources to appeal it. Our Western patent system does very little to acknowledge our transfer of genetic resources and indigenous knowledge. Thus the charge of biopiracy may be seen as at least partially a concern about how the criteria of novelty and nonobviousness can be understood within an international, cross-cultural context.

Plant and animal patents are also criticized for being overly broad (16). Patents are often characterized as a trade-off between the interests of inventors to recoup their costs and the interests of society in having inventions widely and cheaply distributed. Patent advocates argue that this trade-off works to the best interests of society in the long run. However, overly broad patents may strike

the wrong balance in this trade-off, and allow patent holders too much control over the development of further research. Since *Diamond v. Chakrabarty*, "anything under the sun that is made by man" can be patented in the United States. Unfortunately, there is a tendency for patent applications to claim not just anything but rather everything under the sun. Allowing broad patents may not be in the best interests of developing nations, some of which already depend on genetically engineered rice or other crops to feed their growing populations and perhaps produce a small amount for export. Even within one country, very broad patents may benefit one firm at the expense of other firms and often also the public interest (27).

Finally, patenting unmodified genes (human, animal, or plant) is often criticized, even by those engage in biotechnology research. The U.S. PTO allows patenting of cloned, unmodified genes. Cloned genes are not the same as their counterparts in nature, and cloning involves an inventive step. Thus, when scientists discovered and cloned a human breast cancer gene, they patented it. Likewise animal and plant genes can be patented, even if they have not been genetically modified. Patent lawyers explain that GMOs and genes that have been isolated and purified (cloned) are a technical solution to a technical problem, and do not occur naturally. Contrary to the No Patents on Life! petition, they are not "part of the natural world" and do not exist without considerable human innovation.

To patent critics, this is the height of absurdity. Inventions are patentable, but not discoveries. We say that scientists *discover* a gene, not that they *invent* it (49). Patent critics, some of them scientists who understand what goes into cloning a gene, remain unconvinced by the answers of patent lawyers. The general public may not have the knowledge of patenting criteria to express their concern properly, but they are appealing to an intuitive form of the idea that patents must meet the criteria of novelty, utility, and nonobviousness. A cloned gene seems to them to fail the criterion of novelty and/or nonobviousness. To say that scientists *invented* an animal hormone or a breast cancer gene seems inaccurate and arrogant, even though most of these critics would applaud breast cancer research. Although the location and sequence of the gene might be nonobvious, once a gene is discovered, is the process of cloning it nonobvious and the resulting cloned gene novel?

Finally, many critics of plant, animal, and gene patents are perfectly happy to accept process patents. It is a mistake to attribute all of these criticisms to concerns about the research itself, or concerns about animal welfare (7). Rather, they are often genuine concerns about what should constitute a valid patent, what constitutes a criterion such as novelty, and whether it makes sense to say that an entire line or species of animals or plants could be owned (as opposed to owning individual animals and plants). Patent advocates often oversimplify the views of critics, whose spokespersons often are not philosophers or legal scholars and whose arguments are sometimes admittedly not fully developed. However, there are indeed serious issues here. The patenting of animals and plants raises many moral and social issues that have not yet been resolved and that will only become more important in an increasingly globalized information economy.

ACKNOWLEDGMENT

This article was made possible in part by an Iowa State University Science, Technology, and Society grant. Some of the initial research was made possible by a research leave at the Social Philosophy and Policy Center. I would also like to thank my capable and well-organized research assistants, Stacy Bastian and Katherine Dresser, my secretary Janet Krengel, and my husband Richard for playing single dad while I hid in my office.

BIBLIOGRAPHY

1. Council for Responsible Genetics, Fall 1998.
2. S.R. Crespi, in S. Sterckx, ed., *Biotechnology, Patents and Morality*, Ashgate, Brookfield, VT, 1997, pp. 220–221.
3. *Diamond v. Chakrabarty*, 447 U.S. 303, 308, 309 (1980).
4. R. Dresser, *Jurimetrics J.* **28**, 399–435 (1989).
5. U.S. PTO Commissioner D.J. Quigg, Ex parte Allen, April 7, 1987.
6. U.S. Pat. 4,736,866 (April 12, 1988).
7. M. Svatos, in S. Sterckx, ed., *Biotechnology, Patents, and Morality*, Ashgate, Brookfield, VT, 1997.
8. M. Fox, *Atlanta J.* (1987), cited by S. Bent, in W.H. Lesser, ed., *Animal Patents: The Legal, Economic, and Social Issues*, Macmillan, New York, 1989.
9. B.A. Brody, in W.H. Lesser, ed., *Animal Patents: The Legal, Economic, and Social Issues*, Macmillan, New York, 1989.
10. R. Merges, *Maryland Law Rev.* **47**, 1051–1075 (1988).
11. R. Dresser, *Jurimetrics J.* **28**, 399–435 (1989).
12. B. Rollin, *The Frankenstein Syndrome*, Cambridge University Press, New York, 1995.
13. L. Walters, Committee on the Judiciary, 1988.
14. P. Singer, *Animal Liberation*, Avon Books, New York, 1990.
15. T. Regan, *The Case for Animal Rights*, University of California Press, 1983.
16. J. Rifkin, *The Biotech Century*, Putnam, New York, 1998.
17. M. Vandenbosch, in S. Sterckx, ed., *Biotechnology, Patents and Morality*, Ashgate, Brookfield, VT, 1997.
18. B. Belcher and G. Hawtin, *A Patent on Life: Ownership of Plant and Animal Research*, International Development Research Centre, Ottawa, 1991.
19. V. Shiva, in *Biopolitics*, V. Shiva and I. Moser, eds., Zed Books, London, 1995.
20. J. Densberger, *J. Bioethics* 91–115 (1984).
21. M.J. Hanson, *Hastings Center Rep.* (November, December), 1–18 (1997).
22. A. Garry, *Soc. Theory Practice* **4**, 395–442 (1998).
23. L. Kluver, in S. Sterckx, ed., *Biotechnology, Patents and Morality*, Ashgate, Brookfield, VT, 1997.
24. B. Mitchell, in *Patenting of Biological Entities, Proceedings of the ITEST Workshop*, ITEST Faith/Science Press, St. Louis, MO, 1997.

25. N. Hettinger, *BC Environ. Affairs Law Rev.* **22**, 267–305 (1995).
26. Council for Responsible Genetics, *Synthesis/Regeneration* **17**, (Fall 1998).
27. M. Svatos, *Social Philosophy and Policy* **13**(2), 113–143 (1996).
28. B. Weutrich, *Sci. News* **144**, 154–157 (1993).
29. D. Nelkin, *Science as Intellectual Property*, Macmillan, New York, 1984.
30. Committee on the Judiciary, 1988, pp. 64–65, quoted in B.A. Brody, in W.H. Lesser, eds., *Animal Patents: The Legal, Economic, and Social Issues*, Macmillan, New York, 1989.
31. L. Kass, *Toward a More Natural Science*, Free Press, New York, 1985.
32. S. Bent, in W.H. Lesser, eds., *Animal Patents: The Legal, Economic, and Social Issues*, Macmillan, New York, 1989.
33. R. Pistorius and J. van Wijk, *The Exploitation of Plant Genetic Information*, Academisch Proefschrift, Amsterdam, 1999.
34. RAFI Communique, *Bioserfdom: Technology, Intellectual Property and the Erosion of Farmers' Rights in the Industrialized World*, March, April, 1997.
35. R.A. Steinbrecher and P.R. Mooney, *Ecologist* **28**, 4 (1988).
36. R. Weiss, *Washington Post*, February 3, p. A01 (1999).
37. T.G. Palmer, *Hamline Law Rev.* **12** (1989).
38. U.S. Department of Agriculture, *Fact Sheet Why USDA's Technology Protection System (a.k.a. "Terminator") Benefits Agriculture—A Discovery to Spur New Crop Improvement*, Government Printing Office, Washington, DC, 1998.
39. RAFI Communique, *The Terminator Technology*, March, April, 1998.
40. C. Montecions, Centro de Education y Tecnologia, Chile.
41. M. Svatos, Ethical and social issues raised by the 'terminator' seed, and technological fences for intellectual property: Comparing the 'terminator' seed to software copy protection unpublished manuscripts, 1999.
42. M. Crouch, *J. Agri. Environ. Ethics* **4**, 151–158 (1991).
43. J. Doyle, *Altered Harvest*, Viking Press, New York, 1985.
44. K. Dawkins, *Gene Watch* **12**(4), 1 (1999).
45. C. Fowler and P. Mooney, *Shattering: Food, Politics, and the Loss of Genetic Diversity*, Tucson: University of Arizona Press, 1990.
46. H. Shand, *J. Agri. Environ. Ethics* **4**(2), 131–142 (1991).
47. P. Drahos, *A Philosophy of Intellectual Property*, Dartmouth, Brookfield, VT, 1996.
48. J. Vogel, *Genes for Sale*, Oxford University Press, New York, 1994.
49. E.R. Gold, *Body Parts*, Georgetown University Press, Washington, DC, 1996.
50. F. Machlup and E. Penrose, *J. Econ. Hist.* **10**(1), 279–314 (1950).
51. J. Kloppenburg, *First the Seed: The Political Economy of Plant Biotechnology, 1492–2000*, Cambridge University Press, New York, 1988.
52. L. Guenin, *Theoret. Med.* **17** (1996).
53. P. Ducor, *Patenting the Recombinant Products of Biotechnology and Other Molecules*, Kluwer, Boston, 1998.
54. B. Cannon, *Cornell Law Rev.* **79**, 735–765 (1994).

See other entries MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT; see also PATENTS AND LICENSING entries.

PATENTS AND LICENSING, POLICY, PATENTING OF INVENTIONS DEVELOPED WITH PUBLIC FUNDS

DANA KATZ
JON F. MERZ
University of Pennsylvania
Philadelphia, Pennsylvania

OUTLINE

INTRODUCTION

Should governments fund scientific research? If so, what types of research, basic or applied? How should taxpayers' money be allocated? What priorities should the government adopt, and how should those priorities be set? If government research yields commercializable products (other than solely for military purposes), how should those developments be moved into the market? Should private parties be permitted to have exclusive rights to publicly funded inventions? And, ultimately, who are, and who ought to be, the beneficiaries of publicly funded research?

To analyze these questions, we look here at the development of research funding policy in the United States. Because of its free-market roots, the United States has approached these questions with trepidation. Over the last 200 years the country has moved incrementally from funding only research having military importance to current policies that provide substantial support for basic biomedical sciences, among many others, and permit private commercialization of and profit from the resulting intellectual property. This intellectual property is heralded as the means to economic prosperity both domestically and in the global marketplace. Though this can be considered standard free-market industrial policy, the public that invests in scientific innovation conceptualizes its benefits in ways other than economic, such as through the development of life-saving medical technology or an increase in general knowledge. Although there are cases where both economic and social benefits of science can be achieved concurrently, technological and industrial growth, particularly in the area of biotechnology, has increased the potential for these goals to compete. Today the government promotes scientific innovation with marketplace incentives which, while serving to promote technological advance and the development of public goods, raises numerous ethical concerns, such as its perpetuation of inaccessible health

care for many individuals and the privatization of basic research knowledge.

EARLY INNOVATION

Although in 1776 there was little tangible evidence of the benefits of scientific advancement, the founding fathers intuitively recognized the importance of innovation for a developing nation. It was argued during the Constitutional Convention that the wording of the constitution ought to reflect the new republic's "duty and ability to encourage progress in the arts and sciences" (1). To this end, some argued the need for constitutional provision for technical schools, societies, seminaries, and a national university. Those who saw pecuniary incentives as the cornerstone of innovation sought constitutional provisions for patenting rights, as well as rewards, prizes, and direct subsidies for citizens who endeavored creatively to promote agriculture, commerce, and other social goods.

Ultimately innovation was afforded very limited constitutional protection for fear that broader commitments would strengthen the central government and thereby increase the potential for later abuse of power. The only explicit constitutional provision for innovation grants Congress the power to:

> . . .promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries (2).

A. Hunter Dupree, in his seminal work *Science in the Federal Government*, interprets this clause as sanctifying the creation of national scientific institutions (1). Prior to its first comma, the clause states the intention to promote advancement of the arts and sciences. The intention is qualified by the English practice of affording inventors exclusive ownership of their work as an incentive for further innovation and development. This qualification, Dupree argues, is neither a prohibition of, nor suggestion for, publicly funded research because the Constitution does not address the concept directly. For the century following, however, opponents of government support of research argued that Congress was permitted to do only what the Constitution explicitly authorizes so that the absence of a provision permitting research funding prohibits Congress from doing so.

Although the language of the Constitution falls short of reflecting its framers' high regard for innovation, the government was constructed to rely on scientific expertise. From its inception, the basic operations of the federal government have relied on technical skill, such as surveying new territory and establishing coinage, weights, and measures. However integral to its function, the role of science in the federal government was informal and would remain so for generations.

According to Dupree, the nation's inherent regard for science stems from the invocation of natural law used to deliver its founders from the political and clerical despotism of England (1). With the same rationalist approach used to create the new republic, the founding fathers sought to expand natural history, philosophy, political theory, and science in the New World (1). The widespread dislike of bureaucratic and oppressive rule, however, led the federal government to call on scientists only when needed rather than to expand its organization to employ them full-time. As a result the pursuit of knowledge was slated to occur largely outside of the government domain.

The nation's early scientists were upper-class, professional males who pursued scientific inquiry as a hobby. Though their science was amateurish, their powerful positions in society allowed them to effectively introduce scientific concepts and traditions into American society. Their observations rarely concentrated on one aspect of the natural world, such as chemistry, physics, or botany, and this lack of specialization limited the potential for scientific accomplishment (1). With little distinction between pure and applied research, "science" became the catch-all category for products of systematic investigation that ranged from philosophical theories of electricity to the construction of ornamental clocks. Although it was recognized that science offered potential benefits, the most significant innovations of the era were the results of trial and error by those unacquainted with the scientific method (1).

In 1791 the federal government received its first request to fund scientific research. In seeking a patent for navigational calculations, a surveyor requested that Congress finance an arctic voyage to test his theories. Though there was some question as to whether the matter should be deferred to state legislators, the larger ethical question of whether the government ought to use public funds for such uncertain ends went unasked. Many supported the concept, seeing the issue only in terms of the nation's need to improve geography and navigation. Others, while reluctant to prematurely reject a potentially beneficial enterprise, advised against the hasty support of philosophical patent applications (1). Ultimately the request was refused on account of the poor success rate of arctic voyages.

A federal patenting board had already been instituted in 1790 to process what was fast becoming an overwhelming number of patent applications. Arguments that either a panel of experts or ordinary citizens ought to compose the board were defeated on the belief that a few politicians, under the guidance of Thomas Jefferson, would offer inventors the best protection. Authorized to withhold patent rights for inventions lacking novelty, utility, or importance, the nation's first patent board demonstrated how government regulation could be used to promote the general welfare.

Jefferson understood that the fundamental task of a patenting board is to advance society by delicately balancing two often-competing goals: the desire to encourage innovation versus the need to prevent destructive monopolies. Patents essentially function as a government-endorsed monopoly in that they confer exclusive rights to use, produce, and sell an invention. The provision of exclusivity serves as an incentive for individuals to generate useful, novel, and important inventions and develop them into public goods. By eliminating the threat of competition for a fixed period of time, innovative and enterprising individuals are able to recoup their investment in product development and reap the rewards of their ingenuity.

The patent system benefits society by providing a mechanism for new, technologically advanced goods to enter the marketplace and spur economic growth.

However, if patents were issued indiscriminately, the market could become overrun with monopolies that would impede the competition necessary for promoting technological and economic growth. In the absence of a competition-driven marketplace, the distribution of wealth would be inequitable, the prices of goods would remain artificially inflated, and there would be little incentive to improve existing products. To prevent rampant exclusivity, it is necessary to impose criteria for patentability so that incentives remain for collaborative efforts to expand general knowledge. Otherwise, technology, and in turn society, is unlikely to advance as rapidly. Jefferson sought to protect the public from monopolization that would impede innovation by formalizing the distinction between pure and applied research.

On the belief that exclusive ownership of scientific theories would hinder rather than promote innovation, Jefferson declared theoretical discoveries, or pure science, as ineligible for patent protection. While far from its intention, Jefferson's stipulation transformed the congressional sentiment toward scientific research from one of relative support to complete disinterest in any form of innovation that was not reducible to a "machine of potential cash value" (1). Congress, consistent with its aversion to bureaucratized government, used Jefferson's patent policy to free itself from any formal involvement in scientific research. It contended that the constitutional provision empowering Congress to issue patents limited the congressional role in innovation to that function. The federal government thereby distanced itself from basic science by declaring the subsidization of basic research unconstitutional and its findings ineligible for patent protection.

Jefferson, in contrast, believed Congress ought to exercise its constitutional power to encourage innovation by subsidizing the advancement of pure science. He understood that applied science would prosper in the hands of private enterprise because the potential for profit would attract investors. Pure science, however, would be less likely to prosper through private support because it offers a public good that has little profit potential. It was therefore in the nation's interest to invest pubic funds into pure scientific research because innovation ultimately stems from the progress of both pure and applied sciences.

HISTORICAL MILESTONES

Congress never failed to see the potential benefits of the scientific enterprise. To avert bureaucratization, it was simply preferable to fund science on an "as needed" basis rather than formalize an expensive, open-ended commitment. The occasional research project was conducted within the traditional units of a federal government that was focused almost exclusively on military and industrial needs. As a result federally funded research began through the Departments of Commerce and War.

Jefferson, when he was elected to the presidency in 1801, paved the way for military research by securing congressional support for the Lewis and Clark expedition. While convincing Congress that a technical exploration of the west was essential for military and commercial gain, he disclosed to the Spanish and French, who controlled western trade, that the mission was for scientific discovery only. The explorers, trained by the nation's scholars in topography, botany, zoology, agriculture, geology, astronomy, paleontology, and ethnology, returned with a body of knowledge so great that it was used to establish an American corridor to the Pacific and a trading presence in the northwest. The success of the expedition inspired Congress to fund, under the auspices of the War Department, further expeditions and requisite technical training. Arguably the earliest arm of the federal government receptive to scientific research, the early nineteenth-century American military gave rise to a national weather service, the first accurate observations of digestion, and research studies of diet and nutrition. (It is through the military that the National Academy of the Sciences was established in 1863, during the Civil War, and the National Research Council was formed in 1916, during World War I.) The research, however, was neither explicitly nor entirely publicly funded. While the government provided scientists with materials and laboratory space, direct funding was considered unconstitutional. As a result the processes of data analysis and publication depended on private support.

This mix of publicly and privately supported research did not confront the issue of ownership rights, for the scientists of the time, though increasingly specialized and respected, "had no aptitude for applying their knowledge to a downstream product (1)." Nevertheless, the potential relationship between the science reported in academic journals and the duty of the patent office was recognized. Henry L. Ellsworth, appointed the nation's first Commissioner of Patents in 1836, expanded the patent office into a "central depository" for innovation in the belief that the eventual merger of science and invention would serve the nation's economic interests (1).

Ellsworth collected patent application fees and used the revenue to acquire patented mechanical models as well as unpatented models, specimens, manufactures, minerals, seeds, and scientific findings culled from various collected publications. He drew from his library various solutions to the nation's agricultural crisis, becoming the first to use statistical analysis in large-scale problem assessment. In his promotion of a recently published German study calling for the use of chemistry to remedy soil exhaustion, Ellsworth recognized what was in fact "the first direct application of pure scientific research to agricultural success" (1). As well as devoting his career to promoting the utility of science, his efforts on behalf of the congressionally neglected farming community laid the foundation for the Department of Agriculture. In his striving to develop a great scientific bureau, the first Commissioner of Patents exceeded the traditional scope of both the patent office and the federal government.

Meanwhile, in 1838, S.F.B. Morse had presented Congress with his invention of the electric telegraph, another striking example of the convergence of pure and applied scientific advancement. Without a niche

in the nation's economy and few precedents regarding the development and ownership of his invention, Morse lobbied for four years to receive a $30,000 grant for construction of a telegraph line connecting Baltimore to Washington, DC (Congress justified the expense under the constitutional provisions for commerce and postal service). In 1845 Morse issued a proposal to extend the line to New York City, but it was voted down in Congress because the southern contingency thought it unconstitutional. Eleven years later the same fate was suffered by the electric motor; after an initial research grant, federal support was withdrawn despite extremely successful results.

Congress justified its recoil from research support with the contention that laying any groundwork for a research bureau would open a Pandora's box of new and unwanted federal responsibilities. It was argued that federally subsidized research would inappropriately risk public funds, while it forced the government into the role of examining agent for both the scientific enterprise and any devices it produced. In addition the deluge of research proposals requiring evaluation and, to exclude charlatans, investigation, would be too expensive and expansive to take on. Thus it came to be that the mid-nineteenth century would see Congress neglecting many opportunities to formally engage in the "industrial development of devices born from scientific discovery" (1).

At the same time Congress had stepped up its support of basic science through its technical explorations of the North American continent. Increasingly basic scientists were becoming government employees, but they were hired only for specific tasks and not for the construction or maintenance of a permanent research bureau. In 1834, when the Coast Survey, traditionally overseen by the Secretary of the Treasury, was reassigned to the Navy Department for cost-containment reasons, the scientists revolted by refusing to share their work with nonscientists. Two years later the Coast Survey was returned to the Treasury; the scientists succeeded in securing pay commensurate with their abilities while establishing civilian control of research.

As the power and prestige of the scientific community grew, congressional sentiment toward their research evolved accordingly. By 1842 Congress had even assumed responsibility for publishing the findings of a military expedition. However, the publication process was a fiasco and the finished product ridden with mistakes. It was clear that congressional committees were incapable of effectively administrating the highly technical enterprise that science had become. It is ironic that by the time Congress came to explicitly appreciate the utility of science, a government of nonscientists was no longer intellectually equipped to oversee its advancement (1).

INDUSTRIALIZATION

Despite the federal government's express regard for science, it had no intention of expanding its narrowly conceived administrative interests. By the end of the nineteenth century, industrialization, urbanization, and mobility had imposed vast and unfamiliar problems on a small-town, agrarian nation that had seen its traditional institutional framework erode (3). While Congress saw issues of social welfare as falling outside its sphere of responsibility, the early industrialists acted to address the issues that were a side effect of their success. Beholden to the Protestant work ethic, they developed philanthropic foundations in the belief that as stewards of their wealth they had a moral obligation to manage it in a socially redeeming way. Technology and standardization were the tools of choice for the industrialists-turned-philanthropists who strove to enhance social welfare in a manner that would serve their industrial interests.

Through addressing the educational and medical needs of the masses, the industrialists established a national infrastructure that would, for years to come, foster industrial and technological growth. Private universities, by educating middle class men in science and engineering, were able to equip an emerging professional stratum with the skills necessary to manage the future of industry. To advance the fields of science and engineering, universities and foundations funded the research of pure scientists as a means to further their respective disciplines.

The expansion of the scientific enterprise proved particularly beneficial to the field of medicine. Standardized and cost-effective, the application of science to medicine, called biomedicine, employed a reductionist approach to disease that complemented the political and economic agenda of industrial capitalism. While diverting attention from the environmental and social roots of working class illness, biomedical practitioners were able to restore health to the level of functionality necessary to labor toward capital accumulation (4). Consequently biomedicine was afforded tremendous industrialist support, for while it sustained a productive work force, its approach to research dovetailed nicely with the broader trend of scientific development.

By the 1920s biomedical research conducted throughout the nation's universities and foundations required coordination in order to make significant progress. Leading industrialists, as part of a larger trend of tying the nation's technologically driven "capitalist infrastructure to government function," requested that Congress act to centrally organize the nation's efforts in medical research (3). The government was open to the proposition. Despite its traditional reluctance to expand in organizational structure or function, it had been taking incremental, yet affirmative, steps toward establishing a formal commitment to national health for over a century.

The government's first intervention in public health came in response to the late eighteenth century scourges of Yellow Fever in the nation's port cities, for it had impeded federal revenue collection. The correlation between the presence of trading ships, communicable disease, and the effect on revenue, properly classified the matter as one of interstate commerce and hence one of the principle responsibilities of the federal government. The Department of Commerce enacted quarantine laws and established hospitals initially intended to care only for merchant seamen, and shortly thereafter for the military as well. By the twentieth century the government had organized a national, comprehensive medical research effort intended to investigate "the origin and causes of epidemic diseases" and established the Hygienic

Laboratory for bacteriological research in its Staten Island Marine Hospital (5). In 1912 the nation's Public Health Service was established to coordinate the government's involvement in issues of public health that quickly expanded to include numerous research studies on various facets of communicable and noncommunicable disease.

With the backing of powerful industrialists (who had already committed to large donations upon the success of the effort), scientific societies, the American Medical Association, and the life insurance industry, Louisiana Senator Joseph E. Ransdell, sought to establish out of the Public Health Service a central office to coordinate the nation's medical research. The 1930 enabling act of the National Institute of Health (NIH) changed the name of the Hygienic Laboratory to NIH and relocated it to a new building. Significantly, the act authorized the implementation of a graduate-level fellowship program to train young scientists in biomedical research. The Secretary of the Treasury was permitted to accept the philanthropic donations that would support these newly organized efforts in medical research. The congressional record of the enabling act, filled with rhetoric about doing God's work, is a reflection of the traditional interests and priorities of the federal government and the influence of industry in its decision making (6). In addition to arguing that science and medicine are intrinsic goods worthy of support, the industrialists convinced Congress of the relationship between health status and productivity to explain how health maintenance is in the nation's economic interest. In order to maintain national health, public and private efforts in research and data collection needed to be coordinated on a national scale. Further, with the nation still recovering from World War I, the fact that German progress in biomedical research appeared to be outpacing the United States served as an added incentive to engage more American scientists in government endeavors. Chemists, in particular, were understood as valuable to military and industrial research, but without an expansion in employment opportunities, they would be unable to apply their expertise to biomedicine. Fellowships, instead, would provide opportunities for this specialized graduate training. Last, perhaps the most attractive aspect of the Ransdell Act was that the NIH was expected to advance biomedicine ostensibly through private funds, with periodic support from the public treasury as needed (6).

BUREAUCRATIZATION

By the 1930s the promise of both public and private support of medical innovation made its potential appear boundless. Dupree's observation of the early republic, however, would remain applicable centuries later, for science is "a group activity carried on by limited and fallible men, and much of their effectiveness stems from their organization and the continuity and flexibility of their institutional arrangements" (1). The twentieth century would not only foster the growth of scientific research, but of corresponding organizational structures that would inevitably dilute its social and economic benefits. This process of bureaucratization permeated America's scientific enterprise whether conducted through foundations, the government, universities, industry, or combinations thereof.

While the foundations of the 1920s and 1930s funded scientific research, the scientists themselves were responsible for resource allocation. Over time, however, a stratum of science-administrators developed who, according to Barry Karl, were "not content in the function as mindless vehicles for the products of the academic mind" and so aimed to create their own professional identity (3). This pseudoprofessionalization of the scientific bureaucracy was counterproductive to advancement. Pure science was no longer conceptualized as simply a public good but as a means to enhance the careers and reputation of the administrators and their foundations. By the 1950s the foundation bureaucrats had taken to "reshaping the substance and form" of research proposals and placing undue demands on scientists who, due to a tight university job market, had few alternatives for employment (3).

During this period of the 1920s through the 1950s, the agenda of the federal government evolved to rely on both pure and applied scientific advancement. While an era of progressivism helped to focus congressional attention on public health needs, an effusive international agenda led to massive scientific undertakings, particularly in the areas of aeronautics and atomic physics. Starting with the National Advisory Committee for Aeronautics in 1920 and continuing through the Atomic Energy Commission and the National Aeronautics and Space Administration, new scientific agencies were established to exploit the quasi-military technology that would have remained uncultivated if development were left to the existing federal agencies and the private sector (7). It is clear that this organizational expansion of scientific support was initially justified by the technological needs of both World Wars and the cold war. Beyond assisting the military, however, the new agencies served the nation's economic interests and, in doing so, demonstrated that pure research is a viable investment in the nation's economic future.

The size and complexity of the government's scientific endeavors, particularly since World War II, required it to extend beyond its own laboratories to cull specialized expertise from the nation's universities, foundations, and industries. No longer was there pervasive opposition to such government expansion. Post–World War II organizational theorists heralded bureaucratization as a product of modernization, "a rational management technique that, while striving for effectiveness and efficiency, would leave the sources of intellect and power under popular control" (3).

By the 1960s the federal government became the nation's primary supporter of pure scientific research. A welcome employment alternative to the politicized foundations, scientists were eager to both receive government grants and organize the process of resource allocation. Like the foundations, however, the government produced a stratum of science-administrators whose professional aspirations would come to intrude on the enterprise itself. The function of government bureaucrats was to

reconcile the goals of the scientific community with the "intentions of Congress and sensitivities of its committees" (3). This differentiated them from the foundation bureaucrats who, at the very core, remained guided by the mission to serve humanity. Government bureaucrats had no such moral agenda but instead endeavored to engage in political maneuvering as a means to climb a burgeoning bureaucratic hierarchy. Karl attributes to T.S. Kuhn the observation that science-administrators are essentially "manipulated into producing results that are determined elsewhere; they have no interest in the process beyond its efficient progress toward its stated goal" (3). Karl explains that the role of such administrators is not intended to interfere with knowledge development, for an administrator's knowledge of the discipline he manages may be limited to that which is directly related to his job function.

This concept of isolating the administrator from knowledge development proved problematic for academic institutions whose researchers were the recipients of the government's expanded support of science. A need developed for a "scholar-administrator" to mediate between the university's internal needs and interests and the increasing external demands placed upon it. Karl explains that university faculty members discovered quickly that "even a relatively brief career as an administrator could jeopardize their scholastic careers" (3). Academic research science was faced with the challenge of finding administrators who, in addition to having basic institutional affection and technical skills, had a professional understanding of the discipline to be managed (3). The typical manager of academic research was either a "critical compromise" between the faculty and university administration, a rejected tenure candidate, a doctoral candidate who never completed a dissertation, or a legal or business professional (3).

As the practice of federally funded, privately managed research grew, it introduced an unprecedented level of government intrusion into university matters (e.g., investigations for equal opportunity and audits for accountability in the use of public and private resources). Left with the very nebulous distinction between public and private affairs, the university relied on its bureaucracy to protect the once independent, private academic system from imposition by the government. Similar to the experience of the foundation and the government, the expansion of academic research institutionalized the bureaucrat into intellectual life (3). Over time these bureaucrats became permanent stakeholders in the American scientific enterprise and became the individuals largely responsible for the mechanics of technology transfer.

DEVELOPMENT OF TECHNOLOGY TRANSFER

In his 1968 book *The Government of Science*, Harvey Brooks discusses public concern over the priorities of the federal government (7). During the 1960s the government was perceived to have "deprived the civilian economy of its sources of technological innovation" with its cold-war preoccupation with military and space technology, while neglecting to apply this innovation to the improvement of social welfare (7). Brooks attributes the government's reluctance to resolve social concerns, such as pollution, urban transportation, and disease, in part, to its inability to reach consensus regarding: (*1*) the degrees to which issues are federal responsibility, and (*2*) the means and ends of proposed solutions to civilian problems. He believes that the feasibility of general agreement on the process and goals of research and development determines how the government uses technological progress to address social needs (7).

Brooks argues that direct subsidization of research and development is an effective means for developing highly technical areas, such as defense and space, since, in reality, their advancement rests on the consensus of only a small number of technocrats. In contrast, civilian technology requires numerous parties, often with divergent interests, to strike a compromise on who, where, when, and what is to be sacrificed in the interest of the general good. To illustrate his point, Brooks describes what the installation of an urban rapid transit system would entail if completed exclusively on government subsidy. Test results of the technology and system would constantly have to be weighed against public opinion. Not only would the process be exceedingly time-, labor-, and cost-intensive, it would be wasteful as well if the result was a system people elected not to use. Brooks questions whether "technical-economic analysis is sufficiently refined to justify large gambles with public funds" (7). Ironically a response to this question requires a public consensus that is, for the most part, impossible to reach.

Many believe that the marketplace is the most efficient and effective tool with which to measure mass opinion of a public good. In addition to turning out the most workable solutions to public demand, it produces the incentives necessary for the demand to be satisfied through private investment. It is widely contended therefore that the government need only provide a "framework of information, incentives, and underlying general technology" for an entrepreneur who, by responding to the marketplace, can accomplish the goals of public policy without risking public funds (7). This indirect subsidization of research and development is often heralded as the means through which innovative technology born from public funds ought to return to the public in the form of social goods. In order for it to be an effective method of technology transfer, however, the government must offer sufficient incentives so private developers will invest in bringing federally funded inventions to market.

The federal government's World War II expansion of research support led to an investigation of the ways in which it could promote private development of its discoveries while, at the same time, retaining an unrestricted right to their use. The Roosevelt administration observed that while the practice of publication protects the government's rights to discoveries arising from its research, those rights would be better secured through patent protection. A 1945 advisory report by the National Patent Planning Commission recommended that the government patent its inventions, but retain exclusive ownership of the titles only in

cases where private ownership would be detrimental to national welfare. The commission advised the government to make its titles generally available, on a nonexclusive basis, to anyone wishing to develop inventions into public goods. It recognized, however, that certain circumstances require the government to offer exclusive ownership rights in order to "induce private manufacturers to commercialize an invention" (8). Roosevelt's commission therefore recommended that exclusive licenses be issued in cases where it was reasonable to assume that the invention would otherwise remain idle.

Rebecca Eisenberg, in her review and analysis of the U.S. history of technology transfer policy, explains that the aim of transferring title ownership from the government to the private sector is to promote economic prosperity by "stimulating innovation, new products and new jobs" (8). Many believe that in order to effectively promote industrial growth, the government must have the freedom to grant private developers exclusive ownership of inventions made at public expense. This is controversial, however, for while exclusive private ownership may be an effective method of putting publicly funded technology to practical use, its method of empowering industry for the promotion of large-scale economic growth often entails the sacrifice of other social goods, such as the sharing of scientific knowledge. On the other hand, the absence of a provision for exclusive licensure would sacrifice potential technological and industrial advancement. A technology transfer policy based solely on nonexclusive licensure would dissuade private developers from investing to develop government inventions and discoveries because, in some cases, their competitors could "copy successful inventions without having shared in the initial cost and risk of making them" (8). What would likely happen is that each developer would generate a portfolio of improvement patents covering specific applications of the basic government-funded invention. The threat of competition in the race to develop marketable products, however, increases the risk of development and thus lowers the value of the basic invention. As a result the nation's best firms might refrain from involvement in government innovation and many publicly funded inventions would remain undeveloped.

In 1947 the U.S. Attorney General, Robert H. Jackson, issued a recommendation for technology transfer that underscored the drawbacks of conferring exclusive ownership rights to private parties. His report called for the government to retain, with few exceptions, exclusive ownership rights to all inventions funded in part, or in whole, with public funds. To encourage the development of these inventions, he recommended that rights to government inventions be licensed to private parties on a nonexclusive basis only. In cases where development hinged on the provision of exclusive rights, Jackson contended that the government itself should finance product development rather than endorse the monopolization of a publicly funded invention. He further advised the government against charging royalties for the use of this technology.

The recommendation of the Attorney General was intended to encourage the commercialization of government-held inventions while, at the same time, protecting the public's equitable claim to the technology created at their expense. When a private enterpriser has exclusive rights to a publicly funded invention, the public is required to "pay twice for the same invention — once through taxes to support the research that yielded the invention, and then again through higher monopoly prices and restricted supply when the invention reaches the market" (8). This hurts both consumers and small business because it concentrates innovative technology, and its attending economic power, in the hands of beneficiaries of "government favoritism" (8). The practice of issuing exclusive licenses demands that the government undergo a certain level of bureaucratic expansion in order to orchestrate an application process, "select a licensee, police its operations, and detect and prosecute patent infringement" (8). If the licensees are continually large technology firms, entrepreneurs of limited means are prevented from competing in an increasingly technologically driven global market. As a result the practice of doling out exclusive rights to government inventions may contribute to a progressive concentration and centralization of power. Nonexclusive licensure, on the other hand, would allow publicly funded inventions to be used by many firms, thereby introducing them to a competitive, as opposed to monopolistic, marketplace. This would somewhat level the playing field for small business and benefit consumers through reduced product costs.

The congressional response to the 1945 and 1947 advisory reports for technology transfer policy was to refrain from enacting any governmentwide policy for over 30 years. From the 1940s to the 1980s, the federal agencies involved in research were broadly encouraged to license their inventions to private developers. Actual policies, however, were instituted only in response to particular agency-specific issues. It was believed that the tremendous disparities between federal agency missions, collaborator agendas, and the type and commercializability of federally funded inventions, made a standardized policy both "unfeasible and undesirable" (9). As a result much of the technology transfer legislation during this time was directed toward authorizing federal agency heads to manage collaborative research and development efforts in whatever manner best suited their agency's operations. The perception that disparities in agency needs and practices precluded the institution of a uniform policy was supported by a 1965 study commissioned by the Committee on Government Patent Policy. The study found collaborator decisions regarding whether or not to invest resources in government research and development to rest primarily on the commercial potential of specific research endeavors and inventions, as opposed to the particulars of licensing agreements. In fact, until the 1980s, overall commercial utilization of government-sponsored inventions was "very low, regardless of who held the title" (8). The 1965 study, while reluctant to issue a blanket recommendation, asserted that in some cases the provision of exclusive rights would promote the development of inventions better than acquisition of title by the government.

The Kennedy administration moved to standardize technology transfer policy by issuing a memorandum

outlining those circumstances in which the federal government would retain ownership rights to patent titles, and those where rights should be licensed to private developers. The situations in which the government was to retain title ownership included (*1*) when the products of research were intended for the public's health, welfare, or commercial use; (*2*) when the contractor was coordinating a government owned facility or operation; and (*3*) where the government was the principle developer or leading authority in the field of expertise and granting exclusive rights to a contractor would designate that contractor as the dominant figure in the market. Eisenberg explains that the contractor, in turn, is to acquire ownership rights to an invention when exclusive rights are essential for development, and

> . . .where the contract research is to build upon existing technology to develop information, products or processes for use by the government, and the contractor has acquired technical competence and established a nongovernmental commercial position in the field, the contractor would normally acquire title, subject to a non-exclusive, royalty-free license in the government (8).

Although Kennedy, and later Nixon, took affirmative steps to improve and standardize technology transfer policy, their attempts toward establishing governmentwide uniformity were negated by designating agency heads, with their disparate policies, to administer the practice. Nevertheless, these efforts served to improve government assessment and oversight of federal patenting and licensing practices. Significantly, exclusive owners of publicly funded inventions were required to issue progress reports on their commercialization process. If, after three years, they failed to take reasonable steps to bring their invention to practical application, the government was entitled to terminate the contractor's right to exclusivity.

The tragic flaw of the technology transfer policy from the 1940s to 1980 was not that it varied from agency to agency but that it failed to provide adequate incentives for government contractors and grantees to pursue research with commercial potential. In their paper, *Technology Transfer Laws Governing Federally Funded Research and Development*, James V. Lacy, Bradford C. Brown, and Michael R. Rubin attribute the large numbers of government-owned, unlicensed patents to a "lack of statutory basis for royalty sharing" (9). They argue that the absence of a legal provision that entitled government collaborators to a portion of the profits generated by their invention, denied researchers any incentive to create "commercially viable technology" (9). Whereas private sector researchers were motivated by goal structures and profit-oriented management techniques, government researchers and grantees were motivated by salary alone (9). Consequently one of the key problems with technology transfer during this period was not simply the low commercial potential of many government-held inventions, but the lack of incentive for government employee-inventors to transfer any inventions to the private sector for development.

In the late 1970s Congress, in an attempt to improve the nation's low economic productivity, set out to resolve the deficiencies of technology transfer. A "series of bipartisan initiatives" were enacted to "revise government patent policy, reduce legal and bureaucratic barriers, and create incentives to improve federal technology transfer to the private sector" (9). At last, Congress understood that "it was not enough to fund, invent, and patent inventions. The government had to actually make its way into the market in order [for technology transfer policy] to produce positive economic results" (9).

In passing the 1980 Bayh-Dole Act, Congress sought to improve the practice of technology transfer by aligning federal research policy with the nation's economic needs. The Act entitled small business firms and nonprofit organizations collaborating in government research to retain ownership rights in subsequent inventions. According to Eisenberg, the conspicuous omission of large firms was a reflection of the Carter administration's "strategy for improving the industrial competitiveness of the nation" (8). She explains how many of Carter's supporters believed small business to be "innovative, adaptive, risk-taking, entrepreneurial and competitive, yet [inequitably] burdened by the practice of obtaining case-by-case waivers of title from sponsoring agencies" who were traditionally reluctant to grant research funds and patent rights to small businesses (8). This made it difficult to compete with large firms that during this time period, Eisenberg recounts, were often painted as "short-sighted, risk-averse, and predatory — more likely to suppress new technologies than to adopt them, yet savvy and powerful in their dealings with government agencies" (8). It was a hallmark of the Carter presidency to blame the nation's large firms for the decline in the global position of U.S. industry. Thus a policy for technology transfer that promoted small business growth as a means to enhance the American marketplace was a reflection of the Carter era. By 1984, however, the Reagan administration, with its markedly different economic agenda, extended the provisions of the Bayh-Dole Act. This revision, which remains in effect today, entitles *all* private enterprisers in government collaboration, including large firms, to own any inventions generated in whole or in part with public funds.

Through the Bayh-Dole Act, and related legislation throughout the 1980s, Congress was seeking to improve the competitive position of the United States in world markets. The ideal policy for technology transfer would see to it that every dollar invested in scientific research would, in essence, be a dollar invested in national economic prosperity. In order for the policy to meet this objective, Congress modified the earlier system of technology transfer in three key areas. The first was discussed previously: Policies were instituted so that private parties contributing to publicly funded research would retain the right to develop any subsequent inventions. The second was the establishment of an incentive system to motivate the employees of government-owned, government-operated laboratories to make and license commercializable inventions. Third, a legal basis was provided for favoring American over foreign industry in conferring ownership rights to publicly funded technology.

The policy motivating government agencies and employees to invent and license technology was established as part of the 1986 Federal Technology Transfer Act, which made technology transfer a top priority for agencies involved in research. Employee-inventors are now required to actively seek licensees for their inventions and are evaluated on their ability to do so. In certain cases the inventors are permitted to assume ownership rights and pursue commercialization. A system for royalty-sharing was created that gives the agencies and employees of government-owned, government-operated laboratories a financial stake in the inventions they create. Royalties from commercialized inventions are collected by the sponsoring agency and shared with the employee-inventors. A portion of the remaining revenue is put toward the inventing laboratory's budget for the next year and the rest used for activities that encourage technology transfer within the agency.

Particular provisions of the Bayh-Dole Act ensure that the economic benefits of federally funded research are enjoyed primarily by the United States. Agencies are to favor U.S. industry when a developer (*1*) is not located in the United States, (*2*) does not have a place of business in the United States, or (*3*) is subject to control of a foreign government. Further, in order for a developer to assume exclusive rights to an invention, the "products embodying the subject invention or produced through use of the invention" must be manufactured substantially in the United States (9). Exceptions are made, however, if domestic manufacture is either patently infeasible or not possible at the time. Violation of these terms entitles the government to terminate the licensing agreement.

The Bayh-Dole Act also provides the government with residual rights to all publicly funded inventions to ensure its access to the technology in certain circumstances. The government retains the freedom to employ a licensed invention, royalty-free, for its own use or on behalf of a foreign organization and federal agencies are permitted to retain additional rights. To maintain fairness in the marketplace, the technology transfer policy aims to minimize the monopolization of publicly funded inventions by encouraging the use of nonexclusive licenses. Agencies are only permitted to issue exclusive licenses when they are proved to be in the best interest of the public. Exclusivity is beneficial only when it is a necessary incentive for development, and does not threaten competition or concentrate a particular technology in a specific geographic area. To further protect fair competition, when federal agencies issue licenses, they are required to give first preference to small businesses who have adequate resources for successful commercialization. In addition the government is entitled to exercise "march-in" rights and terminate the exclusivity of a contract if (*1*) The licensee has taken, or is not expected to take in a reasonable amount of time, effective steps toward developing an invention, (*2*) requirements for public use specified by federal regulations are not being reasonably satisfied by the licensee, or (*3*) action is necessary to alleviate health or safety needs that are not reasonably satisfied by the licensee. These march-in rights have never been exercised.

PUBLICLY FUNDED RESEARCH AND BIOTECHNOLOGY

The current policy of technology transfer is a reflection of the federal government's traditional conceptualization of scientific research. Although it offers an array of societal benefits, the primary reason science receives extensive federal support is that its advancement has become vital to national military and economic progress. Thus biotechnology, which is believed to hold great potential for the field of medicine, is primarily conceptualized by the government as a means to promote industrial growth. This is reflected in the 1989 argument by Michael A. Andrews before the House of Representatives in a plea to secure future funding for the Human Genome Project:

> The United States has a soaring trade deficit. We are slowly awakening to a growing weakness in international competition.... The Japanese are developing automated sequencing devices. The English have almost completed the mapping of the roundworm genome. The West Germans and the French have set up international reference data banks to collect the results of genome research. International competition has often spurred the United States into action on major scientific endeavors: Sputnik caused us to put a man on the moon, World War II brought about the Manhattan project.... I believe that international competition will shore up a commitment of the United States to the Human Genome Project more than any other single factor.... One by one, we have watched the pillars of our economy fall: the steel industry, the auto industry, the electronics industry, and the energy industry. Biotechnology is one area where the U.S. can have a clear lead (10).

The government apparently heeded Andrews' advice. Not only was extensive funding secured for the Human Genome Project, but by 1994, federal laboratories were the nation's leading inventors and enablers of new technologies for the biotechnology and pharmaceutical industries (11). The Public Health Service continues to lead the nation as the organization with the largest number of "therapeutics in active development, both in terms of those licensed out and those being developed internally" (11). This vast federal support of biotechnology raises special concerns because discoveries in this field hold value beyond their contribution to industrial growth. Biotechnological innovation carries the potential to improve the lives of the public, who, through their investment of tax money, and by serving as clinical research subjects, make medical research possible. Arguably, because the public is so uniquely invested in the products of medical research, particular aspects of the technology transfer policy ought to be reconsidered.

The absence of price controls, for example, allows private developers to set the price of publicly funded medical technology beyond what portions of the population can afford. In other words, current technology transfer policy confers on private parties the right to ration publicly funded therapies according to what the market is willing to pay. The government refrains from exercising its right to "march-in" and terminate exclusive licenses on behalf of the public health because the ultimate goal of technology transfer is not necessarily to improve social well-being but to serve the nation's economic interests.

The circumstances surrounding the implementation and repeal of the "reasonable price clause" illustrate this point. In 1984, when HIV was identified as the virus that causes AIDS, a screening program was initiated where drug companies submitted shelved drugs to NIH for testing against the retrovirus. Burroughs-Wellcome submitted a drug called AZT that was invented in 1964 by the National Cancer Institute (NCI). In 1985 AZT was deemed effective against HIV in vitro and was therefore worthy of clinical investigation. With NCI supplying the thymidine necessary for AZT production, Burroughs-Wellcome provided NIH with AZT to run the clinical trials necessary for FDA approval. A year later AZT, having successfully prolonged the survival of AIDS sufferers, was approved by FDA. To promote its development, Burroughs-Wellcome was granted a seven-year exclusive marketing privilege and patent rights until 2005 for its use in the treatment of HIV. AZT was introduced to the public in 1987, at the exorbitant price of $10,000 to $12,000 per patient/per year (12).

Outraged, the public demanded to know why the price of AZT was so high when both the initial discovery and later recognition and research of its modern application were publicly sponsored. In his recounting of the AZT controversy, Baruch Brody asks whether "the public's need for the drugs [is] being served by allowing drug companies to charge that much for drugs?" (12). A plausible argument can be made that the needs of the public should take precedence over the promotion of technology transfer. In response, some would argue that technology transfer never actually takes precedence over public needs because it serves those needs in the long term. Allowing companies to set high prices is society's way of rewarding them for transferring the public's scientific research investment into important public goods. The question arises, however, whether a public good is provided when a product is largely inaccessible to its sponsors.

In 1989 the government sought to improve access to products of technology transfer by promulgating a "reasonable price clause" in exclusive licenses arising from Cooperative Research and Development Agreements between government and industry. The clause mandated that the price of inventions must reasonably reflect the health and safety needs of the public and their investment in the product (13). In his explication of the reasonable price clause, Brody outlines its numerous presuppositions. For one, it assumes that the government is entitled to a financial return on the intellectual property rights conferred to private developers. Reducing product prices according to the degree of public subsidization will result in Medicare and Medicaid savings. In addition the clause assumes that access to the products of technology transfer should not be determined by price alone but weighed against public health and safety needs. Last, the reasonable price clause implicitly assumes that its price control measures would not deter private developers from investing in publicly funded inventions. Brody believes that the assumptions contained in the reasonable price clause are problematic in that a pricing policy based on the degree of public funding may yield an entirely different price than a policy focused on accessibility. Furthermore, it is presumptuous to assume that either approach offers the return on investment necessary for private developers to engage in technology transfer (12).

In 1995 the reasonable price clause was repealed, for numerous reasons. NIH claimed that it was not only difficult to enforce but that it had a chilling effect on industry–government research collaboration. The repeal was clearly influenced by the pharmaceutical/biotechnology lobbyists who were vehemently opposed to price controls, by the patient advocacy groups who were incensed over a possible delay of new products, and by a Congress that was quick to abolish regulations without instituting safeguards to secure the public's equitable claim to the products of technology transfer.

The resulting policy omission exemplifies the federal government's prioritization of economic interests over any genuine commitment to improve the fundamental problems of the nation's (and the world's) access to health care. As patients, providers, and payers alike struggle with the burgeoning costs of medicine, the current technology transfer policy propagates the belief that "the proliferation of new technology developed at public expense is an unqualified good" (14). William Sage questions what he has dubbed "the conventional wisdom" of the policy in light of its contribution to health care inflation (14). He points out that most new products introduced by the biotechnology and pharmaceutical companies neither prevent nor immediately cure illness. They tend, rather, to "palliate suffering and prolong life" which, coupled with their high price tags, serves to funnel a significant portion of increasingly scarce health care dollars into industry pocketbooks (14). The government incurs the escalating expense of medical technology through Medicare and Medicaid. To contain costs, it chooses not to question the cost-effectiveness of new technology but rather to reduce provider reimbursement rates and tighten Medicaid eligibility criteria. Whereas it could fund cost-effectiveness studies with the royalties collected from technology transfer, the government has little incentive to do so because any policy that restricts the market for new technology would be detrimental to the pharmaceutical/biotechnology companies that are regarded as key to national economic prosperity. Private insurers, on the other hand, and the employers and individuals who pay insurance premiums, are increasingly reluctant to cover expensive technology when cost-effectiveness has not been demonstrated. At this point, however, programs to evaluate cost-effectiveness are too expensive and controversial to permit their widespread implementation.

For the 44 million Americans without health insurance, the industrialization and bureaucratization of health care has rendered rudimentary care, let alone technologically advanced treatment, virtually inaccessible. Sage argues that to retard the trend of using scarce health care dollars to purchase products that, in the long run, are a cost rather than a benefit to the system, new medical technologies, particularly those arising from public funds, should be brought to market only after determined to be cost-effective (14). Whether an invention's cost-effectiveness is sufficient for it to be brought to market

ought to be based on a societal consensus to purchase the technology for all, regardless of their ability to pay. To increase access, royalties could be used to subsidize the cost of technology for the indigent. Additionally the government could require private patent holders to provide their product to the disadvantaged gratuitously or at a discount. Sage concedes, however, that it is unlikely these suggestions would adequately offset "the incentives for unbridled innovation and consequent cost pressures created by the current technology transfer policy" (14).

Whereas royalties can potentially patch some of the holes in the system, some believe they create more problems than they solve. It is unclear, for example, why the government collects royalties from licensed inventions if the objective of technology transfer is to promote the private development of public inventions. Royalties make publicly funded technology a less desirable investment because they function essentially like a tax on development that reduces the overall profit potential of government-funded technology and increases the risk of the investment. In the end it is the consumer who bears the cost of royalty agreements because product prices must offset royalty expenditure, and higher prices reduce the accessibility of public goods. Again, the public pays twice for inventions.

The Bayh-Dole Act encouraged university ownership as a means to promote academe–industry collaboration. Underlying this decision was the perception that universities are well suited to determine which research results ought to be patented and developed, and which would best serve the interests of science as part of the public domain. University–industry collaboration appeared to optimize the potential for scientific advancement. Industry would provide university researchers with incentives to generate inventions that would benefit society in the form of public goods and economic productivity. However, in order to simplify their administrative burden, universities tend to prefer to grant exclusive rather than nonexclusive licences, which undermines the rights of the public to relatively unrestrained access to publicly sponsored inventions and discoveries (15).

The financial return on commercialized inventions would provide future funding for innovative academic research. The collection of royalties by universities therefore may provide a much-needed source of revenue for sustaining research initiatives and other institutional needs. It would be unwise, however, for all universities to rely on such revenue, since it is only a minority that can generate a sufficient amount of royalty revenue to reliably sustain institutional functions. Further there is some debate over whether the patents to publicly funded inventions should have been transferred to universities in the first place because they offer no advantage over the government in that both are unable to develop inventions into public goods.

The university–industry research collaboration has changed, and it is likely to continue to change the culture of academic science. It is a tradition in academic research to promptly share information with the scientific community with the incentives for doing so largely taking the form of professional accolades for furthering the advancement of knowledge. Commercial sponsorship threatens this tradition by motivating researchers to privatize the knowledge they generate in order to ensure both patent eligibility and gain a competitive edge. This trend is reflected in a study conducted by David Blumenthal that found biotechnology faculty with industry sponsors to be more than four times as likely as colleagues without commercial support to report that they had kept research results a secret in order to protect their proprietary value (16). His study also found these researchers to be five times as likely to report than they had conducted research the results of which were the property of private sponsors and could not be published without the sponsor's consent (16), reflecting a serious sellout of academic freedom. The incentive to patent research results increases the likelihood that researchers will engage in the lengthy patent application process, during which time they may be reluctant, or even contractually forbidden, to share information regarding their inventions. Thus the industrial support of academic research imposes barriers that hinder the sharing of knowledge among colleagues. Yet it is this cooperation that has traditionally advanced science.

In biotechnology patentability is still a gray area, and publicizing information, such as sequence data, may render an invention "obvious" and thus ineligible for patent protection. When patents are granted in young fields like biotechnology, earlier inventions are often granted broader patent coverage than those that follow because "patent claims are drafted to encompass not simply what the inventor has done, but the idea which underlies the specific detail. Sometimes, in the absence of much detail, patent rights may be granted which many regard as excessively broad" (17). The uncertainty that surrounds biotechnology patents is exemplified by the 1988 patenting of the Harvard OncoMouse, a transgenic mouse produced for carcinogenicity testing (18). The OncoMouse patent raises numerous questions, including whether the patent was limited to transgenic mice, or whether it extended to transgenic rodents, or even to transgenic mammals in general (19). When the boundaries of intellectual property are nebulous, the stage is set for complicated turf battles that will likely send potential developers in search of investment opportunities with less potential for complication.

Broad patent claims, when applied to basic research tools like the OncoMouse, may chill entire areas of research. Licenses to use patented research tools may be unaffordable to some institutions, or researchers may be unwilling to purchase licenses that would essentially allow their research to be shaped by the interests of the patent holder. By owning and exploiting the rights to research tools, industry can gain considerable control over the nation's research agenda. This was demonstrated in the controversy over Cre-lox (20), a recombinant technology owned by DuPont that was used freely in NIH genetic research for years (17). In what was thought to be a prudent business maneuver, DuPont began to require researchers to purchase licenses for the use of Cre-lox. Recognizing that these licenses would be unaffordable for some institutions, DuPont permitted researchers to use Cre-lox with the understanding that DuPont would retain

ownership of all their inventions that either incorporated the technology or used it somewhere along the line. Researchers were prohibited from transferring Cre-lox technology to unlicensed colleagues and were required to send all Cre-lox-related papers to DuPont for review prior to publication. Ultimately the research community struck a less restrictive agreement with DuPont for the use of Cre-lox, but the circumstances reflect that the strong presence of industry, and its ownership of technology, can wield such power over research that the government has to intervene, a heretofore unprecedented action.

Companies are involved in the development of products when they are still in the early stages of research. Researchers, including those working under federal grants, often receive private compensation for their ideas in the form of income, equity interest in the developing company, a seat on its board of directors, or a percentage of future sales (14). Admittedly, scientists are entitled to benefit financially from their inventions, and the technology transfer system is constructed to tolerate personal gain at taxpayer expense in order to encourage the development of public goods. However, when universities and individual faculty members have a direct profit motive to invent a product, it "represents a clear departure from past practices, and creates real risks for universities and the functions they are designed to serve in society" (14). When academic research begins to take on market characteristics, competition may tempt scientists "to circulate misinformation about a project's likelihood of success or even to commit outright fraud" (14). In the past it appeared harmless to allow researchers to earn royalties from their inventions because this occurred only in exceptional cases and after a prolonged period of research and development. Today a financial stake has been driven through the lab bench by companies who pay universities and researchers large sums of money up-front for the long-term rights to their inventions. Sage explains that "when the enrichment of scientists is directly related to the success of the scientific endeavor, society runs the risk that researchers will knowingly influence the outcome of neutral scientific inquiries" (14). The compromise of scientific integrity for greater reward violates, particularly in the case of medical research, the scientist's fiduciary duty to the public who entrusts the research community with their health and safety (14).

Royalty incentives, much like researcher equity positions in end products, comprise a facial conflict of interest. The federal government, however, relies on these incentives to drive technology transfer (21) and so excludes them from federal conflict-of-interest regulation. The use of royalty incentives in the absence of safeguards to protect the public from unsound science questions whether the government is using the public's funds responsibly and in the interest of their welfare. The pervasive use of royalty incentives raises the additional concern that if the scientific community is motivated by financial reward, it may cause basic research, because its products lack immediate commercial potential, to become undervalued and underdeveloped. Basic science, however, is essential to societal advancement. It has been understood since the nation's inception that the greatest breakthroughs often occur out of investigations into uncertainty. The information derived from basic research is the foundation for applied research which, in contrast, is defined by certainty in that it usually requires a set of "unambiguous facts" and specific targets toward which to work (22). There is a prevailing fear that as the industrialization of science progresses, the breadth of scientific inquiry may narrow to areas with foreseeable market potential. Thomas Jefferson believed that the government is obliged to promote innovation, and it ought to do so by subsidizing basic science because the goods it produces, and the incentives it requires, cannot be adequately sustained by the private sector. In that case the driving of publicly funded research toward the invention of commercializable goods is not what the government's role in science should be. Without the comprehensive support of basic science, the socioeconomic structure of the nation may suffer as it will be denied the benefit of scientific uncertainty and its resulting discovery. Arguably, the government tilts the nation's research agenda away from basic science through the policy of technology transfer.

If the government funds basic research, then what happens to the results should depend on the commercial potential and usefulness of the results. For example, some space research might hold military importance, some long-term strategic importance (e.g., claiming bases on Mars for future uses), and even some potential commercial utility (e.g., for commercial satellite, communication, or other types of businesses). Other research might be pure science performed in an attempt to understand the universe. Arguably, if the public sponsors the former, then some mechanism should be found to give a commercial preference to domestic companies, which will, in turn, return direct benefits to the country by employing citizens and paying taxes. In the latter case, however, the intellectual contribution of understanding black holes may not and should not be considered property but knowledge. Knowledge should be more freely shared with all. The government, through its policy for technology transfer, promotes the patenting and development of technology that provides society with public goods and economic productivity, but it does so at the expense of the collaboration necessary for overall scientific advancement (23). This is consistent with the federal government's conceptualization of science. Because it is valued as a means to spur economic growth, the measures used to promote this growth may threaten the proliferation of knowledge is not such a pressing concern. This view is rather short-sighted because it is indeed the proliferation of knowledge that ultimately gives rise to overall societal advancement.

Aspects of the technology transfer policy run counter to the ideal of a "more perfect union" where the federal government fulfills its constitutional duty to "establish justice," "ensure domestic tranquility" and "promote the general welfare" (2, preamble). The investment of public funds in the life-enhancing field of medical research, without ensuring the public's equitable claim to resulting therapy, is arguably unjust, despite the aggregate economic benefits. An expensive, bureaucratized, and largely inaccessible health care system is not in the interest of the general welfare, yet aspects of technology transfer perpetuate these systemic flaws.

As this review shows, the evolution of technology transfer policy reflects a difficult balancing between promoting development and use of government-funded inventions and discoveries, and protecting public welfare. One constant throughout this historical account is the military or economic justification for government support of research. A second constant is that as science becomes increasingly complex, it is becoming industrialized and, as a result, bureaucratized. All the while, however, it is not necessarily becoming more accessible to the public. A third constant is the concern about monopoly: Exclusive rights can benefit the powerful at the expense of the public, but exclusive rights sometimes are the best way to move inventions off the shelf and into the market.

BIBLIOGRAPHY

1. A.H. Dupree, *Science in the Federal Government: A History of Policies and Activities to 1940*, The Belknap Press of Harvard University Press, Cambridge, MA, 1957.
2. United States Constitution, Art. 2, §8, cl. 8.
3. B.D. Karl, *Daedalus* **105**(4), 129–149 (1976).
4. H. Baer, *Soc. Sci. Med.* **28**(11), 1103–1112 (1989).
5. *NIH Almanac*, September 1998.
6. Congressional Record H 9180-9183 (daily ed. May 19, 1930), Congressional Record S 9258-9265 (daily ed. May 21, 1930), Congressional Record S 10358-10362 (daily ed. June 10, 1930), Congressional Record H 10950-10952 (daily ed. June 16, 1930).
7. H. Brooks, *The Government of Science*, MIT Press, Cambridge, MA, 1968.
8. R. Eisenberg, *Virginia Law Rev.* **82**, 1663–1726 (1996).
9. J.V. Lacy, B.C. Brown, and M.R. Rubin, *Pepperdine Law Rev.* **19**, 1–28 (1991).
10. 101 Congressional Record H, E1418 (1989).
11. NIH Fall 1994 press release/study. Originally released 8/22/94; revised 9/8/94 including new data for 43 NIH CRADAs. <http://www.bioinfo.com/fed_biotech_transfer_study.html>
12. B. Brody, *Hastings Center Report* **26**(2), 5–11 (1996).
13. *NIH Backgrounder*, April 11 (1995).
14. W. Sage, *Virginia Law Rev.* **82**, 1737–1752 (1996).
15. A. Schissel, J.F. Merz, and M.K. Cho, *Nature* **402**, 118 (1999).
16. D. Blumenthal, *J. Am. Med. Assoc.* **268**, 3344–3346 (1992).
17. D. Wood, *Whose Life is it Anyway*, InterPharma (1996). <http://www.interpharma.co.uk/iss3.htm>
18. U.S. Patent No. 4,959,317 (1988) Site-specific recombination of DNA in eukaryotic cells.
19. E. Russo, *Scientist* **12**(18), 1 (1998).
20. U.S. Patent No. 4,736,866, Transgenic nonhuman mammals.
21. M. Witt and L. Gostin *J. Am. Med. Assoc.* **271**(7), 547–551 (1994).
22. T. Lewis, *The Lives of a Cell: Notes of a Biology Watcher*, Viking Press, New York, 1974.
23. *The Organization of the National Research Council*, Available at: <*www4.nationalacademies.org/arc.nsf*>

See other entries HUMAN GENOME DIVERSITY PROJECT; MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT; OWNERSHIP OF HUMAN BIOLOGICAL MATERIAL; see also PATENTS AND LICENSING entries.

PATENTS, ETHICS, HUMAN LIFE FORMS

LOUIS M. GUENIN
Harvard Medical School
Boston, Massachusetts

OUTLINE

INTRODUCTION

It might be supposed that morality operates as a side constraint on patentability. On this view, even though a process or device might meet conceptual and scientific criteria for recognition as an invention, moral considerations might override so as to deny a patent. Or again, it might be held that morality, genetics, and biotechnology so intertwine that whenever we construct criteria of patentability with respect to "genetic inventions," we perforce impose some moral view.

EXPLOITATION AND MONOPOLY OF CLONED DNA SEQUENCES

Whereas U.S. patent authorities formerly declined to issue patents on gambling devices and phony medicines, the U.S. Patent Code of 1952 dissociates law and morality. It leaves to other laws the matter of restraining the use of inventions. Efficiency alone commends this division of labor, since many patents are never exploited. For example, Pasteur obtained a U.S. patent in 1873 on a yeast for making beer. But, so far as we know, he never developed a commercial product (1). If immorality of use does not count as an objection to a patent application in general — even if the proffered device be mischievous — should immorality of use count as an objection to patents on life forms?

We might answer this with another question. If a patent on a gene or other human life form confers ownership over something human, and if, on moral grounds, we reject claims to own humans, are not patents on genes and human substances illegitimate? This question probes not an invention's use but the appropriateness of the patent privilege itself. A patent lawyer will reply, reprovingly, that a patent does not confer ownership, that a patent merely grants for a term of twenty years the privilege to exclude others from making, using, or selling an invention. This reply does not end the discussion. For various circumstantial reasons any policy on biological patents brings moral controversy in its train. In the first instance, allowing commercial entities to wield even limited monopolies on things human will seem morally problematic to many observers. Some will regard such privileges as threats to the autonomy of persons (as discussed below for clinical settings). Others will point to various economic consequences of wielding patents, among them high prices and restricted output of end products. When a DNA sequence patent issues but the patentee fails or declines to introduce a product predicated on the sequence, the only benefit of the patent, if one may call it that, is to prevent the patentee's competitors from exploiting the sequence. It may be granted that for some the welfare loss of squandering an opportunity to improve beer production, especially for a mere scientific career, is cause for lament. But if a patentee shelves a human gene patent and denies society an opportunity to develop beneficial drugs or to perform gene therapy, the cost may be human suffering. As we shall see, good reasons obtain to resist the generalization that biological patents enhance aggregate welfare. In respect of the foregoing concerns, one hears not merely the voices of patent examiners and courts — unlikely arbiters of morality in any event — but a variety of moral views held among citizens to whom accountability for governmental decisions is owed.

Because the decision to award a patent may be publicly perceived as at least implicitly a decision to condone any and all uses of the invention, it may behoove us first to resolve objections concerning morally problematic uses of certain biotechnological innovations before we attempt a consensus on monopoly of the innovations. If prudence commends this two-part agenda in the United States, the European patent system demands it. The European Patent Convention of 1963, whose criteria of patentability are otherwise roughly coincident with the American, proscribes patents on inventions whose commercial exploitation would be "contrary to *l'ordre publique* or morality." This phrase was long considered so vague as to lack teeth. But as adopted in 1998, the Directive on the Legal Protection of Biotechnological Inventions of the European Parliament (the "European Directive," or "ED") declares unpatentable, on the ground that their commercial exploitation would be contrary to *l'ordre publique* or morality, the following: human germ line intervention, "cloning" humans, commercial use of embryos, and both somatic and germ line genetic intervention in animals that is "likely to cause suffering without any substantial medical benefit to man or animal" (2). To this the ED curiously adds, "exploitation shall not be deemed to be so contrary merely because it is prohibited by law or regulation."

Anticipated Benefits of Transgenesis

Mankind has bred plants and animals for millennia. Since Mendel, breeders have exploited knowledge of dominant and recessive alleles. Moral controversy about genetic engineering stems not from manipulation by breeding, but from recombinant DNA. It is not that recombinant techniques clearly violate any moral view in particular. Rather it is the case that recombinant DNA technology poses questions not previously raised within any traditional moral theory.

Transgenesis consists in isolating a gene of one living being and inserting the gene at an early embryonic stage, before somatic and germ cells separate, in such a way that the gene enters the germ line of another living being of a different species. The insertion may be accomplished by introducing the foreign gene into (i) a retrovirus that infects an embryo, (ii) a plasmid microinjected into the pronucleus of a zygote, or (iii) cultured embryonic stem cells injected into the cavity of a blastocyst. The inserted genes are usually few and manifest themselves in only a small subset of an animal's phenotype. Transgenesis enables improvements in the growth, heartiness, and yields of animals and plants as sources of food, vaccines, and other compounds, affords models for study of diseases, the immune system, and gene regulation and expression, holds promise for direct therapeutic use in humans, allows the "pharming" of animal organs so that, upon transplant to humans, they will not be rejected, and allows production of cotton, plastics, and other industrially valuable compounds. A vaccine-enriched transgenic banana holds promise as a vehicle for surmounting economic and practical obstacles to vaccine delivery in many regions of the world.

Reservations Concerning Transgenesis

As encouraging as these prospects may be, they are not without their detractors. Objections to transgenesis include the following. Even if genes insert at a targeted locus, in animals the effect of transgenesis may be suffering, a theme frequently rehearsed in European discussions. The usual defense of animal experimentation (as in the ED) adverts to collective human benefit. A net

increase in aggregate human preference satisfaction is all that need occur to satisfy a utilitarian; the second form of Kant's categorical imperative permits using even humans as means so long as they are not used solely as means. But risk-averse humans worry about their own welfare in eating transgenic plants and animals — even assuming full disclosure in the grocery store. To introduce a vaccine into a banana crop raises questions about imposed risk-taking and paternalism when informed consent may not be feasible. Risk of human suffering is sometimes cited as a consideration against human gene therapy. Risks about where and in how many copies genes insert and whether a procedure will otherwise work are chanced by any single recipient of somatic cell therapy; to this germ line intervention adds the risk that an untoward result may burden future generations who, it may be said, have no voice in what is done to their ancestors' genomes. Even when gene therapy achieves an intended result, the long-term effect may be a population less diverse, a gene pool that is diminished. A suite of controllable genetic characteristics may eventually generate a canon. By reference to that canon, persons lacking certain traits may be treated by others as inferior. Perhaps indeed we shall cavort down a slippery slope from disease-related therapies to frivolous enhancements. To engage in germ line intervention, it is decried, is to "play God."

Defense of Human Germ Line Intervention

A defense of human germ line intervention might run as follows. Genetic engineering may be a way to improve man's contribution as co-creator in God's work (3). One might argue that God would wish caring physicians to use it. Gene therapy will not invent discrimination, a practice already thriving with respect to many traits. We should not count even against enhancement that someone will be born with a trait more desirable than another trait. Instead a temptation to be invidious should remind us, as do Kant and many religions, to recognize the dignity of each person. As for an effect on descendants, that is not a new category of moral responsibility. Future generations are already affected by innumerable influences on one's germ cells of how one lives. One may imagine a complaint of wrongful life by someone born with a genome adversely affected by something that went wrong in a gene therapy procedure, but in another case, the same tort might be committed by failing to attempt gene therapy. It would appear dubious to bar a physician from using an available method of averting disease in a consenting patient's offspring if, for example, the odds of the method's success exceed those of any treatment after birth. It is also difficult to expect a family or society to forgo eradicating a lethal gene if that be possible. A few decades from now, germ line intervention may be considered routine, its provision the duty of a competent physician, its inclusion a requirement of the health care a just society ought to provide.

The slippery slope to enhancement is not fairly ascribed to gene therapy (though perhaps to recombinant DNA). Recombinant human growth hormone, for example, is already dispensed. At least at present, the imagined efficacy of gene therapy is limited to diseases involving single genes, and among those, to diseases mediated by recessive genes, because an inserted gene's locus may not be controllable and any dominant defective gene remains in the patient. On the other hand, someday prospective parents may be speaking of a "designer child" polymath 9′ basketball player. In any event a distinction between "therapy" and "enhancement" may not be sharp or necessary. If a slippery slope connects therapy and enhancement, the transition from first violation of any ban on genetic enhancement to widespread violation thereof may be an avalanche. A treaty presented for adoption by members of the European Union bans enhancement (4). As of this writing, human germ line intervention for enhancement purposes is not feasible, but upon its availability, one may expect the following. Though a government may ban genetic enhancement, as soon as one person manages to procure an enhancement, others acting rationally will likely rush to procure enhancement in order to remain competitive (5). These aspirants will either violate the ban or migrate to a sovereignty that lacks one. Sovereignties will likely behave the same way, rushing to follow the first innovator for fear of being dominated by superiors. Unless one contemplates intrusive "audits" comparing parental and progeny DNA, a ban on enhancement seems unenforceable.

The foregoing brief account reveals moral concerns about uses that insinuate themselves into discussions whether to approve monopolies of uses. We may now turn to moral concerns about a patent privilege itself.

Autonomy and Patent Claims Against Parents and Children

The effect that a patent might exert on individual autonomy may be studied through a dramatic example. This first requires that we explain the rationale for patents on transgenics.

Patented Transgenic Organisms. Although Pasteur's yeast long before gained a patent, modern recognition of a nonplant life form as patentable occurred in a decision with respect to a bacterium into which were introduced plasmids rendering the bacterium capable of decomposing oil (6). Patentability was later confirmed for multicellular organisms (polyploid oysters) (7). A furor ensued over ethical concerns, and for a time, the U.S. Patent and Trademark Office ("PTO") imposed a moratorium on animal patents. Thereafter the PTO issued a patent on the Harvard mouse (8).

Designed for the study of cancer, the Harvard mouse contains DNA sequences, comprising an oncogene such as *myc* and a promoter, that effect a high proclivity to form tumors when the mouse encounters carcinogens. The introduction and expression of the *myc* gene in the mouse was innovationary. One could not have presumed that a zygote acquiring the oncogene would survive its insertion and expression, or if the zygote did survive and a mouse were born, that the offspring would be fertile. The Harvard mouse invention was exhibited for patent purposes by deposit of DNA in a plasmid. But the patent extends not merely to the inserted DNA, not merely to the oncogene's introduction and expression, but to the whole "oncomouse."

Why should a patent embrace an animal? Two arguments might be mustered. First, there has endured throughout the history of patents the notion that patents should not be available on nature's extant treasures — in a phrase attributed to Thomas Jefferson, on any "product of nature" — but should be available only on what humans manufacture. The discovery of uranium garnered no patent, but the PTO issued a patent to Glenn Seaborg on curium and isotopes of americum, transuranic elements believed to exist on earth only in a cyclotron or reactor as a result of human efforts. Of course such elements may be abundant in stars. A precise statement of patent eligibility would not exclude from patentability every naturally occurring thing. Rather we may state patent eligibility by the following proposition, which we may call the "unpatentability of nature": to be eligible for a patent, a thing must be such that there obtains only a very low probability that, without human intervention, the thing exists near the surface of the earth or on other astronomical bodies to which humans travel. What constitutes a "very low" probability requires specification. Since patent lore speaks confidently of things that "exist" or "do not exist," no guidance on probability may be found within it.

It is logically possible for there to evolve an organism whose genome is identical to some modern transgenic organism. Exchange of genes across species occurs and any mutation is possible. Yet the probability may be extremely low that a creature will contain genes of two given organisms that do not mate. In such case a patent examiner may treat a transgenic genome as if it does not naturally occur. An animal possessing such genome is then seen not as an unpatentable product of nature but as a patentable "manufacture" or "composition of matter" (9). With respect to the European Patent Convention, it would be said that such an animal is not an unpatentable "variety" (10). In general, the fruits of breeding programs are considered varieties, but transgenic animals are not considered varieties because transgenesis was unknown in 1973 when the European concept of a variety was introduced.

Second, when introduced into a recipient's germ cells, transgenes pass to descendants. A transgene will not be expressed in all offspring of the first generation, but those in which it is expressed will be selected for further breeding. Were a patent to cover only cells expressing a trait, it would not capture the invention, which, by virtue of being genomic, appears in every cell. Were a patent to cover only an inventor's process of introduction and expression, anyone who purchased one transgenic animal could breed others without infringement; natural reproduction is not the same process as laboratory transgenesis. A purchaser of transgenic agricultural livestock could breed the livestock through unlimited generations. Hence inventors are accorded the protection of a patent on the transgenic genome, which is effectively a patent on the animal. Breeding descendants of a patented transgenic animal without license would just as clearly constitute infringement as would duplicating a patented laboratory process for inserting transgenes into an embryo. The patent system assimilates reproduction, whether natural or artificially aided, to "making" a duplicate.

The Human Qua Infringement. A moment's reflection reveals that if the foregoing two grounds (a claim to originate a manufacture; self-reproducibility of a recipient) entitle an architect of transgenesis to a patent on recipient and progeny, then in the case of human germ line intervention, infringement claims will lie against the birth and existence of humans. To call human birth or life a "patent infringement" seems perverse. But on what principled grounds should we reject such claims? Even in the somatic cell case, as scientists perfect the manufacture of yet more human enzymes and other proteins, as they progress to substantial tissues, should society continue to grant patents on human "parts"? Manufacture of a liver or other major organ, or someday even of a brain, may confound previous thinking.

Abjuring the Human Qua Infringement. To resolve the solecism of the human qua infringement, we may reason as follows. We do not imagine infringement claims against any plant or animal. Instead we recognize claims against people who control breeding. We do so because we recognize a farmer's ownership of plants and animals. When the "designer" of a transgenic organism applies for a patent, the contest concerns only which humans (or corporations they represent) own property in the nonhuman species. When the issue is which of two humans owns a human, we say that humans own themselves. They do not own each other. Human births, we hold, are not analogous to breeding, to manipulation by owners of mating subjects. Hence we may decline to recognize property rights in humans.

The premise that humans do not own each other, that we each enjoy a "bodyright" (11), is not categorically held in all societies, and given that the common law describes an adolescent's maturity as "emancipation," perhaps it is not unequivocally held anywhere. Defense of the premise often comes round to some distinction between humans and animals. According to a Cartesian distinction, man is a singular creature possessed of reason. According to Kant, only man and angels are capable of reason. To say that a human being's existence infringes property rights would seem inconsistent with the second and third forms of the Kantian categorical imperative, which together enjoin that we treat each person not solely as a means but as an end-in-himself in a kingdom of ends. If we allow an ownership claim on a person, we condone treating the person solely as a means. We condone interfering with the person's autonomy. Were someone to assert an ownership claim that purportedly extended to only part of a person, the claim would appear indistinguishable from a claim on the whole. Bodily parts are integrated. For the same reasons, conception and birth, the instantiation of human nature, may be held immune from claims of others. We may also say that conception and birth are private.

Distributive Justice and Patent Claims on Extracorporeal Compounds

Although we may thus deny the permissibility of exerting dominion over, impairing the autonomy of, or disrespecting an individual, a different case, actual in biotechnology, is the following. There is adduced a substance that is human

in the sense of being found in the human species, but which has been made outside the human body and is not ascribable to any individual. Were a patent to issue on that substance, the patent would not appear to interfere with any individual.

To this it must be added that it is not easy to steer clear of the DNA sequences that distinguish an individual or that make any individual akin to another. Only about three million of the three billion base pairs in the human genome account for individual differences, but the genetic code is redundant, the most interesting traits are polygenic and beyond present understanding, and mutation never ceases. For now, individual identity, to the extent it is genetic, is genomic. We have not demarcated a nonindividuating subset of the genome that we may cede. We do know that individuality is greatly affected by a relatively small number of regulatory sequences that control which genes are expressed. Such sequences are indeed used in biotechnology manufacturing unless a bacterium's or other host's regulatory sequences effect expression.

Let us assume for the moment that it is possible to grant monopolies on proteins and DNA sequences without there resulting any interference with the autonomy of any individual. The PTO effectively allowed as much when it began to grant patents on human DNA sequences despite its earlier declaration that a patent on a human would violate the prohibition of slavery in the Thirteenth Amendment to the Constitution of the United States. It contravenes common usage to say that a nonpossessory interest in a protein or gene constitutes slavery.

Were one confident that a system of limited monopolies would lead to advances that prevent or alleviate human suffering, one might decide that the conceptual coherence of the patent system should give way to the promotion of aggregate welfare. If patents on genes contravene the unpatentability of nature, so much the worse for that premise. It seems that one perforce resorts to that stance for the defense of patents on plant antibiotics. One might go so far as to say that there should issue any patent, even if the patenting process is purely piecemeal, that results in net aggregate welfare gains.

Whether compromise with intellectual purity be systematic or piecemeal, and even assuming any contingent results that an advocate of such conceptual indulgence predicts, this talk of welfare effects presupposes a criterion for discerning improvements in welfare. That in turn implicates some version of a social welfare function. A social welfare function is a function that yields or induces a positioning of possible resource allocations on which one may predicate a claim such as "α is a welfare improvement over β." The specification of a social welfare function is the main problem of distributive justice. For this reason, what begins as a moral problem concerning respect for personal autonomy, which arguably is tractable by virtue of the ability to eschew interfering with any individual, endures as a challenging problem of collective morality.

To the extent that patents are distributive mechanisms, this problem arises within an economy in respect of any patent. But concerns abound with respect to the welfare effects of biotechnology patents. Some may espy unacceptable burdens and risks for the human species as a whole from various patents on molecular or structural human life forms. For instance, when a patent owner sets what seems an exorbitant price for a vital drug, one observes an arguably undesirable effect of market power conferred by an unqualified government-created monopoly.

PRODUCT PATENTS ON HUMAN DNA SEQUENCES

Supposing that the prospects of collective benefit or some other morally persuasive consideration have justified the alteration and use of life forms, why confer exclusive control on one party? The orthodox *quid pro quo* of a patent is that, instead of keeping an invention a trade secret, the patent teaches the details. When a patent expires, the invention will be in the public domain, and even during its term, what others learn from its teaching may foster other innovations. Whether the patent's revelations are in fact valuable will depend on whether one may easily infer the invention by reverse engineering (12). An alternative and more familiar rationale asserts that a patent provides an incentive that fosters ingenuity and effort. Or as Bentham put it, "He who has no hope that he shall reap will not take the trouble to sow" (13).

Isolation-Purification Rationale for DNA Product Patents

Organic compounds found in humans are, *ipso facto*, naturally occurring. Suppose that an organic chemist discovers a way to synthesize a protein in a purified form not found in humans. If the protein appears extractable from another organism, then perhaps we should not regard the protein as distinctively human. But in fact the human version of a given protein is unlikely to be identical with that of another organism. Through mutations in duplicate genes, species have evolved a variety of genes coding for different versions of proteins that we call by single generic names. The notion of isolation and purification (a creature of case law, not statute) was popularized by product patents on inorganic chemicals. (In patent parlance, a "product patent" is a patent on a thing as opposed to a patent on a process.) The notion was then borrowed in support of patents on the products of biological processes, including purified human adrenalin, prostaglandins, vitamin B_{12}, and, most recently, human DNA sequences. For the last, investigators' counsel have persuaded patent examiners that investigators have "isolated and purified," which is to say cloned, human genes.

There is reason for scepticism whether a patent must be available in order to induce a given result. "The large amount of research that has already occurred when no researcher had sure knowledge that patent protection would be available," noted the Supreme Court of the United States in affirming the patent on the oil-eating bacterium, "suggests that legislative or judicial fiat as to patentability will not deter the scientific mind from probing into the unknown any more than Canute could command the tides" (6). There arrived for filing a spate of plant patent applications in Europe, many presumably

from European companies, quickly after the first plant patent was allowed there in 1989, which suggests that the research had long before been done. And rivals face effort and expense to follow a first entrant into the market. In the United States the lead time that an imitator of a drug or medical device would need to obtain approval from the Food and Drug Administration ("FDA") for selling the product provides a period of *de facto* exclusivity to the product's originator once the originator obtains approval. There may also be observed a tendency after a favorable experience for physicians to continue prescribing, and consumers to purchase, the first drug of a genre. Assertions about the necessity of incentives can be facile, but evidence is lacking.

Inventions concerning nonliving phenomena make use of materials that mankind has long exploited with an aplomb perhaps attributable to the mistaken belief that we cannot alter earth's vastness. Biological inventions obviously effect alterations of nature. If we approve the engineering of some protein, we might view the circumstance that it is found in humans as a reason against monopoly. Hence the isolation–purification rationale originated for chemistry in general cannot be assumed to carry the day for human compounds. One might add that to adopt such construct for humans would not follow the model of chemistry faithfully enough. Patents have been granted to those first to synthesize chemicals, but courts tend to find evidence that chemical patents have been infringed only insofar as a patentee's process has been copied.

The probable benefits of recombinant innovations may in a given case outweigh the acknowledged detriments. But each wave of innovation evokes a new comparison of risks, costs, and welfare gains.

Unpatentability of Nature and DNA Patents

An enduring challenge for the molecular biologist is to understand a disease or bodily function, to identify a protein related to it, to ascertain the nucleotide sequence of a gene coding for the protein and the protein's amino acid sequence, to locate the gene on a chromosome, and to explain the gene's regulation and expression. Once sequenced, a cloned gene may be preserved as complementary DNA ("cDNA") in a vector. When vectors transform and infect, not only do they multiply an inserted gene, but the gene can integrate into the transformant genome, causing such host to produce the protein for which the gene codes. It is by growing such transformants under suitable conditions that a biotechnology manufacturer may produce a protein in high volume.

A typical patent claims at least three inventions: (a) an isolated and purified DNA sequence encoding some protein, (b) any vector that contains that sequence and any transformed host possessing that sequence, and (c) one or more processes. The patent system indulges the notion that the cloning of genes produces "inventions" that do not naturally occur. By contrast, a detailed examination of what occurs in the laboratory, though providing ample evidence to confirm our admiration for scientific achievements, reveals no *entity* that mankind creates (14).

A gene encoding any human protein exists in nature. It is embodied in a chromosome. Its transcript also exists in mature messenger RNA ("mRNA"), a single strand of DNA-complementary nucleotides. Transcription of DNA into mRNA, followed by splicing that eliminates introns, is nature's own "isolation" of the coding sequence (with uracil in place of the thymine of DNA). This alone seems to tell against the argument that only an inventor has achieved isolation. Is "purification" then the inventor's trump over nature? The process of making cDNA is not thought to occur naturally in humans (though many viruses that infect humans make DNA from RNA). But once a gene is known, the laboratory process of making cDNA can be routine. Perhaps then vector and host deserve credit as the ingenious embodiments of purification? Where a host and a donor of foreign DNA are members of species that do not mate, asserted Stanley N. Cohen and Herbert W. Boyer in teaching the first recombinant process, a recombinant host "could not exist in nature" (15). This imagined impossibility of course is an exaggeration. Any mutation is possible. As Bernard D. Davis observed in debating the hazards of recombinant DNA research, bacteria absorb foreign DNA from lysed cells of their host, including the human gut. The rate of bacterial absorption of foreign DNA is low, but the number of bacteria in the gut is enormous, and bacteria have thrived for millions of years. There is some probability for any given human gene that it has already occurred in a bacterium. We may not observe that gene today, since the gene may have conferred no selective advantage on the bacterium, and the strain vanished (16). On the other hand, what is the likelihood that a plasmid or bacterium naturally contains a given human gene? That probability may approximate that of unicorns existing. Cohen and Boyer evoked a sense of mythological improbability when they called an altered plasmid a "chimera."

When a patent claims a chimera and host, the two are usually notable only in one respect: they contain a DNA sequence that the applicant purports to have invented. The plasmid and host are effectively the sequence's housing and factory. Once an investigator has selected a sequence as described above, the process of cloning it, and hence the "invention" of the sequence in a vector and transformed host, is mechanistic. Advanced techniques for sequencing proteins may make straightforward the selection of probes and primers, and hence the discovery of a known protein's gene. The "invention" of a protein variant may also follow straightforwardly from a variant, experimentally achieved, of a gene sequence. An investigator who finds naturally occurring genes and proteins merits accolades. But it remains to be shown why a patent should issue, and if so, on what.

What Should Suffice for a Biotechnology Product Patent?

As necessary conditions for award of a product patent, consider the following: (*1*) a claimed invention is such that it is highly improbable that we shall find it as such on earth or on any other astronomical body to which human travel is possible, and (*2*) the invention is ingenious. The "as such" phrase in (*1*) would allow some particularly convenient forms (e.g., a vector with a foreign DNA insert) to gain

recognition apart from a natural form. In (*2*), "ingenious," which shares an etymological root with "genetics," is a placeholder for what constitutes an invention, of which more later. Sequences fail (*1*) if they are found in chromosomes and mRNA. Chimeras, transformed hosts, and cDNA meet (*1*) but fail (*2*) when they are mere mechanistic steps from discovery of a sequence. The vectors and hosts of microbiology are not as fastidious as the species united in mythological chimeras: they accept DNA inserts regardless what creature originates them.

An "artificial gene" may satisfy both (*1*) and (*2*). In theory such a gene may be constructed of any codon-containing sequence one likes; in practice the amino acid sequence of a protein may inspire the sequence. Some caution may be needed in characterizing a sequence as "artificial" because if a sequence codes for a human protein, then either that sequence, another differing only by substitution of alternative codons for the same amino acids, or yet another that is insignificantly different exists somewhere in the genome. Because of the phenomenon of overlapping genes, the sequence may also be part of another gene that its discover has not even envisioned. On the other hand, it may be that an "artificial gene" meets (*1*) and that its gene product is in some sense superior. (The artificial gene product might, for example, lack contaminants usually found in the natural gene product.) Not every such sequence will be ingenious. Courts have often declared DNA sequences inferred from protein sequences to be obvious (17, p. 50).

In view of patents on algorithms — a departure from previous conventional wisdom that ideas are not patentable — one might appeal to the notion of patents on information as a defense of DNA patents. But this defense would seem to fail insofar as any information encoded in cDNA is encoded in naturally occurring DNA and mRNA.

Adverse Welfare Effects of DNA Patents

If the autonomy of no one in particular is threatened by a product patent, the autonomy of everyone together might be.

The PTO in 1987 granted a product patent on isolated and purified natural erythropoietin. Merely four months later it granted a second patent to another party relating to a recombinant DNA technique for making the protein. The second patentee had cloned the gene after screening a genomic DNA bank with two sets of probes. It then produced the protein in transformed hamster ovary cells. The first patent blocked the invention of the second. This portended that patients would be deprived of a recombinant method of producing erythropoietin in high volume at low cost. As a group, patients were saddled with the first patentee's production method (extracting extremely low yields of the protein from thousands of gallons of urine). When, four years later, the first patent was invalidated on unrelated procedural grounds, the second became a barrier against any better recombinant process employing the claimed sequence (18). Similarly did a biotechnology firm discover the human gene for factor VIII:C by probing a human cDNA bank, inserting the gene in plasmids, transforming hamster kidney cells with the plasmids, and producing factor VIII:C. This recombinant advance was blocked by an earlier patent on factor VIII:C itself (19). The patentee's process not only required enormous amounts of donated blood plasma for a small yield, but in contrast with the recombinant method, it risked contamination. Contamination was a critical risk because many hemophiliacs who received contaminated factor VIII:C died of AIDS. The recombinant's manufacturer protested unsuccessfully that the patentee had not invented factor VIII:C (though, given the chance, the manufacturer might have argued for its own invention of the recombinant). The erythropoietin and factor VIII:C episodes illustrate how product patents may frustrate society's interest in encouraging, at the same rapid pace at which biomedical research is otherwise moving, helpful innovations in processes for making therapeutically valuable human compounds.

Farmers raise the specter that, burdened by the high cost of patented animals and crops, they may turn to cheap unpatented strains, that crops will become less diverse, and that more crops will succumb to pests. In transgenesis, often an investigator cannot control the place within a genome at which a foreign gene inserts or the number of copies that insert, or in the case of plants, the weeds or other unwanted plants to which a transgene may migrate via airborne pollen. We also have reason to rue "blind promotion of technological innovation" (20). Agriculture, after all, is an industry afflicted with overproduction. A patent granted by the European Patent Office on all manner of genetically engineered soybeans has been criticized on the ground that soybeans are among the world's most important crops and monopoly of soybeans will threaten "world food security." Suppose that a patent on a critical crop, organism, or substance has been conferred on an enterprise that becomes bankrupt. Or suppose that the patent is acquired by some foreign entity that is involved in international intrigue, that uses the patent as leverage for some disreputable purpose, or that otherwise seems to control output contrary to the common good. As exemplified by experience with the anti-AIDS drug zidovudine (or "AZT"), the price of a patented product is a monopolist's price.

Scientists have become acutely aware that availability of patents on DNA sequences may be generating a patent race that misallocates resources and delays publication of results. This would run contrary to the hope underlying the Human Genome Project that disseminating chromosomal mapping and sequence data will foster growth in collective knowledge. It took four years after a gene implicated in breast cancer, *BRCA1*, was mapped to chromosome 17 before one of twelve rival collaborations found the gene, a feat they all recognized as a "discovery" (21). Yet the winner immediately sought a patent on *BRCA1* and related diagnostic processes. About a year later, one of the competing groups contributed to a public database the sequence of a large portion of chromosome 13 where *BRCA2* was thought to repose. "It will not be helpful to medicine," the scientist John Sulston was quoted as saying, "if, by the year 2003, control of every single gene is tied up by one company or another for twenty years. That would be an enormous ball and chain. . . .[F]or

the good of humanity, we should try to keep these things in the publicly exploitable domain" (22). The group contributing the chromosome 13 sequence data urged that DNA sequences be public information. Within a month thereafter, *BRCA2* was found (23). This seemed to exemplify the rapidity of progress when results are shared. Thereupon the discoverers of *BRCA1* filed for a patent on *BRCA2*, launching a dispute over who found *BRCA2* first. Seemingly ignored was the untenability of claiming to invent parts of nature's storehouse.

One cannot dismiss objections to product patents as the outpouring of any single, disputed moral view. Even without an appeal to morality, it may be argued on exclusively scientific and economic grounds that patents on human DNA sequences violate the unpatentability of nature. Many moral views assign significance to the aggregate welfare consequences of that violation. In such case the moral case against human DNA sequence patents reprises the scientific.

ALTERNATIVE INCENTIVES FOR BIOTECHNOLOGICAL INNOVATION

Measures for Holding onto the Availability of Product Patents

The legal criteria for patentability are that a "process," "manufacture," or "composition of matter" be "new," "useful," and "nonobvious" (24). According to a conservative article of faith espoused by patent practitioners, these criteria possess such protean qualities as to suffice for the resolution of all questions that arise from time to time. The criteria need only be interpreted by the courts. In reply to this, it must be said that, under prevailing interpretations, "new" and "useful" erect only minimal thresholds. "New" eliminates from patentability only what has already been published. "Useful" eliminates from patentability only the utterly useless, a rare creature among proffered inventions anyway. (Scientists at the National Institutes of Health, NIH, dramatized the weakness of the "utility" requirement in 1991–1992 when they ostensibly satisfied the criterion by citing a seemingly trivial use for parts, "expressed sequence tags," of cDNA sequences. The applicants conceded ignorance about the feature of usual biological interest, viz., what the sequences encode or regulate, and ventured only that the tags could be useful as genetic markers, primers, or probes in diagnostic kits for unnamed diseases. But any DNA sequence may be a marker in genomic mapping.)

To resolve a question of patentability, two sobriquets, "manufacture" and "nonobviousness," must carry most of the load. In fact nonobviousness must do all the work. For it is considered settled in U.S. patent law that cloned DNA sequences, as fruits of "isolation and purification," constitute a patent eligible genre. Being a "product of nature" is now seen as no impediment to patent eligibility; the question is whether a sequence constitutes a new, useful, and nonobvious manufacture or composition of matter (25). Thus stood on its head is Jefferson's use of "product of nature" for the unpatentable. But the point is only semantic: since every extant thing's ingredients are naturally occurring raw materials, every extant thing may be called a "product of nature" in some sense. The semantic point entails no practical consequence if some other provision insures the unpatentability of nature. (As defined earlier, the unpatentability of nature is the premise that to be eligible for a patent, a thing must be such that there obtains only a very low probability that without human intervention, the thing exists near the surface of the earth or on other astronomical bodies to which humans travel.) We might think that the statutory term "invent" secures the unpatentability of nature. But instead for the domain of biotechnology though not for others, we observe patent examiners and courts effectively either rejecting the unpatentability of nature or exhibiting remarkable restraint as they construe the premise. The DNA sequences that they pronounce patentable are sequences that *chromosomes of living beings contain*. The only apparent way to reconcile this with some version of the unpatentability of nature is to emphasize that a given cDNA sequence corresponding to a chromosomal sequence differs from the chromosomal sequence insofar as the chromosomal sequence is littered with introns. Still it must be said that the chromosomal sequence *includes* the cDNA sequence. That is to say nothing of the chromosomal sequence's uninterrupted transcript in the form of mRNA. Courts and patent examiners keep faith with only a weak version of the unpatentability of nature.

Given that DNA sequences are recognized as a patentable genre, whether a given sequence garners a patent turns on whether the sequence is deemed obvious. When courts first struggled with arguments about recombinant DNA technology, it seemed obvious that what was obvious was not obvious. As courts came to recognize recombinant techniques as commonplace, they bent over backwards to conclude that newly discovered cDNA sequences were nonobvious. If the prior art did not enable a method of finding a proffered sequence "with a reasonable prospect of success," a court would pronounce the sequence nonobvious. (This move vindicated a patent on the sequence encoding erythropoietin, a sequence found by screening a genomic DNA bank with two fully degenerate sets of oligonucleotide probes.) As further reasons to sustain a verdict of nonobviousness, courts have even recognized circumstances extraneous to the intellectual process of discovery and invention, including commercial success, long-felt need, failure of others, unexpected results, and the scepticism of rivals (17, p. 19). As critics would have it, one influential judicial decision saves the day for cDNA patents only by tortuously construing "obvious" so as effectively to declare patentable *per se* any DNA sequence found to encode a protein (26). According to one observer, the obviousness of many purported cDNA inventions is betrayed "in the very attitude of the persons skilled in the field. Today, if a researcher discovers a new protein and its probable properties, he usually does not publicize the information until he has found the corresponding gene. How to explain this in a community whose motto is 'publish or perish' save that it would be obvious to another research team to pick up the information, and clone the gene?" (17, p. 90). In hopes of securing future DNA patents against a tide of progress that may render

ever fewer cDNA sequences nonobvious, it has been suggested that the nonobviousness requirement, to the extent not already emasculated by the aforementioned judicial decision, might be weakened. If the steeplechase jump proves too high for the average contestant, lower the bar. Where a nucleotide sequence is itself a drug, as with anti-sense RNA or the use of a DNA sequence to achieve expression without integration into the genome, it has been suggested that obviousness might be replaced with superiority over prior art in therapeutic efficacy (17, pp. 141–143, 148). Without some such move, it is urged, future application of the obviousness standard may thwart the availability of product patents on the expectation of which the biotechnological industry arose. Here an appeal is made to the biotechnologist's familiar prediction that a world without DNA patents will be a world without therapeutic innovation. But rehearsing that prediction does not provide evidence for it.

The issue remains whether, all things considered, more DNA product patents should issue. By virtue of considerations mentioned earlier, the answer may be in the negative. In such case, what might replace such patents?

Categorical Prohibitions

In the Biotechnology Patent Protection Act of 1995 (27), procured at the behest of the biotechnology industry to allow a biotechnology process claim to piggyback on a product claim, precedent was set for legislation that speaks in biological parlance about patents concerning molecular biology. Despite the supposed protean generality of the patent conceptual scheme, the door has been thrown open to discipline-specific rules. What most commonly seems to flow through that door is a stream of *ad hoc* prohibitions, usually categorical and often embracing uses as well as monopolies on uses. One favorite in the U.S. Congress and the European Parliament is a moratorium on a given sort of patent or research. Sometimes the rationale for a prohibition will appear to assimilate a property interest in an extracorporeal molecule to invasion of personal autonomy, which earlier we found reason to distinguish. As the portion of the human genome claimed by patents expands, motivation arises to prohibit any more such patents. Were a ban imposed, the control experiment of life without patents would run in real time. Biotechnology firms would compete with no intellectual property save for trade secret protection of whatever they managed to keep secret. Such competition might produce salutary results. It might also diminish aggregate welfare unless some mechanism replaces at least some of the incentives fostered by product patents.

Compulsory Publication of DNA Sequence Data to Thwart Patents

When without first filing for a patent, a scientist publishes a DNA sequence, no one may obtain a patent on the sequence. Mere citation of that publication as prior art will spike anyone's claim that the sequence is "new." Mindful of this, some scientists acting of their own volition and other scientists acting in compliance with funding mandates have promptly and systematically released DNA sequence data as discovered. A concerted effort so to publish could thwart most new DNA patent applications. Thereupon it becomes open season for any and all to explore therapies predicated on all unpatented portions of the genome. The benefits of expanding the universe of potential investigators would seem apparent. To the extent that research motivated by profit may contribute applications that might not flow from academic laboratories, incentives must now be sought elsewhere.

Subsidies

When a public good is underprovided, as is familiarly the case in perfect competition (e.g., as to education and national defense), government may step in to provide it. Valuable public ideas are intellectual public goods. Suppose then that no further patents issue on DNA sequences other than artificial sequences. Instead, the government systematically subsidizes biotechnological research. Subsidies are awarded not only to academic institutions but to nonprofit biotechnology research centers. Specifically organized for the pursuit of applied as well as basic research, these centers tackle applications that might not be pursued, or pursued with less zeal, in academic laboratories. This scheme could implement coordinated decisions, reached with benefit of expert extramural advice, concerning which fields of fundamental biomedical and biotechnological research should be pursued and to what extent. The scheme entails substantial expenditures and may importune taxes earmarked for research (20). But the subsidies assure that society gains the benefit of valuable innovations.

Were it widgets that society sought to encourage, subsidies for institutional laboratory research might not succeed in coaxing the same innovations as would market incentives for entrepreneurs experimenting in their shops. When the desired innovations are biotechnological, it happens that academic laboratories constitute society's most fertile source of ideas. What academic laboratories do not pursue by way of applications may be pursued in the research centers. Were a share of sales revenues promised to any laboratory originating an end product, market incentives could also be brought to bear within both academic laboratories and research centers.

In the marketplace, with valuable discoveries being contributed to the public domain and available for exploitation, firms would now compete less on the basis of their discoveries and more on their efficiency in production. As with any subsidy, it may be difficult to ascertain whether the extent of biotechnological innovation induced is optimal. But one could at least compare the extent of technological innovation during the present era of product patent availability with the extent of technological innovation under a new regime.

Exclusive FDA Approval for a Term of Years

A government may also engraft an incentive mechanism upon the process by which, with a view to public safety, the government grants approval for the sale of medical products and devices. The Orphan Drug Act (28) affords a model for such an incentive scheme. According to that statute, if the FDA grants a manufacturer approval to

sell a drug targeted at a disease that affects fewer than 200,000 persons in the United States, or whose likely sales cannot reasonably be expected to recoup the costs of development, the agency must refrain, for seven years after such approval, from approving sale of the drug by anyone else for use against that disease. Routine delay in obtaining FDA approval for any drug affords to the first party who gains FDA approval some period of *de facto* postapproval protection against imitators; in respect of an orphan drug, the first party to gain approval enjoys seven years of *de jure* postapproval exclusivity. The orphan drug scheme is not without its complications. For purposes of identifying which compounds are blocked for seven years by an approved orphan drug, it has been necessary to define what constitutes "the same drug." The FDA defines a new drug to be the same as a previously approved orphan drug if the new drug has the same "principal molecular features" — unless the new drug is "clinically superior" to the approved orphan drug (29). This seems to rehearse, though with variations, a judicial patent doctrine that a claimed invention is obvious if the prior art includes a structurally similar compound — unless the claimed invention possesses an unexpected property (17, pp. 145–148).

This incentive scheme could be extended. From orphan drugs it could be extended to any genre of products that seem likely to serve the public interest — indeed to any and all biotechnology products. To specify the genre of products for exclusive approval, the government could rely on advice from extramural scientific panels. Such a scheme would spare the costs, burdens, and uncertainties of patents. It would reward the development of valuable products without tying up the human genome with property claims. It would respect the unpatentability of nature. The number of years and other terms of the exclusive sale privilege are of course variable. One might also replicate the provison of the Orphan Drug Act that allows the FDA to approve sale of an orphan drug by a second applicant if the original manufacturer "cannot assure the availability of sufficient quantities of the drug to meet the needs of persons with the disease or condition" (30).

Human Methods Patent

Ambivalence between patents on products *vis-à-vis* patents on processes has been evident since recombinant DNA technology began. The Cohen-Boyer patent protects a process. It was followed, as the technology developed, by many process as well as product patents (31). The Cohen-Boyer application also sought claims on recombinants, but no product patents issued until 1984 (on plasmids) and 1988 (on plasmid-transformed hosts). Stanford University's licensing of the process patent thrived beginning in 1981 before Stanford acquired any product patent (32).

A possible resolution of the ambivalence would be to require hereafter that to secure a patent pertaining to a DNA sequence, one must invent some new process that can be performed in respect of the sequence rather than claim to have invented the sequence or its gene product. A new form of patent predicated on this principle has been proposed (14) in the following statutory phrasing:

> There shall be allowed a patent pertaining to a human life form (a "human methods patent" or "*HMP*"), the scope of which patent shall not exceed the least inclusive description of an ingenious process. Such a process may consist in the production, use, alteration, amplification, or attenuation of human life forms outside the human body. An *HMP* may include an additional claim on nonhuman reproduction of any transgenic and its progeny if and only if (a) the ingenious process produces such transgenic, (b) such transgenic produces a human life form, and (c) without reference to such human life form, the process would not be patentable.
>
> No product patent shall be allowed on a human life form or anything in which it is included. The foregoing shall not preclude a patent on a synthesized, fully explicated nucleotide sequence or protein that is not present, consecutively or otherwise, in the human body.
>
> Research in a nonprofit institution for nonprofit purposes shall be exempt from any claim of infringement.

The significance of an *HMP* may be made more clear by the following observations.

(i) Interspecies homology is only similarity to a degree, not identity of nucleotide sequences. Absent evidence of identity with a nonhuman form in a given case, "human life form" may be assumed distinct.

(ii) The confinement of an *HMP* to the least inclusive description of an invention protects against the detriment of overbreadth as illustrated by experience with erythropoietin and factor VIII:C. Such limitation would depart from the law's tendency to allow contributors a claim on a whole — as when a farmer obtains a claim on grain in an elevator with which the farmer's is commingled, or a security interest in a part attaches to a mass in which the part is commingled or assembled. Reasons for parsimony obtain concerning the human genome.

(iii) Suppose that an *HMP* claims "a method for obtaining DNA sequence $b_1, b_2, \ldots b_n$ from genomic DNA as follows: . . . ," and describes the ingenious method by which the sequence was discovered. Without more, such a patent would afford little protection. Everyone may now read the sequence disclosed by the patent. Free of infringement, anyone may then obtain the sequence by employing any process, including the polymerase chain reaction, other than the patented process. To avoid this vulnerability, the discoverer might seek to claim "cloning of the sequence in vector v and transformation by v of host h that results in production of protein p_i as follows: . . . ." Perhaps this investigator has ingeniously devised a way to use a new v to produce p_i in some mammalian h never before used to produce human substances. In general, it will not be ingenious to clone an identified sequence, nor to produce a protein by means of a known gene. The principle of least inclusiveness allows a claim on only so much of the process as is ingenious. The discoverer's successors may find it unnecessary to use the process first used to discover the thing, and may proceed to "event around" the process. This is true about the discovery of any natural thing. Successors may also invent methods by which to use, alter, or promote or attenuate the effect of the thing.

Process patent opportunities still await — in protein chemistry, insertion of foreign DNA, transformation and infection, gene expression, and protein-manufacturing techniques. One process might describe a technique for

making a protein, another how to use it. Or a firm might use knowledge of a gene not to produce but to curtail the effect of a given protein, including a newly discovered protein.

(iv) It may be possible to state certain minimal conditions for work to be ingenious. If a claimed method is predicated on a human life form, it may be unlikely to exhibit an advance over present knowledge unless the life form is fully explicated. "Fully explicated" entails, in the case of DNA, that a specific nucleotide sequence (the "explicated sequence") is identified, including all regulatory sequences necessary for any exons in the explicated sequence to be transcribed into RNA and for a gene to be expressed, that all such sequences have been inserted into a vector or maintained in some stable form, that it is known what the explicated sequence encodes or regulates (or perhaps only that it is implicated in the etiology of a disease), and that the process succeeds in expressing or preventing expression of such gene. For a protein, full explication would embrace biological function, amino acid sequence, and encoding gene sequence.

(v) One could circumvent a patent thus far described if, for example, one were to pay a royalty in order to perform a patented transgenic process of producing a human hormone in a pig, and then, without paying any more royalties, one were to breed a line of pigs. Thereby one could obtain copious amounts of the hormone. Natural breeding of course produces naturally occurring progeny, and, except for plants in the United States, such progeny would seem unpatentable. To prevent the foregoing circumvention, which would defeat an inventor's reasonable property expectations, the *HMP* allows a claim on growing or nonhuman breeding of a transgenic if the transgenic is a result of the invented process and the transgenic produces some human life form without reference to which the process would not be patentable. The additional claim may be defended as a claim on reproduction of an "unnatural" organism, one not likely to be found in nature. Such scheme resolves the predicament to which the self-reproducibility argument for the Harvard mouse patent is directed. It allows no claim on a human life form itself. Nor may the additional claim encompass human reproduction. Should the invented process happen to be one of artificial human reproduction, remedies may be provided (as discussed in the next section) against infringing providers, but never against a parent or child as such.

Whether an "artificial gene" or the protein it encodes will qualify for an *HMP* is contingent on how close a variant or equivalent the gene may be to what is found in the human genome. Will this contingency discourage fruitful research on the genome? Significant disincentives seem unlikely unless firms so greatly prefer product to process patents that they choose to pursue the more difficult task of sequencing proteins rather than finding naturally occurring genes encoding for proteins. Where proteins may be sequenced automatically, a disincentive may occur. But if the therapeutic value of an artificial gene product is not sufficient, it will not be an appealing product no matter what the patent availability. At least the products of naturally occurring genes have known worth.

Objections to the *HMP* and replies thereto include the following. Industrialists preferring product patents often contend that recombinants are more potent and free of contaminants, that recombinants thus differ from natural isolates and from each other, and hence that product patents will not prevent new advances from reaching the market. This conjecture seems belied by the history of erythropoietin and factor VIII:C in particular, and in general by the hegemony of any product patent over improvements. Whichever industrialist happens to be first in time will often hoist another on the petard of contradiction. In scientific publications and in advertising, sellers of recombinants are wont to describe their products as virtually identical to the corresponding natural isolates. But when forced to defend against a claim of patent infringement, the same sellers may be heard invoking the "reverse doctrine of equivalents," which, under U.S. patent law, excuses some literal infringements if the accused product displays differences in specific activity and purity from the patented product. (The doctrine of which this is called the "reverse" sustains a claim of infringement against an accused protein somehow differing from a patented protein if no functional differences obtain between them.)

A more orthodox industrialist objection to the *HMP* would be to say that without product patents, businesses will not invest the millions of dollars needed to find a gene and to produce a protein by recombinant methods. This bluff is handy because it is counterfactual. As earlier indicated, when one looks at the relatively scant evidence of inventive behavior without patents, and then conjectures about what happens if only process and not product patents are available, one may be sceptical about the claim that biotechnology cannot thrive without product patents. The effective protection afforded by process patents depends on how easy it is to design around a process. Large, complex proteins found in humans may be more difficult to design than, say, pharmaceuticals. Biotechnology patents are replete with process claims. It appears that firms have found ways to protect their intellectual property even though patent examiners vary in their view of product *vis-á-vis* process claims, and even though, given how often courts invalidate them, the status of any product patent is contingent. It must be granted that process patents are often less convenient to enforce because a patentee must show what transpired in a rival's plant. Even so, if a patent has been issued on a recombinant process, ordinarily the recombinant result has only a very low probability of naturally occurring. The patent holder may invoke this probability to refute a defendant's claim to have bred transgenics without using a patented process and without using offspring of the patented process.

One previous motivation for a U.S. product patent is now obviated. When the Harvard mouse emerged, anyone could avoid infringement of a U.S. process patent by performing the patented process in a foreign country outside the reach of U.S. law and then importing the end product into the United States; no such move would defeat a product patent. A statutory amendment changed this by declaring that any such importation is

an infringement of the process patent (33). By virtue of the Biotechnology Patent Protection Act, one may obtain a process patent on a recombinant process that uses or makes a patented product, although this piggyback rule will be moot if product patents become unavailable. Instead of this piggyback rule, it might better be declared that a patent is available on an invented process if what the process uses or produces would be patentable but for the fact that the product is a human life form. Such is the effect of the *HMP*. It allows a process claim to be predicated upon a human life form while allowing no claim on the life form itself.

It remains necessary to show an ingenious process. An industrialist may object that there seem to be few new processes to invent, that current biotechnology employs standard processes that differ only by genes expressed. Mere substitution of a different gene in a known process may indeed be perfunctory. It would not seem to state an argument for product patents to say that innovation is difficult. Opportunities for process innovations abound. The Cohen-Boyer patents expired in 1997. It may simply be that the challenge of finding genes commands more attention at present.

The *HMP*, subsidies, and a period of exclusive FDA drug approval could be implemented separately or together.

ANCILLARY MECHANISMS

Compulsory Licensing

A patent subjects society to the vagaries of a monopolist's choices and fortunes. A possible protection against such risk with respect to biological patents is compulsory licensing according to which anyone may use a patented process upon payment of no more than some reasonable royalty. Another protection is ceilings on the prices of goods made by patented processes. As early as the federally supported Cohen-Boyer research, the NIH considered seeking patents on funded innovations. NIH asserted patent rights to AZT based on the research contributions of NIH intramural scientists, all with the declared purpose of restraining prices of products. This prompts the suggestion that a government agency other than the patent office be empowered to determine what events trigger, and the royalty rate of, a compulsory license established as a condition of any biotechnological patent. A further condition might empower the agency to set maximum prices on goods produced and processes performed in the practice of the patent. An ideal scheme would foster commercial incentives and allow a reasonable return on investment while preventing exorbitant prices.

Such a scheme, it may immediately be objected, would interfere with markets. The industrialist might contend that governments should not restrain returns on genetic inventions since they do not restrain prices of patented artificial hearts or organ transplants. One might reply that when a government grants the privilege of selling a drug or medical product, or of enjoying a monopoly on anything importantly related to human health, the public interest may justify conditioning the privilege on end product price restraint. Compulsory licensing would also protect against disasters with respect to things other than price. As earlier noted, the patentee of the sole therapy for a serious disease could become bankrupt or for other reasons decline to practice or license the invention. The common weal may demand that the invention be available. The industrialist's appeal to the case of an organ transplant does not provide a persuasive counterexample against a compulsory license because organs are donated and recipients pay only for services. An artificial organ is not perfectly analogous to a gene since the organ lacks person-defining genetic information. In any case there may be good reasons to interfere concerning any commerce in human parts.

Expert Guidance

A U.S. patent is only presumptively valid. Since courts often invalidate patents, no one knows for sure that a patent is valid until and unless it is upheld in court. Consider how numerous are the courts within the sovereignties that comprise the international biotechnology market. Trial courts decide only questions placed before them by a flow of cases that is nearly stochastic. The same is true for appellate courts on which depend the prospects of resolving conflicts among trial courts. In contrast to scientists for whom dialogue is a way of life, judges of different courts do not, as a matter of decorum, communicate with each other on pending cases. The science on which they rule is also limited to that practiced a few years, if not a decade, before trial. This obtains because time of invention is the reference point for what is obvious. Hence judicial decisions provide uncertain guidance about patentability of today's scientific processes. Moral issues, as we earlier saw, are not even tackled.

It seems improbable that any one word such as "nonobviousness" or "ingenuity" can bear the load of defining what is a sufficient feat to merit a monopoly. For instance, a claimed invention might be a *tour de force* of genetic engineering, even though the investigator knows neither a sequence's chromosomal locus nor the sequence's coding or regulatory function, if the investigator correctly infers that the sequence is involved, by homology or otherwise, in the etiology of a disease. To transform "ingenious" from placeholder to admission ticket, we may have to settle for a notion of family resemblance. For if ingenuity were to admit of precise definition, would anything be ingenious?

To meet the difficulty of recognizing ingenuity as science progresses, to overcome the lag between research and adjudication, and to improve upon the limited expertise brought to bear in patent adjudication, a mechanism could be confirmed for introducing scientific expertise. A government agency, otherwise involved in scientific research, could exercise authority continually to revise published standards for patenting life forms in reliance on recommendations of expert scientific panels. For purposes of judicial review, the law could preserve the practice of judging a patent by the standards in effect at the time of alleged invention. From such expertly framed standards, the biotechnology industry could obtain guidance more

current and systematic than case law or statute is likely ever to be.

IMPLICATIONS OF PATENTS IN THE CLINIC

Introduction of Human Substances Outside the Germ Line

Ex vivo somatic intervention involves removing patient cells (e.g., tissue-infiltrating lymphocytes or bone marrow stem cells), growing them in culture, transferring genes into them using nonvirulent retroviruses or otherwise, and reintroducing the cells into the patient's body, not necessarily at the site of their effect. *In vivo* intervention is exemplified by the introduction of retrovirus vectors containing human genes at the site of the condition to be overcome. The ED, which would allow patents on substances isolated from the human body, would permit, while the *HMP* would deny, a patent on such cultured cells or vectors. They are ineligible for an *HMP* because they are or contain human life forms. Indeed the cultured cells are grown from the patient's. Except for attempted enhancement, the cultured cells would be unlikely candidates for "inventions" anyway. They are not intended to be innovations. The goal of therapy is to insert a normal gene. The ultimate achievement is homologous recombination. Thereby a normal gene replaces a defective one rather than entering the genome at an indeterminate locus.

Somatic interventions involve medical procedures on patients. Medical treatment, surgery, and diagnosis are not patentable in Europe (34). Their eligibility for U.S. patents has been dubious since 1862 when a patent was sought on the use of ether. It has seemed to many that it would be wrong to discourage physicians, on pain of infringement, from deploying in the relief of human suffering the most efficacious procedures they can muster under the exigencies they face. Hence one might conclude that the only processes of somatic intervention that may qualify for an *HMP* are ancillary laboratory processes. Similarly might patents be confined to laboratory processes with respect to tissues or organs grown in cell culture — especially if, as may be typical to avoid rejection, the cultured cells are grown from the patient's. Opportunities for process innovations would appear abundant when one apprises present difficulties in somatic cell therapy and the challenge of growing tissues and organs.

A contrary moral view might be that the foregoing is too generous. Suppose that one opposed patents on reproduction of any sort. One might assimilate the culturing of cells to reproduction, thereby reversing the patent law's assimilation of reproduction to manufacture. One might add that a laboratory process ancillary to a medical treatment is indistinguishable for these purposes from the treatment. The contention that human reproduction cannot be an infringement does not entail any claim about what is human reproduction. One might conclude that a patent on growing cells outside the human body does not threaten any patient's autonomy so long as there is no claim on the cells themselves. The difficulty of developing successful methods of somatic cell therapy, and of cultivating tissues and organs, suggests the benefit of patent incentives. One need not claim that laboratory processes ancillary to medical procedures are in general nonmedical. One need only allow some of them to be patentable.

Human Germ Line Intervention

Germ line intervention affects reproduction in two ways. (a) It alters the genome, an offspring's complement of genes that appear in all cells including the gametes. (b) In order to achieve (a), it is performed before germ and somatic cells of an individual differentiate, i.e., on zygotes and early stage embryos. A moral objection might be lodged against a patent on any such method because of these links to reproduction. As noted, the ED would allow no patent on any method of human germ line intervention. Again a reply may be that collective benefit could result from creating patent incentives on certain laboratory processes. It is also noteworthy that a patent on gene therapy would not be a patent on *in vitro* fertilization. Therapy is subsequent to fertilization. The choice to conceive may be seen as a different choice than the choice whether to intervene genetically for the health of a child whose conception has been chosen, even if the former is contingent on the latter. On the other hand, the opposite may be the case if eggs fertilized *in vitro* are screened for genetic defects or traits, thereby exercising a choice of which shall live. Two *in vivo* methods also merit mention. One consists in altering an embryo *in utero* by retroviral infection. Another consists in causing adult testes or ovaries to produce genetically engineered gametes (35). A requested European patent on the latter technique was criticized as contrary to *l'ordre publique* or morality (36). For these also one may ask whether the prospect of collective benefit suffices to warrant property claims on ancillary laboratory processes of medical procedures.

If government grants patents on any germ line interventionary process, does that comport with the stance that human reproduction cannot be infringement? The answer lies in stipulating that no remedy will lie against a parent or child as such. Damages and preconception injunctive relief could be made available against unlicensed providers of patented processes. If Mr. and Mrs. Thurston, learning of Mendipulate Inc.'s patented technique for germ line manipulation, arrange with their physician for the technique but no one pays the royalty, a damage remedy may lie against the providers. We can scarcely imagine a suit by Mendipulate against Mrs. Thurston, her daughter or granddaughter, or their physicians or hospitals, complaining of the conception of a child, not to mention injunctive relief, i.e., an order for an abortion. Mere pragmatism makes clear that Mendipulate's interests require no remedy against a patient. Drug manufacturers do not sue patients who infringe by "using" an infringing drug. They sue rival manufacturers and distributors who "make" and "sell" the drug in quantity.

Mendipulate may protest that if it cannot obtain a product patent, every Thurston descendant will benefit

from Mendipulate's invention without paying for it. Mendipulate is correct that the *HMP* allows claims on reproducing the progeny of transgenesis only for nonhuman reproduction. But consider that Mendipulate will advertise a patented process of germ line therapy as a method to remedy a genetic defect. It cannot tenably assert that if it had a product patent, many Thurston descendants would become good-paying customers when they inherit the defect! Moreover, whether the process is therapy or enhancement, Mendipulate's twenty years of monopoly will run before any transgenic Thurston reaches adulthood. Mendipulate may still complain that if Mr. or Mrs. Thurston undergoes a patented Mendipulate process that causes them to produce genetically engineered gametes, no more compensation will be gotten by Mendipulate if the Thurstons have a dozen children than if they have one. This of course overlooks the difference between having children and copying a patented contraption for profit. People are not motivated to have children because they can copy a gene for free. Mendipulate may anticipate fecundity when it prices the royalty for its laboratory process.

Since interventions will be performed by physicians, enforcement of a process patent will require showing what happened in the doctor's office. To Mendipulate this will seem inconvenient. It would prefer a product patent whose infringement it could establish by comparison of parental and progeny DNA. Such a comparison would be peculiar, to say the least, as it would be mustered in support of a complaint that a child is healthy or possessed of some enhancement. It should suffice to protect Mendipulate that licensed specialists may generally be expected to pay royalties on patented processes. What would be troublesome would be the enterprising move of a patient who sells gametes that contain altered genes. This concern may be minimized for the moment by realizing that only enhancement genes, not corrected disease-causing genes, would be likely to be marketable.

Society might deem the collective benefit of enhancement to be insufficient for allowing a patent. If concerns about playing God and discrimination prevail, refusing patents on enhancement would be a means to discourage the practice. A contrary view might be that if we demarcate certain interventions to be outside the physician's armamentarium for maintaining health, no public policy will be disserved by a patent.

There remains possible a product patent on a synthetic gene nowhere found in humans. To use such a gene might depart from the present vision of installing normal in lieu of defective genes. The prospect of such departures no doubt explains the habitual mention of Frankenstein when observers discuss germ line intervention. Regardless, the immunity of parents and children as such from claims of infringement would control. The inventor of a human genetic intervention surrenders the product of the process for integration into an unownable being. If a process alters an early stage embryo, integration occurs into a human in gestation. If alterations are made in gametes or the means of their production, integration occurs into the body of the patient.

CONSISTENCY OF POLICY FOR PLANTS AND ANIMALS

Unless policies about forms of life evince a consistent understanding of innovation and reflect generalizable moral principles, a stable consensus seems unattainable. Conditions (*1*) and (*2*) above stated for a biotechnology product patent — low likelihood of finding the claimed invention in nature, and ingenuity — appear applicable to any life form patent. Some transgenic plants and animals may be improbable of natural occurrence and recognizable as the products of ingenuity. Others may possess transgenes from members of their own species for which the odds of acquisition by mutation are better than trivial, or as to which the process of transgenesis is not ingenious. Where a product patent would be unwarranted, a process patent could be available. As may an *HMP*, a process patent could claim a process by reference to an identified plant or animal life form. It could add a claim on the breeding of any plant or animal that the patented process produces and without which the process would not be patentable. Such an additional claim would obviate the self-reproducibility rationale for a transgenic product patent.

One may argue for bounding a patent's enforceability by operation of a "farmers' privilege," a derogation imposed for plants in the United States and often proposed for animals there and in the ED. This permits a farmer to breed patented animals to the extent needed to replenish stocks on the farm, or to plant seeds generated by transgenic plants grown on the farm. A farmers' privilege would avail a typical farmer who does not seek to compete with breeders in selling varieties as such but who wishes to sell what is raised on the farm. The derogation would entail that, as Mendipulate must do concerning the Thurstons, commercial breeders must collect their royalties on the first generation.

ACKNOWLEDGMENT

This article is adapted from the author's *Norms for Patents Concerning Human and other Life Forms* (14).

BIBLIOGRAPHY

1. P.J. Federico, *Science* **86**, 327 (1937).
2. Directive 98/44/EC of the European Parliament and of the Council of 6 July 1998 on the Legal Protection of Biotechnological Inventions, *Off. J. Eur. Commun.* **L213**, 13–21, Article 6.1 (1998).
3. M.A. Ryan, *J. Med. Philos.* **20**, 419–438 (1995).
4. Council of Europe, Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine, Eur. Treaty Ser. 164, Article 13.
5. W. Gardner, *J. Med. Philos.* **20**, 65–84 (1995).
6. *Diamond v. Chakrabarty*, 447 U.S. 303 (1980).
7. *Ex parte Allen*, 2 U.S.P.Q. 2d 1425 (1987).

8. U.S. Pat. 4,736,866 (April 12, 1988), P. Leder (to the President and Fellows of Harvard College).
9. 35 U.S.C. §101.
10. European Patent Convention Article 53.
11. M. Tallacchini, *Bibl. Della Liberta* **33**(147), 21–50 (1998).
12. M.S. Greenfield, 44 *Stanford Law Rev.* 1051–1094, 1058 (1992).
13. J. Bentham, *A Manual of Political Economy* in J. Bowring, ed., *The Works of Jeremy Bentham*, vol. 3, Russell and Russell, New York, 1962, p. 71.
14. L.M. Guenin, *Theor. Med.* **17**, 279–314 (1996).
15. U.S. Pat. 4,237,224 (December 2, 1980), S. Cohen and H. Boyer (to Trustees of Leland Stanford Jr. University).
16. B.D. Davis, *Science* **193**, 442 (1976); *Am. Sci.* **65**, 547–555 (1977); *The Genetic Revolution*, Johns Hopkins University Press, Baltimore, MD, 1991, pp. 21, 243.
17. P.G. Ducor, *Patenting the Recombinant Products of Biotechnology*, Kluwer, London, 1998.
18. *Amgen v. Chugai Pharmaceutical Co., Ltd.*, 706 F. Supp. 94 (D. Mass. 1989); 13 U.S.P.Q. 2d 1737 (D. Mass. 1989); *aff'd in part and reversed in part*, 927 F. 2d 1200 (Fed. Ct. 1991), *cert. denied*, 112 U.S. 169 (1991).
19. *Scripps Clinic and Research Foundation v. Genentech, Inc.*, 666 F. Supp. 1379 (N. D. Cal. 1987); 678 F. Supp. 1429 (N. D. Cal. 1988); 707 F. Supp. 1547 (N. D. Cal. 1989); *aff'd in part and reversed in part*, 927 F. 2d 1565 (Fed. Cir. 1991).
20. M. Svatos, in S. Sterckx, ed., *Biotechnology, Patents and Morality*, Ashgate, Aldershot, 1997, pp. 291–305.
21. Y. Miki et al., *Science* **266**, 66–71 (1994).
22. D. Dickson, *Nature* **378**, 424–425 (1995).
23. R. Wooster et al., *Nature* **378**, 789–792 (1995).
24. 35 U.S.C. §101–103; European Patent Convention Article 52 (expressing the last two criteria as "industrial application" and "inventive step," and, corresponding to the notion of "manufacture," excluding "discoveries" of nature).
25. *Merck & Co., Inc. v. Olin Mathieson Chemical Corp.*, 116 U.S.P.Q. 484 (4th Cir. 1958).
26. *In re Deuel*, 34 U.S.P.Q. 2d 1210 (Fed. Cir. 1995), discussed at length in (17), pp. 70–90, 141–142.
27. P.L. 104-41, adding 35 U.S.C. §103(b).
28. 21 U.S.C. §§360aa–360cc.
29. 21 C.F.R. §316.3 (13).
30. 21 U.S.C. §360cc.
31. D.W. Plant, N.J. Reimers, and N.D. Zinder, eds., *Patenting Life Forms*, Cold Spring Harbor Lab., Cold Spring Harbor, NY, 1982, pp. 75–78, 95–106.
32. U.S. Pat. 4,468,464 (August 28, 1984) and 4,740,470 (April 26, 1988), S. Cohen and H. Boyer (to Trustees of Leland Stanford Jr. University). See N. Reimers, *J. Assoc. Univ. Technol. Managers* **7**, 25–47 (1995).
33. 35 U.S.C. §271 (g) added by P.L. 100-418.
34. European Patent Convention Article 52 (4).
35. R. Brinster and M.R. Avarbock, *Proc. Natl. Acad. Sci.* **91**, 11303–11307 (1994).
36. *Nature* **368**, 572 (1994).

See other entries HUMAN GENOME DIVERSITY PROJECT; MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT; OWNERSHIP OF HUMAN BIOLOGICAL MATERIAL; see also PATENTS AND LICENSING entries.

PHARMACOGENETICS

MICHAEL M. SHI
MICHAEL R. BLEAVINS
Warner-Lambert Company
University of Michigan Medical School
Ann Arbor, Michigan

OUTLINE

INTRODUCTION

Pharmacogenetics is a new and quickly evolving scientific discipline that studies how genetic differences determine an individual's response to therapeutics. Recent advances have shown that many drug-metabolizing enzymes have genetic variations in expression and regulation. When these genetic variations affect the enzyme's function significantly, different clinical outcomes can occur among people exposed to a particular drug. Pharmacogenetics can help to individualize dosing regimens, thereby maximizing a drug's therapeutic effect and minimizing toxicity. The rapid development of genotyping as a molecular diagnostic tool nevertheless raises ethical, legal and policy issues, as have been the case with other DNA-based testing. Although pharmacogenetics shares concerns with other genetic research in clinical practice, this discipline also has unique objectives

and goals. It therefore is important to address the issues associated with this emerging discipline. The purpose of this article is to describe the concepts and advances in pharmacogenetics, discuss social implications of this field and provide recommendations in this area.

PHARMACOGENETICS

Genetic Variations in Drug Response

Significant variations in drug response exist among both populations and individuals. These variations can be due to genetic and/or environmental factors. Age, gender, body size, diet, alcohol or tobacco consumption, pregnancy, kidney or liver dysfunction, concurrent disease states, and drug interactions can all modify the bioavailability, distribution, protein binding, metabolism and excretion of drugs (1). Interindividual variations in drug response also result from genetically based differences in drug metabolism. Individual differences in the absorption, distribution, metabolism, and excretion of therapeutics can alter the effects of a given dose, leading to a spectrum of responses ranging from clinical benefit to adverse effects or therapeutic failure. When taken up by the human body, foreign compounds typically undergo metabolism via Phase I and II enzyme reactions (2). Phase I reactions involve hydrolysis, reduction, and oxidation. These enzymes introduce a functional group (i.e., hydroxyl, amino, sulfhydrol, or carboxyl) to the compound, usually increasing water solubility. Within the category of Phase I enzymes are the cytochorome P450 multigene family, AND(P)H quinone oxidoreductase, and aldehyde dehydrogenase. Phase II metabolism includes glucuronidation, sulfation, acetylation, methylation, glutathione conjugation, and amino acid conjugations. Most Phase II reactions dramatically increase hydrophilicity, thus greatly enhancing the excretion of foreign compounds. Phase II enzymes include UDP-glucuronosyl transferases, glutathione-*S*-transferases (GSTs), sulfotransferases, catechol-*o*-methyltransferases, phenol-*o*-methyltransferase, and thiol methyltransferases (3). Recent research has identified functionally important variations in most Phase I and II enzymes, which can lead to different metabolic profiles among individuals.

Genetic Polymorphisms of Drug-Metabolizing Enzymes

Genetic polymorphism refers to distinct traits derived from a single gene that exists in more than one form. These polymorphisms are transmitted from generation to generation, sometimes with striking differences in allele distributions among different ethnic groups (4). A majority of the genetic polymorphisms do not affect protein function and therefore have no phenotypic importance. Other polymorphic genes encode drug-metabolizing enzymes with dysfunctional or nonfunctional activity. As a result the subgroup of the population with the genotype(s) will metabolize drugs that are eliminated via this pathway differently from individuals with the normal (wild-type) genotype. A classic example of genetic polymorphism is class 2 aldehyde dehydrogenase (ALDH2). Approximately 50 percent of the Asian population has a single amino acid change of Glu^{487} to Lys^{487} in this enzyme, causing impaired acetaldehyde metabolism (5). These individuals can rapidly convert ethanol to acetaldehyde but only slowly metabolize acetaldehyde to acetic acid. Affected people experience of flushing syndrome after alcohol consumption due to the release of catecholamines triggered by the sustained high blood acetaldehyde levels, which does not normally occur in people with the fully functional form of the enzyme. Pharmacogenetics emerged as a discipline to study genetic variations in drug-metabolizing enzymes that may determine an individual's responsiveness to therapeutic agents (6,7). Advances in this area have important clinical implications and practical values for the design of dosing regimens. It is important to recognize that the genetic variations in drug metabolism can lead to significantly different therapeutic responses, including either low or exaggerated pharmacological effects or side effects.

An increasing number of drug-metabolizing enzyme polymorphisms have been identified in recent years (8). Examples of Phase I and II enzymes with functionally important polymorphisms are listed in Table 1 (9–22). A well-characterized drug metabolizing enzyme with functionally important variants, cytochrome P450 2D6 (CYP2D6, debrisoquine hydroxylase), will be highlighted. The genetic polymorphisms of CYP2D6 are perhaps the most well-established alterations with known clinical significance. This microsomal isozyme is responsible for the oxidative metabolism of approximately 50 clinically important drugs of varying therapeutic classes (2). Its substrates include widely used antiarrhythmics, tricyclic antidepressants, β-adrenergic blocking agents, neuroleptics, and other classes. These drugs frequently have narrow therapeutic windows, meaning that slightly lower than targeted plasma concentrations will not have the desirable therapeutic effect while only somewhat higher concentrations cause toxicity.

CYP2D6 has three clinically distinct phenotypes: (*1*) the normal (or extensive) metabolizers, (*2*) slow (or poor) metabolizers, and (*3*) fast (or ultraextensive) metabolizers. The same dose of a drug metabolized via CYP2D6 will result in plasma concentrations that vary greatly among these individuals. In normal metabolizers, steady-state plasma concentrations will normally fall within the desired therapeutic range and toxic effects will be non-existent or minimal. In fast metabolizer individuals, steady-state drug levels will be below therapeutic concentration and this group of patients is unlikely to respond to standard treatment regimens. It has recently been reported that the cholesterol-lowering drug simvastatin did not reduce plasma lipid levels in CYP2D6 fast metabolizers at standard doses (23). While it may be possible to increase the dose given to these patients and achieve the same therapeutic effect as normal metabolizers, this approach also increases the potential for undesirable side effects, particularly those not related to CYP2D6 metabolism. In slow metabolizing individuals, plasma drug concentrations will be significantly above therapeutic levels when

Table 1. Example Polymorphic Enzymes in Drug Metabolism and/or Disease Susceptibility

Enzymes	Enzyme Reaction	Phenotype
CYP1A1	PAH oxidation	FM associated with lung cancer in smokers (9)
CYP2C9	Tolbutamide hydroxylation	SM for tolbutamide (10)
CYP2C19	S-Mephenytoin hydroxylation	SM for mephenytoin and other drugs (10)
CYP2D6	Debrisoquine hydroxylation	SM and FM for over 50 clinically important drugs (11)
CYP2E1	Chlorzoxazone hydroxylation	Associated with lung cancer (12)
ADH2	Ethanol metabolism	SM for alcohol metabolism (5)
NQO1	Quinone reduction	Associated with urological (13) and lung cancer (14)
GSTM1	Conjugation of epoxide	Gene deletion associated with lung (15) and bladder cancer (16)
NAT2	Acetylation	SM and FM for isoniazid and other drugs (17) associated with bladder and colon cancer (18)
UDPGT1A1	Bilirubin conjugation.	Deficiencies in Crigler-Najjar (19,20) and Gilbert syndrome (21)
TPMT	Methylation	Deficiency associated with mercaptopurine and azathiopurine toxicity (22)

Abbreviations: FM, fast metabolizer; GSTM1, Glutathione S-transferase M1; NAT2, *N*-acetyltransferase 2; NQO1, NADPH-quinone oxidoreductase 1; PAH, polycyclic aromatic hydrocarbons; SM, slow metabolizer; TPMT, Thiopurine S-methyltransferase; UDPGT, UDP-glucuronosyltransferase.

conventional doses are used. In this case undesired toxicity can proceed or mask the desired pharmacological effects, and these individuals are likely to suffer adverse side effects. This is particularly true for many antipsychotic and antidepressant agents (11). In other instances, the parent drug requires biotransformation to an active form. For example, the analgesic effect of codeine largely depends on its conversion to morphine through *o*-demethylation, and so adequate analgesic effect cannot be achieved in CYP2D6 slow metabolizers (24). Therefore clinical practice in the future may benefit from dose individualization to avoid toxicity or achieve optimal therapeutic benefit.

The human CYP2D6 gene consists of nine exons and has been mapped to chromosome 22 (25). After transcription, the premature mRNA undergoes splicing and only the exonal region encodes the protein synthesis. The poor metabolizer phenotype occurs in 7 to 10 percent of the Caucasian population (11) and results from autosomal recessive inheritance of nonfunctional alleles (7). In addition to the wild-type gene (CYP2D6* 1), over 20 different alleles of CYP2D6 are associated with deficient, reduced, or increased enzyme activity (26). The most frequent inactivating mutation among Caucasians is the CYP2D6* 2 genotype, a splice–site mutation involving $G_{1934}A$ transition in the 3′-end of intron 3, leading to a mis-splicing of the premature transcript, and loss of enzyme activity (27,28). The CYP2D6* 3 mutation is a 1-bp A_{2637} deletion in exon 5 leading to a frame–shift change in the translation of CYP2D6 mRNA (29). The CYP2D6* 5 mutation is a deletion of the entire CYP2D6 gene (30). The CYP2D6* 2 mutation constitutes about 75 percent of all mutant alleles, with the CYP2D6* 5 mutation responsible for 14 percent and the CYP2D6* 3 mutation for 5 percent. Together these three polymorphisms account for approximately 95 percent of the slow metabolizer genotypes (29).

Detection of Genetic Polymorphisms

Standard procedure for evaluating metabolic capacity involves administration of a probe compound and measuring the ratio between the parent drug and its metabolite in urine and/or plasma. This procedure involves analytical techniques and often requires a week or more before a conclusion can be drawn. Metabolic phenotyping has additional disadvantages in that results can be influenced by sample stability and that conversion of the drug can be affected by external factors, including age, nutrition, general health, and other medications. Furthermore some poor metabolizers experience unpleasant side effects of the probe drugs (11). Therefore metabolic phenotyping is not widely used in clinical practice. These limitations can be circumvented for many enzymes by genotyping the patient at a centralized laboratory.

Genotyping is relatively easy to perform and generally requires only a sample of peripheral blood from patients. Therefore it is potentially less invasive than phenotyping and is not influenced by drug–drug or drug–food interactions. If a polymorphic site changes the DNA recognition sequence of a restriction enzyme, or if the genetic polymorphism involves a large deletion or insertion, the genetic polymorphism can be identified using polymerase chain reaction (PCR) coupled with restriction fragment length polymorphism (PCR–RFLP) analysis. In this approach, DNA sequences containing the polymorphic site are amplified by PCR, followed by restriction digestion and gel electrophoresis. PCR-RFLP tests have been developed to detect most of the CYP2D6 mutations (26). Other genotyping methods include allele-specific PCR (11), fluorescent dye-based high throughput genotyping (31,32), and the recent gene chip technology (33,34). Genotyping for genetic polymorphisms in drug metabolism can help explain drug toxicity or therapeutic failures, and help predict potential drug interactions. The addition of pharmacogenetic testing will

help clinicians to manage drug therapy, especially for drugs with low therapeutic indices.

Prescriptive Medicine

Interindividual variability in uptake and metabolism of many drugs make clear dose–response relationships for these compounds difficult to predict and toxic side effects a real possibility. A dose that produces the desired therapeutic response in one individual may be toxic or subtherapeutic for another person. Therefore it would be valuable to know in advance the dose of medication to prescribe based on each individual's metabolic capabilities. Pharmacogenetic testing can provide a powerful tool for optimizing therapeutic efficacy and reducing drug toxicity for those compounds known to be metabolized via pathways with functionally important genetic polymorphisms. It has been estimated that a typical marketed drug is efficacious in approximately 20 to 40 percent of patients (35). The same substance is likely to have no therapeutic benefit for 20 to 40 percent of patients, and to cause significant side effects in 20 percent of patients (35). Adverse drug effects negatively impact the health of patients and their quality of life, as well as adding substantial financial burden to society in terms of costs and acceptance of the compound to treat disease. More effectively predicting drug efficacy and toxicity offers significant benefits to society and may be described as "prescriptive medicine." Genotyping for relevant DNA markers may help physicians prescribe the most efficacious drugs for a given individual and disease, while minimizing adverse drug side effects. Patients also could be provided an opportunity to choose among therapeutic agents, both prescribed and over the counter, to obtain suitable medication and avoid undesired side effects. If genotyping and phenotyping can predict which patients are most likely to benefit from a specific treatment, only those people need be exposed, with individuals likely to experience adverse or ineffective responses not administered the drug. In addition health care costs can be reduced, especially for drugs that must be taken for extended periods, such as hypocholesteremic, antihypertensive, and antidiabetic drugs. Once validated scientifically and proved to be cost effective, pharmacogenetics will provide significant benefits to both patients and society. In the future physicians may have the opportunity to prescribe the most effective and safest drug based on the patients' genetic blueprint.

Pharmacogenetics may also help drug makers to design therapeutic agents that specifically target patient subpopulations. A recent study (36) found that a drug given to 400 Alzheimer's patients had no statistically significant effect. However, when patients were stratified according to ApoE subtype, a clinically significant response was demonstrated. Due to the heterogeneity of human populations, genetic stratification can be the difference between a drug's success and failure. Pharmacogenetics thus can assist in identifying safer and more efficacious drugs by targeting subpopulations of optimal responders and predetermining those most at risk of undesirable reactions. It also holds great potential for accelerating the drug discovery process by providing clearer answers, as well as reducing the length and cost of clinical trials.

Ecogenetics and Preventive Toxicology

Ecogenetics is a broader definition of interindividual differences in response to environmental toxic chemicals (37). Just as in the metabolism of therapeutics, some slow metabolizers might detoxify environmental or occupational agents significantly slower than normal populations. By the same token, fast metabolizers may more readily activate some foreign agents to their toxic intermediates. Therefore certain allelic forms of drug-metabolizing enzymes could render an individual either more sensitive or more resistant to the toxic effects of specific classes of foreign compounds. For example, molecular epidemiology studies have identified associations between specific genotypes of CYP1A1, CYP2E1, and GST-M1 with a variety of cancers (9,12,38). By defining these susceptibility genes, those people at increased risk can be advised to avoid certain exposures and safer standards established for workers and the public. An informed decision to avoid exposure to some occupational hazardous materials could be made by individuals whose genotypes had been associated with cancer or diseases. Clearly, there are also uncertainties and controversies in the use of these genotyping and association studies. For example, an association between CYP2E1 DraI genetic polymorphism with increased lung cancer susceptibility was suggested in a Japanese population (12) but was not observed Caucasian populations (39,40). This discrepancy may be due to a significantly low frequency of CYP2E1 Dral polymorphism in Caucasians (41). Clearly, more critical research has to be done to effectively use the research and epidemiological data emerging from this area of study.

ETHICAL, LEGAL, AND POLICY ISSUES ASSOCIATED WITH PHARMACOGENETICS RESEARCH

Most genetic research and tests share similar ethical, legal, and policy issues. The most common concerns involve risks of psychological distress, loss of insurance or employment, as well as confidentiality of genetic information. Since the objectives and goals of pharmacogenetics are different from other genetic research, it is important to discuss these issues separately. Up until now, only a few ethical issues have been briefly touched upon (42). There are potential controversial aspects such as informed consent, confidentiality of genetic information, sample and data ownership, potential discrimination against people identified as genetically "deficient," and access to human genetic materials and information. This section will review the current status in the genetic field and provide some recommendations for pharmacogenetics.

Informed Consent

As pharmacogenetic research requires population-based sampling for genetic variation and gene–environmental interaction studies, it frequently involves collection of large numbers of volunteer and patient samples. In order to obtain the testing materials, researchers need the informed consent of test subjects. Informed consent involves a process of education and counseling that facilitates voluntary, reasoned decision making. The

prospective participants or patients must understand the purpose and the nature of the study or test, understand his/her role in that study, and be cognizant of the benefits and risks that may result from the study. The document should be comprehensive, easily understandable, and serve as the means to protect both patient and care provider.

There is a growing belief that genetic information is particularly sensitive and that some people may not want genetic information about them obtained, even in therapeutic indications. These concerns must be addressed in order to protect individual's rights, pursue important research, fulfill medical ethics, satisfy regulatory requirements, and benefit society. Significant controversy also surrounds the ethical issues associated with archived blood and tissue samples used for molecular genetic testing for either basic research or clinical parameters (43). With PCR technology, even material from archived paraffin-embedded tissue or frozen blood can yield sufficient DNA for genetic analysis. These materials have tremendous value for pharmacogenetic research, particularly when a new genetic polymorphism is identified. Researchers then have the possibility of retrospectively identifing banked samples and conducting genotyping that establishes whether or not a genotype/phenotype correlation exists. This raises concerns about invasion of privacy, loss of individual autonomy, and stigmatization if test results were released. At the same time, however, a balance needs to be achieved that allows researchers access to these human samples for improving disease treatment and test validation. Another compelling reason to use archived DNA is that some historical samples are extremely valuable in that it may not be possible to reconduct the same study or collect new samples. In addition, with the continuing progress in human genetics, these specimens may have significant value in furthering medical discoveries beyond even those currently envisioned. A recent article offers the suggestion to treat the test samples, however obtained, as anonymous by keeping the patient-specific portion of molecular genetic test results confidential from even those scientists conducting the evaluation (44). A general guideline on informed consent for genetic research has been provided by the American Society of Human Genetics (45).

Recommendations. When obtaining informed consent for research and clinical testing, it is imperative to clarify the following:

- Purpose. limitation, and potential outcomes of the research
- Methods for maintaining confidentiality of results
- Anticipated use of testing samples
- Duration of storage and disposal of the materials
- Potential for research to lead to new clinical diagnostic tests
- Final publication or disclosure of study results

When it is not possible to give informed consent to an incomprehensive patient, a legally authorized guardian or appropriate decision maker may be substituted. In addition, a general notification of potential future use of the samples should be included. If the patient does not object and if samples are coded or remain anonymous, the DNA may be used for research not specifically defined in the informed consent statement (46). The collection and storage of DNA used for genotyping should follow established guidelines for DNA databanking (47).

Confidentiality of the Research and Test Results

Confidentiality of test results has been and remains a major focus of the ethical, legal, and policy issues related to genetic testing. Pharmacogenetics provides an opportunity to observe a person's molecular genetic makeup independently of the visible characteristics. The genotyping data may reveal asymptomatic conditions that would only manifest with age, or upon exposure to specific drugs or compounds. Genetic testing therefore may allow better diagnosis for disease risk at earlier stages of life.

Pharmacogenetic testing promises to provide value in making diagnostic decisions and assessing medication risk. As described earlier, recent epidemiological studies and animal models also have identified a strong association between some metabolic enzyme genetic polymorphisms and cancer. Although specific associations exist, these studies identify risks, not certainties. In reality the development of cancer is a complicated process and depends on multiple gene–environmental interactions. Obviously more critical research has to be done to clearly establish the biochemical pathways important in carcinogenesis and the role of genetics in susceptibility. On the positive side, knowing this information may help patients prepare for the risk and adjust their work or lifestyle to minimize potential hazardous occupational or environmental agent exposure that could trigger disease. One major concern in this area revolves around the possibility that insurers and employers might regard an increased genetic risk as the final concrete outcome and use this information to establish policies that discriminate among individuals. For example, employers might propose to identify workers with lower genetic risk for toxicity or malignancy from exposure to particular occupational agents and select only those individuals to perform "high-risk" jobs. Genetic information also could be used by employers to predict health care costs or an employee's productivity. These predictions might influence hiring, retention or promotion decisions.

Similarly genotypic information could be used for insurance purposes to weed out individuals at highest disease risk. This information might be used to justify higher premiums or cancellation of policies. It therefore is very important to establish guidelines that prevent abuse of genotyping information. Researchers and clinicians must carefully consider the risks and benefits and the potential impact of genetic information for participants and others. Protection of patient confidentiality may require further protection through legislation (48).

It is advised that patients or volunteers be informed of pertinent aspects related to acquisition, storage, and use of data, as well as the degree to which third parties can obtain access. Researchers should adhere to the principle of least-intrusive disclosure, in which the data are stored

using identifiers such that patient identity and sensitive personal factors connections are not possible, or where the fewest number of investigators necessary to achieve the research goal is maintained (49). A security infrastructure should be in place to ensure the confidentiality of research information, including access control, audit trails, disaster recovery, and encryption of patient-identifiable data before transmission on networks (48).

Recommendations

- Patients should be informed about storage and access to test results
- Study center should safeguard genetic information
- Disclosure of information and access to DNA samples should respect principles of privacy

CLINICAL APPLICATIONS AND ASSOCIATED ETHICAL, LEGAL, AND POLICY ISSUES

Benefits and Risks of Clinical Pharmacogenetic Testing

Recently genetic testing to predict linkage to late-onset diseases such as ApoE in Alzheimer's disease, or BRCA1 and 2 mutations in breast cancer, have resulted in substantial public debate (50). Genetic testing is complicated by uncertainties in predicting and diagnosing these diseases, and more importantly, by the social, ethical, and legal implications of disclosing genotype results.

Unlike the BRCA and ApoE genotyping, where no current therapeutic intervention is available, pharmacogenetic tests may be more acceptable to the public. The reason for this is that there are intervention strategies available, namely either withdrawal of the drug or switching to another compound belonging to a different chemical class. Based on a patient's drug metabolizing genotype, physicians also may be able to adjust the standard doses, thus achieving therapeutic value and avoiding toxicity. In addition to satisfying patients' physical well-being, substantial financial benefits may be achieved by utilizing the data obtained by pharmacogenetic testing. Incentives from managed care organizations and insurers to control health care costs may strongly support these tests. If a pharmacogenetic test accurately predicts that a commonly prescribed drug will be ineffective, or has serious adverse effects for a relative large percentage of patients, these individuals could be given preselected medications to avoid the lack of efficacy or the severe toxicity.

Predictive Value and Limitations of the Tests

Unlike molecular biology tests for pathogens, definitive or absolute results cannot be easily achieved in current pharmacogenetic testing. For example, CYP2D6 has more than 20 polymorphisms leading to altered enzyme activity. Until testing is performed for all polymorphisms, many of which are extremely rare, it is not possible to have an absolutely accurate prediction about every individual's phenotype. It also must be kept in mind that other functional polymorphisms may exist that have not yet been identified. Since pharmacogenetics is a relatively new area, the prediction of results often must be qualified. There are only a handful of pharmacogenetic markers currently available commercially for genotyping, such as CYP2D6, CYP2C19 and ApoE (8). When used properly, they can be valuable for decisions regarding dosing, counseling and prognosis. Some general principles can be expected for clinical testing in pharmacogenetic testing:

- Testing is useful for detecting functionally important polymorphisms of drug metabolizing enzymes
- Techniques used for genotyping are relative simple and noninvasive
- Testing can be expected to benefit the patient and society
- Pretreatment genotyping might be desirable for drugs with low therapeutic indices and high toxicities
- Positive testing results will be more informative than negative ones

Standards in Clinical Testing

Due to the complexity of interpreting genotyping results and the requirement of expensive and complex laboratory equipment, pharmacogenetics is likely to be conducted primarily through service laboratories. Currently the clinical laboratory is required to establish analytic validity and Clinical Laboratory Improvement Amendments (CLIA) certification (51). Yet the rapid progress in human genome research and development in genetic testing technologies have outpaced the quality assurance and quality control for molecular diagnosis. A major concern is that a complete set of specific standards to assure proficiency for genetic testing in the clinical laboratory has not yet been developed, even though genotyping procedures are classified as being of high complexity. A recent report from the NIH–DOE Task Force on Genetic Testing recommended that clinical validity, such as test sensitivity and the predictive value of a positive test result, as well as institutional review board (IRB) approval of the protocol, should be part of the prerequirement of CLIA certification (52). In addition the Task Force suggested an external review before the genetic test can be commercially offered even after CLIA certification (53).

In the future pharmacogenetic testing is likely to fall within the competency testing currently imposed on clinical pathology laboratories as described under CLIA and Commission on Office Laboratory Accreditation (COLA) regulations. In lieu of the current situation, clinical pharmacogenetic tests should be done by laboratories that meet accepted standards of general laboratory quality assurance, including patient test management, personnel requirements, specimen handling, quality control, test validation, and confidentiality. In addition to CLIA requirements, an inspection of the laboratory's competence in performing the genetic test should be conducted by certification agencies. Interlaboratory comparisons of reference samples might be necessary to assure the quality control mechanism, a system already in place for more routine clinical biochemistry and hematology procedures.

Indications and Demands in Pharmacogenetics Testing

Although pharmacogenetic testing could have predictive value in clinical situations, it is not routinely invoked unless patients have previously had severe toxicity and the primary care providers have a knowledge of pharmacogenetics. Predictive tests, such as the association with cancer risk or late-onset genetic disorders, should not be routinely performed until additional research supports a beneficial outcome and effective treatment.

Several important questions should be asked before initiating pharmacogenetic testing. What are the benefits of conducting these tests? What is the most cost-effective way to do this? Are these tests accurate enough to predict the clinical outcome? At present, it is questionable which variant alleles should be routinely genotyped to allow a sufficiently reliable, but still practical, estimation of a person's metabolic capacity. With rising health care costs and the scarcity of resources, it is not acceptable to adopt expensive testing that does not add any value in patient care or new technologies of which the benefits are still unclear (54).

A genotyping test will only be of value to consumers when it can provide predictive value and a reasonably precise answer regarding individual risk. Currently very few pharmacogenetic tests are available that meet these standards. Although some genotyping tests are commercially available, they primarily provide services for research studies and have not been integrated into routine clinical practice. Another limitation is that pharmacogenetics as a discipline has not been integrated into the medical training curriculum and few physicians are familiar with the underlying concepts, benefits, and applications. Giving the progress in pharmacogenetics, it is easily foreseeable that it will soon become another subspecialty in medicine and pharmacy. One complexity is that there can be many genotypes that produce the same phenotype. Attempting to identify all polymorphisms associated with a defective phenotype will significantly increase testing complexity and expense. The practical approach at this time is to screen for the most common genotypes leading to altered enzyme activities. With the rapid development of modern technology in genetic diagnosis, it will be possible to detect multiple polymorphisms in a single test. For example, there is a report indicating a human P450 DNA chip could identify all the currently known polymorphisms of human CYP2D6 and CYP2C19, but its clinical usefulness has yet to be identified (55). However, these chips only detect those variants specifically programmed onto them and modifications to include new polymorphisms are expensive and time-consuming. The progress in fluorescent-based high throughput genotyping and DNA chip technology will definitely add significant value to pharmacogenetic diagnostics. Genotyping tests for the major drug metabolizing enzyme polymorphisms will soon be as easy as a routine blood test.

Cost Effectiveness of Pharmacogenetic Testing

Pharmacogenetics has the potential to be cost-effective in the managed care community. A simple diagnostic genetic test will enable the drug to be selectively prescribed to those patients for whom a drug would be safe and effective. This would provide cost savings to the health care providers by increasing drug efficacy, reduce follow-up and doctor visits, eliminate costly ineffective drugs, and reduce possible drug toxicity at "normal" doses in slow metabolizers. For example, a patient who metabolizes drugs more rapidly than other patients will not respond to the drug treatment under standard dosing. Identifying these rapid metabolizers of the drug could help these patients to either increase the dose or to use other appropriate drugs without undergoing three or four months' trial and error. This will be cost saving for many expensive prescription drugs for treatment of chronic disease such as dislipidemia, diabetes, and Alzheimer's disease.

Avoiding adverse drug effects alone may bring significant savings to society. It is estimated that approximately 3 to 5 percent of all hospitalizations result from adverse drug reactions, and as many as 30 percent of patients hospitalized for other reasons may have an adverse drug effect during hospitalization (56). The cost of treating drug reactions in the United States alone is estimated at approximately $3 billion annually (57).

Current genotyping generally costs up to two to three hundred dollars per test depending on specific assays. With the rapid development of automated high-throughput genotyping technologies, pharmacogenetic testing will likely become a relatively low-cost/high-volume service, just like a routine blood test. It is reasonable to develop cost-efficient pharmacogenetic tests by using multiplex PCR or non-PCR based genotyping technologies. Therefore pharmacogenetics holds great promise for prescriptive medicine, and it is expected that monitoring of pharmacogenetic markers will be routinely used clinically, especially for patients receiving drugs with low therapeutic indexes.

NEW DIRECTIONS OF PHARMACOGENETICS

Genomic Approach in Pharmacogenetic Studies

Studies of families with disease are informative for identifying highly penetrant gene variants. However, other approaches are needed to study less penetrant alleles, which may not be easily identified in family members. This is particularly true for environmental susceptibility genes and drug response genes. Such alleles may identify those people at risk, but who otherwise would only be observable in an exposed population. While some drug toxicities have been identified, the genetic polymorphisms associated with the effects of most drugs have not been characterized. Targeting these genes will be another goal of pharmacogenetics. The new concept of pharmacogenomics will utilize high-density markers to conduct genome scans to better predict drug efficacy and toxicity. In contrast to the candidate gene method, the strength of this system is the ability to scan the entire genome (58). It is becoming increasingly popular to use single nucleotide polymorphism (SNP) for association and linkage analysis, since they are the most

frequent DNA sequence variations found in the human genome. Researchers will be able to conduct whole genome scans for identification of critical drug–response genes in nonfamilial studies. Creation of high-density SNP maps is feasible using high-throughput DNA sequencing (59) and chip hybridization (60). Cataloging common variants in human genes is moving very rapidly (61). It will be necessary for the scientists to prove the technology can work in the real world and transfer the research to clinical practice. Again, with the progress in this area, and as more genes have been identified, additional and increasingly complicated social, ethical, and policy issues will be encountered. Because of potential social consequences, researchers have been encouraged to pursue anonymous testing whenever possible and to ensure that the results of genetic testing are separated from an individual's record.

Education for Clinical Practitioners

The advances in pharmacogenetics will have a significant impact on the practice of diagnostic and preventive medicine. Currently there are only a limited number of medical practitioners familiar with and conversant in this area. It is important for clinicians to understand the concepts and applications of pharmacogenetics, since they will be explaining genetic tests and implications to their patients, determining when testing is appropriate, selecting specific tests, and interpreting the results. It is therefore necessary to develop training programs, that efficiently transfer a working knowledge of this field. Until such programs are available, physicians and pharmacists are advised to contact specialists or consult with colleagues having expertise in pharmacogenetic testing. As pharmacogenetic testing becomes more commonly offered by clinical pathology and reference laboratories, these facilities also will be called upon to provide expertise and appropriate indications.

Patient Stratification in Clinical Trials

For drugs prescribed on a limited basis due to a high incidence of adverse effects, pharmacogenetics may provide means to identify those most likely to benefit therapeutically without serious side effects. By targeting a specific subpopulation, pharmacogenetics offers the possibility of wider and safer drug use by creating a clear prescription path. The use of patient stratification also offers the potential to rescue an existing drug with great promise but undesirable problems, or to provide the data needed to withdraw a dangerous compound earlier. Pharmacogenetics data similarly could be used to logically design clinical trials and to increase the amount of information obtained from these studies. The greatest, but yet unproved, promise of pharmacogenetics is to alter trial-and-error application of a new medication into prescriptive medicine. Differentiating patient groups to improve the risk–benefit ratio of a new drug is already common practice. Therefore the concept of this predictive medicine approach looks attractive and can build on existing principles. But clearly, any tests associated with drug toxicity must be rigorously reviewed before a conclusion is made.

With the emerging global economy, pharmaceutical companies need to market new drugs in multiple countries. Due to the differential distribution of some drug metabolizing enzyme genetic polymorphisms among populations, a well-developed and extensively tested drug might not be suitable for patients with different ethnic backgrounds. In utilizing data obtained from genotyping both ethnic groups, prediction of drug efficacy and toxicity in a different population group become possible, as well as potentially reducing the need to conduct pivotal clinical trials in multiple countries. The most important step for pharmacogenomics now is proof of principle. It is critical to clearly show that pharmacogenetic concepts will yield improved and more predictive results in clinical trials. This process is actively underway.

CONCLUSIONS

The introduction of pharmacogenetic testing into clinical medicine has great promise for affecting the future of prescriptive and preventive medicine. Physicians may be able to prescribe drugs based on genotype, as well as allowing pharmacists to check for potential drug interactions and side effects. It will be important to educate medical practitioners and patients on both the concepts and clinical practice of pharmacogenetic testing. The up-front discussion of social, legal, and policy issues should not be used to block the collection of genetic data, but should serve as a safeguard to benefit and protect patient rights. With the advances from the Human Genome Project and functional genomics, massive increases are taking place in the information available on individual genes and functionally important polymorphisms. These differences hold the potential to improve effectiveness and limit toxicities of the available drugs while providing an understanding of gene/environmental interactions. Consideration of the ethical, social, legal, and policy aspects of accurate genetic prediction, and the design of more specific and safer drugs to meet individual's needs, will be important considerations as we enter the new millennium.

BIBLIOGRAPHY

1. A.S. Nies and S.P. Spielberg, in J.G. Hardman and L.E. Limbird, eds., *Goodman & Gilman's the Pharmacological Basis of Therapeutics*, 9th ed., McGraw-Hill, New York, 1996, pp. 43–62.
2. L.Z. Benet et al., eds., *Goodman & Gilman's the Pharmacological Basis of Therapeutics*, 9th ed., McGraw-Hill, New York, 1996, pp. 3–27.
3. A. Parkinson, in K.C.D. Klaassen, ed., *Casarett & Doull's Toxicology*, 5th ed., McGraw-Hill, New York, 1996, pp. 113–186.
4. U.A. Meyer et al., *Adv. Drug Res.* **19**, 197–214 (1990).
5. H.W. Goedde and D.P. Agarwal, in W. Kalow, ed., *Pharmacogenetics of Drug Metabolism*, Pergamon, New York, 1992, pp. 281–311.
6. W. Kalow, ed., *Pharmacogenetics of Drug Metabolism. International Encyclopedia of Pharmacology and Therapeutics*, Pergamon Press, Oxford, UK, 1992.

7. F.J. Gonzalez and J.R. Idle, *Clin. Pharmacokinet.* **26**, 59–70 (1994).
8. W.E. Evans and M.V. Relling, *Science* **286**, 487–491 (1999).
9. K. Kawajiri et al., *FEBS Lett.* **263**, 131–133 (1990).
10. J.A. Goldstein and S.M. de Morais, *Pharmacogenetics* **4**, 285–299 (1994).
11. M.W. Linder, R.A. Prouch, and R. Valeds, Jr., *Clin. Chem.* **43**, 254–266 (1997).
12. F. Uematsu et al., *Jpn. J. Cancer Res.* **82**, 254–256 (1991).
13. A.S. Wolfgan et al., *Pharmacogenetics* **7**, 235–239 (1997).
14. J.K. Wiencke et al., *Cancer Epidemiol, Biomarkers & Prev.* **6**, 87–92 (1997).
15. J.E. McWilliams et al., *Cancer Epidemiol. Biomarkers & Prev.* **4**, 589–594 (1995).
16. A. Hirvonen, *J. Occ. Env. Med.* **37**, 37–43 (1995).
17. S.P. Spielberg, *J. Pharmacokinet. Biopharmaceut.* **24**, 509–519 (1996).
18. D.A.P. Evans, in W. Kalow, ed., *Pharmacogenetics of Drug Metabolism*, Pergamon, New York, 1992, pp. 95–178.
19. P.J. Bosma et al., *FASEB J.* **6**, 2859–2863 (1992).
20. J.K. Ritter et al., *J. Clin. Invest.* **90**, 150–155 (1992).
21. P.J. Bosma et al., *N. Engl. J. Med.* **333**, 1171–1175 (1995).
22. D. Otterness et al., *Clin. Pharmacol. Ther.* **62**, 60–73 (1997).
23. C. Nordin et al., *Lancet* **350**, 29–30 (1997).
24. I. Johansson et al., *Eur. J. Clin. Pharmacol.* **40**, 553–556 (1991).
25. F.J. Gonzalez et al., *Genomics* **2**, 174–179 (1988).
26. C. Sachse et al., *Am. J. Hum. Genet.* **60**, 284–295 (1997).
27. N. Hanioka et al., *Am. J. Hum. Genet.* **47**, 994–1001 (1990).
28. A.C. Gough et al., *Nature* **347**, 773–776 (1990).
29. R. Saxena et al., *Hum. Molec. Genet.* **3**, 923–926 (1994).
30. A. Gaedigk et al., *Am. J. Hum. Genet.* **48**, 943–950 (1991).
31. M.M. Shi et al., *J. Clin. Pathol: Mol. Pathol.* **52**, 295–299 (1999).
32. M.M. Shi et al., *Res. Commun. Mol. Pathol. Pharmacol.* **103**, 3–15 (1999).
33. M. Chee et al., *Science* **274**, 610–614 (1996).
34. M.M. Shi, M.R. Bleavins, and F.A. de la Iglesia, *Mol. Diagnosis* **4**, 343–351 (1999).
35. V. Glaser, *Genetic Engineer. News* **18**, 1998.
36. F. Richard et al., *Lancet* **349**, 539 (1997).
37. E.J. Calabrese, *Ecogenetics*, J. Wiley, New York, 1983.
38. K.T. Kelsey et al., *Am. J. Indus. Med.* **31**, 274–279 (1997).
39. S. Kato, *Cancer Res.* **52**, 6712–6715 (1992).
40. A. Hivonen et al., *Carcinogenesis* **14**, 85–88 (1993).
41. J.Y. Hong and C.S. Yang, *Envrion. Health Prospect.* **105**, 759–762 (1997).
42. D.W. Nebert, *Am. J. Hum. Genet.* **60**, 265–271 (1997).
43. W.W. Grody, *Diag. Mol. Pathol.* **4**, 155–157 (1995).
44. E.W. Clayton, *J. Am. Med. Assoc.* **274**, 1786–1792 (1995).
45. ASHG Report, *Am. J. Hum. Genet.* **59**, 471–474 (1996).
46. J. Stephenson, *J. Am. Med. Assoc.* **279**, 184 (1998).
47. G.J. Annas, *J. Am. Med. Assoc.* **270**, 2346–2350 (1993).
48. L. Gostin, *Ann. Intern. Med.* **127**, 683–690 (1997).
49. M.J. Mehlman et al., *Am. J. Hum Genet.* **58**, 393–397 (1996).
50. W.C. McKinnon et al., *J. Am. Med. Assoc.* **278**, 1217–1219 (1997).
51. Public Law 100-578, Clinical Laboratory Improvement Amendments of 1988, 42 USC 263a et seq.
52. N.A. Holtzman et al., *Science* **278**, 602–605 (1997).
53. NIH–DOE Task Force on Genetic Testing, Available at: *www.med.jhu.edu/tfgtelsi*
54. S. Gevers, *Med. Law* **15**, 407–411 (1996).
55. M. Cronin, in J. Schlegel, ed., *Pharmacogenetics: Bridging the Gap Between Basic Sciences and Clinical Application*, IBC, Southborough, MA, 1996, pp. 5.1.1–5.1.18.
56. H.J. Jick, *Allergy Clin. Immunol.* **74**, 555–557 (1984).
57. H.C. Slavkin, *JADA* **128**, 1157–1160 (1997).
58. L.M. Fisher, *N.Y. Times*, February 25 (1998).
59. E. Lai et al., *Genomics* **54**, 31–38 (1998).
60. D.G. Wang et al., *Science* **280**, 1077–1082 (1998).
61. M. Cargill et al., *Nat. Genet.* **22**, 231–238 (1999).

PROFESSIONAL POWER AND THE CULTURAL MEANINGS OF BIOTECHNOLOGY

Robert M. Nelson
University of Pennsylvania
Philadelphia, Pennsylvania

Paul Brodwin
University of Wisconsin
Milwaukee, Wisconsin

OUTLINE

INTRODUCTION

Contemporary bioethics arose in the 1960s in the wake of innovations in dialysis, transplants, artificial organs, and assisted reproduction. These biotechnologies sparked debates about the allocation of scarce resources and the quality and limits of life. In response, ethicists developed a set of abstract normative principles — autonomy, beneficence, and distributive justice — which structure professional debates to this day. These core principles of American bioethics take a generic concept of the person and make it the basis for a universal morality. This is, of course, the autonomous individual of Western liberalism: the sovereign individual who acts freely according to a self-chosen plan. However, the very technologies

that sparked early bioethics unsettle this tacit understanding of the person. They have created new ways of exerting one's will and gauging one's present identity and future fate. For example, predictive genetic testing alters the way people calculate their life prospects — the likely mixture of happiness and suffering they will encounter — and it can erase or intensify certain aspects of their identity. New strategies to assess people's subjective experience are demanded by transplantation and mechanical ventilation. Such technologies set in motion profound transformations in our cultural model of personhood and in the ways we experience and enact moral agency. These transformations, as much as the conflict between abstract principles, motivate our deepest ethical concerns.

This article takes up a vexing tendency in our current use of biotechnologies: the replacement of moral discourse by technical expertise. This tendency is magnified when health care professionals must make educated guesses about the subjectivity of technologically altered individuals. If we regard biotechnology as a simple collection of devices — a morally neutral means to the ultimate good of prolonging life — certain procedures become a standard, unquestionable component of care (1). This, in turn, justifies technical discourse as the sole guide for treatment decisions. It provides certainty for medical workers, but it also rules out other ways to understand the experience of, for example, chronically ventilated children and our ethical obligations toward them. This article examines the conflict of interpretation over the subjective experience of the mechanically ventilated child. Too often, the authority to read it one way or the other remains solely in the hands of medical professionals, illustrating how rational, technical expertise can foreclose genuine moral debate.

PROLONGED VENTILATION FOR CHILDREN WITH NEUROMUSCULAR DISEASE

An infant or young child with a neuromuscular disorder, such as nemaline rod myopathy or spinal muscular atrophy, usually presents with generalized poor muscle tone, which then progresses to respiratory failure and the need for assisted ventilation. Often a muscle biopsy is necessary to make a diagnosis, and the first biopsy may be inconclusive as it takes time for the characteristic pathological findings to develop. Nemaline rod myopathy, selected as an illustration of the issues raised in this entry, is a rare and slowly progressive neuromuscular disease that renders a person immobile (2,3). Unable to move and unable to breath, such an infant may undergo a tracheostomy procedure so that the airway is secure and the infant can be ventilated more easily. A tracheostomy, which is a plastic tube inserted into the windpipe through a surgical incision in the front of the neck, is done when the plan is to provide long-term ventilation, perhaps including sending the patient home on a ventilator. In addition a so-called gastrostomy tube may be placed into the infant's stomach through the abdominal wall so that the infant can be fed without placing a temporary feeding tube through the nose. Once these procedures have been performed, the infant is often transferred from the intensive care unit to a "step-down" unit designed for long-term care and is placed on a simpler breathing machine whose primary purpose is for use in home ventilation.

We do not know what causes nemaline rod myopathy, nor do we have any treatment for it other than putting someone on a ventilator. It only affects the skeletal muscle, so the other muscles of the body such as the heart and gut work fine. In its most severe form, the child cannot move, cannot swallow, cannot breathe, and may not be able to move his eyes or close his eyelids. Eyes open and barely moving, an affected infant or child will stare at you without expression, unable to move the muscles of the face to show pain or pleasure, unable to smile or frown, unable to laugh or cry, unable to communicate at all — the face a frozen, expressionless mask. The disease, however, does not affect the brain. An affected infant or child is alert and aware of everything going on around him.

At times, conflict may arise over the continued use of mechanical ventilation, with a parent insisting on the right to remove the ventilator and let the child die, and the professional staff (both nurses and physicians) resisting or disputing this claim. Usually this conflict is addressed in such terms as the child's quality of life, the "value" of living with a severe disability, the child's "best interest," the authority of the physician/state in determining "medical neglect," the parent's authority to make decisions concerning the child's medical care, and so forth. All of these approaches assume that the technology itself is morally neutral: It sets the stage for ethical conflict but does not influence the outcome. In this article we question that assumption. We suggest that once biotechnology is introduced into patient care, it constrains our subsequent moral choices, and this belies the claim that the technology itself is morally neutral. We examine, in particular, tracheostomy and mechanical ventilation, techniques that undercut parental (or nonprofessional) control over the medical care of children suffering from a neuromuscular disease such as nemaline rod myopathy. Because medical professionals control both how to use these technologies and how to interpret their effects upon individual subjectivity, ethical conflicts become prematurely translated into matters of technical expertise.

To examine these issues in the case of a ventilator dependent child, we need to understand how the technology affects the child, how the ventilator defines and redefines the boundary of body, and how the ventilator produces a subject who is both body and machine. Rather than a neutral technology designed to achieve goals that are selected for non-technical reasons, the ventilator seems to impose its own agenda and values. However, we should not reify this technology in a way that obscures and thus privileges the agency of the medical professionals who control it — as in the Wizard of Oz when we are told "don't pay attention to that man behind the curtain." Moreover, we need to pay attention to the organizational and cultural context within which professionals operate. Only then can we understand how professional power infiltrates both the use of technology and the interpretation of its effects.

MEANING OF TEARS

An infant with nemaline rod myopathy often has tears in his eyes, lending support to the belief (usually by a parent or family member) that this constant tearing indicates emotional and physical distress. The professional staff, however, may interpret these tears as a simple result of an inability to close the eyelids, rather than a reminder of an infant's suffering. Such an infant becomes the locus of a contested interpretation. The parents view the infant as too fragile and unable to tolerate activities such as being propped up in a wheelchair or taken out of the hospital on field trips. The apparent suffering of the infant often motivates their desire to stop the ventilator at some point in the future. The staff regards the infant as able to take pleasure in simple things and to sit contentedly in a chair for hours. On occasion, the staff may blame a parent for the very episodes that the parent interprets as fragility, arguing that the infant would cry when his parent arrived and "suffocated" him during a brief visit. This conjures up the image of an overprotective parent who fails to appreciate the strength and ability of her or his child; however, the use of the metaphor of suffocation takes on a more literal and provocative meaning given the staff's suspicion of the parent's (implicit or explicit) desire to stop the ventilator.

Fundamentally the infant's parent continues to see the ventilator as something other than the infant: as a threat, as an invasion of his body, as something foreign. Removing it in order not to prolong the infant's suffering simply returns the infant to a more natural state. In the eyes of the staff, the threat to the infant is not the ventilator but the parent. The staff often regards such an infant not as a body on a machine, but as both body and machine — that is, a machine-human/ human-machine cyborg (cybernetic organism) whose body and machine components are mutually interdependent (4). Therefore, to ask the staff to participate in turning off the ventilator is to ask them to amputate part of the infant's body. Is there a fact of the matter that could settle this dispute about removing the ventilator? Is there a third point of view, acceptable to both parties, from which the relationship between tears and suffering could be "objectively" determined? Glossing the difference between pain and suffering, the absence of other signs of pain such as sweating and a rapid heart rate may support the staff's argument. This is an argument that a parent can only lose; the best one could hope for is for everyone to agree to disagree.

There often is no apparent disagreement over the moral principle that mechanical ventilation can be stopped if the burden of such treatment outweighs the potential benefit (5). What is in dispute is the description of the child as either suffering or simply unable to close his eyelids. The staff is able to substitute a physiologic argument about the presence of pain for the existential question of whether the child is suffering — in effect, shifting a moral argument about the worth of living life on a ventilator to a technical dispute about the interpretation of a physical sign. In addition the staff has the political power to threaten a parent with a charge of failing to provide necessary medical treatment, that is, "medical neglect," thus throwing the matter into court. Consequently the moral and political questions about what to do are transformed into a rational technical discourse that in the minds of the staff is unambiguous. As such, the dependence of knowledge on subject position or "point of view" is implicitly denied; the power to determine objectivity is invisibly exercised. Rather than moral and political discourse about the conflicting visions of the infant's experience serving as the "paradigm of rational discourse," the professional technical discourse determines the political stakes and reduces moral discourse to the vanishing point (6, p. 194).

KNOWLEDGE AND COMMUNICATION

An appeal by the parent to the shared "sense experience" of the child's tears usually will not persuade the staff that the child is suffering. In addition the staff will probably fail to convince a parent that the child's tears simply mean that he cannot close his eyelids. The staff may additionally dismiss the parent's claim to know that the child is suffering as mere subjective opinion. The staff's insistence that the child is not suffering clearly reinforces their professional interest in continuing treatment.

In answer to the question — "What protects knowledge from being [either] the arbitrary expression of subjective desires [on the part of a parent] or the tool of social and personal interests [on the part of the medical and nursing staff]?" — Helen Longino, a philosopher of science at the University of Minnesota, offers an approach she refers to as "contextual empiricism." Longino, as do other philosophers in the pragmatic tradition, grounds "objectivity" or the truth of a statement concerning a sense experience in an "inter-subjective" or "social" process that should ensure "the inclusion of all socially relevant perspectives in the community engaged in the critical construction of knowledge" (7, pp. 200, 202–203). A necessary part of this communicative process is a critical examination of the implicit assumptions that establish the relevance and interpretation of observational or empirical data. The natural world cannot impose one single interpretation, that is, the empirical observation of the presence of tears does not establish the truth of one or the other interpretation. However, differences of power in this social or communicative process may limit the plurality of interpretations to the one that is consistent with the dominant discourse.

This is the outcome, for example, when a particular powerful group or individual constrains the freedom of expression and diversity of legitimate knowledge, or restricts the community of discourse in such a way as to predetermine which interpretation is accepted. This process involved (fragment) discounting the parent's interpretation of a physical sign such as an infant's tears. Parents may also be isolated from outside family and community supports and effectively alone with the medical staff during conversations in the hospital about the care of their child. In this setting medical professionals fail to establish a meaningful community of inquiry concerning the question of a child's suffering.

In discussing contemporary policy debates about technology in general, Langdon Winner observes that this lack of a coherent community of discourse "contributes to two distinctive features ... (*1*) futile rituals of expert advice and (*2*) interminable disagreements about which choices are morally justified" (8, p. 75). The moral uncertainty involved in the application of ventilator technology to the indefinite support of patients cannot be resolved by an appeal to the technical advice and expertise of the physician. Such a "futile" appeal to expert advice will not achieve a consensus. Moreover the lack of an appropriate community of discourse and the resulting disagreement over what choices are morally justified privileges the physician's technical expertise and thus interpretation of the patient's experience. While this does not avoid a conflict of interpretations, it guarantees that the conflict gets resolved in ways that favor the power and interests of the physician.

KNOWLEDGE AND POWER

In response to this professional prerogative which constrains the available choices, how do we empower a parent to make decisions concerning her child's medical care? The notion of personal autonomy or self-rule has resulted in a significant shift of power from the physician to the patient over the past two decades. However, once we abandoned the concept of the child as property, the notion of parental autonomy as a justification for the right of a parent to direct a child's medical care became problematic. Each one of us may have an absolute right to determine our own medical care. A parent has, at most, a prima facie right that is limited by the child's right to life and freedom from serious bodily injury or disability (9). Within this constraint we expect that a parent will make decisions that benefit the child or, in other words, are in the child's "best interest." Thus the parent's vision of the good is imposed on or becomes the child's vision — an imposition we accept given the diversity and, at times, incompatibility of competing visions of the good within our society. This creates a disturbing paradox. On the one hand, we expect a parent to express a decision concerning his or her child not as "what is good for the parent" but rather as "what is good for the child." On the other hand, the only possible way to give "voice to the voiceless" is by articulating adult values and projecting them upon the child (10). An infant with nemaline rod myopathy cannot speak; so when we speak for such an infant, we ask: "If I (the adult) were in this condition, what I would want?"

If we seek to escape this paradox and avoid this imposition of adult values by supporting the child until he is capable of self-expression, we inadvertently reinforce the physician's tendency for the relentless application of life-sustaining technology. Consequently the concept of a child's "best interest" appears to be the arena for an unavoidable expression of adult power on the part of either the physician or the parent. The stakes are high, for if the parent understands the child's "best interest" in such a way as to refuse what the physician otherwise believes to be necessary medical care, the parent may find him- or herself in court facing a charge of medical neglect.

The past two decades have seen a lively debate in the bioethics literature and the courts concerning the withholding and withdrawal of life-sustaining technology (11). Some have argued that removing a person from a ventilator is to choose death based on the judgment that the anticipated quality of life is not worth living. Others, concerned about the potential abuse of quality-of-life judgments, have argued that such decisions are better understood as the choice of how to live while dying (12) or as simply the decision to remove technology that is no longer medically indicated (11,13). The first argument, that of how to live while dying, makes the decision to remove a ventilator dependent on a prior determination that the patient is dying — a determination that the technology itself makes more difficult. A child with nemaline rod myopathy who is on a ventilator may not die for years in the absence of an intervening complication. Thus, once you put him on the ventilator, you cannot remove it unless he is dying, and he is not dying unless you remove the ventilator. The second argument, that technology can be removed when it is no longer medically indicated, either restricts the removal of technology to those situations where more narrow technical goals cannot be achieved or obscures the physician's own determination of an acceptable quality of life behind the veil of professional technical competence. A ventilator is medically indicated when a patient has respiratory failure; it is not indicated when either the patient recovers or the ventilator fails to correct the respiratory failure. Thus, in most cases, the ventilator for an infant with nemaline rod myopathy is medically indicated. If a physician argues that the ventilator is not medically indicated, since correcting the patient's respiratory failure does not contribute to the overall good of the patient, we necessarily must engage the question of what is or is not in the patient's "best interest" — a discussion that cannot avoid questions of the patient's quality of life. The problem then of trying to avoid an explicit discussion of a child's anticipated quality of life is that the physician's power and authority is inadvertently reinforced.

Physicians impose their power by establishing what counts as legitimate and credible knowledge, rather than by forcing a choice for one of either two credible options. In asking whether a child on a ventilator is suffering, a parent and the health care team may disagree over the description of the child's life, not over the moral evaluation of an agreed upon description (14). It is simply not credible to the medical staff that the child is suffering. In discussing the problem of technology as ideology, Robert Pippin asks whether we have "been so influenced by technical instruments ... that our basic sense of the natural world has changed ... so fundamentally that ... possibilities for social existence are seen only ... in terms of such technical imperatives." The physician's reliance on technology "reaches a point where what ought to be understood as contingent, an option among others, open to political discussion, is instead falsely understood as necessary; what serves particular interests is seen, without reflection, as of universal interest; what is a contingent, historical experience is regarded as natural" (15, p. 46). Physicians appear to have lost any sense of the natural or the contingent as a moral category. Rather the natural

serves to mark that domain that resists the physician's intervention, as in "let nature take its course." The natural becomes that which cannot be technically overcome, rather than that which should not be overcome. The natural is subservient to the technical, which in turn resists the explicit introduction of moral and political questions.

DIFFERENCE BETWEEN STARTING AND STOPPING

The belief that technology is a neutral means to whatever ends are selected on moral, political, or more narrow physiologic grounds is a fundamental conviction and ideology of medical practice. For example, the decision to perform a tracheostomy may not be intended as a decision for long-term home ventilation but may be seen as consistent with a desire to defer any decision to limit or withdraw support given any remaining uncertainty about a child's diagnosis and prognosis. A parent may be told that a decision to perform a tracheostomy does not preclude a decision at some point in the future to remove a child from the ventilator—"what is done can be undone." Such a statement is consistent with the widely endorsed bioethical teaching that there is no significant moral or legal difference between withholding and withdrawing treatment (11).

There are a number of important assumptions behind the use of this bioethical maxim. First, it assumes a symmetry in the application and removal of medical technology consistent with the prejudice that technological means are value-neutral. It also assumes a symmetry between an endotracheal tube and a tracheostomy by reducing each to its essential function of establishing an airway for the purpose of mechanical ventilation. However, as an endotracheal tube is inserted either through the mouth or nose, the tape required to hold it in place covers a major portion of the face. A tracheostomy surgically inserted through the front of the neck results in the entire face being visible and thus capable of expression. Second, the maxim appears to ignore any relevant differences that may occur between the moments of application and removal of the technology, apart from any changes in the medical indications. Third, and related to this historically naive stance, is the view that the organizational context in which these decisions are being made is apparently unimportant. After a tracheostomy a child may be transferred out of the intensive care unit and to the ward that houses patients in the home ventilation program. One wonders whether the use of this bioethical maxim that there exists no significant moral or legal difference between withholding and withdrawing treatment is based on a reasoned ethical stance, or used as a rhetorical device to postpone the discussion of more difficult ethical issues to a later date. The latter interpretation is confirmed, for example, when a receiving physician in conflict with a parent over the removal of a ventilator is unable to find any other physician willing to assume the child's ongoing medical care (and thus agree to withdraw support), including any of the physicians who have previously cared for the child prior to the tracheostomy. An appreciation of the value-laden nature of a tracheostomy, along with the importance of time and context, counsel against a premature surgical procedure and then transfer to a home ventilation program.

IS TECHNOLOGY VALUE-NEUTRAL?

The bias that our medical technology is simply a "collection of devices" emphasizes the functional aspects of technology and obscures its social context. As a result of this dichotomy between function and context, our technology appears value-neutral, while only the application of that technology becomes morally problematic. Andrew Feenberg, in an article reflecting on the relationship between technology and power, points out that this "dichotomy of goal [function] and meaning [context] is a [contingent] product of functionalist professional culture" rather than a necessary component of technology (16, p. 9). Echoing criticisms of the common view of science as value-free, Feenberg asserts that this ideology of technology as value-neutral reinforces the dominant forms of power that compose the cultural horizon and social meaning of technology. Feenberg refers to this as the "bias of technology" by which "apparently neutral, functional rationality is enlisted in support of a hegemony," that is, the professional dominance of those who control the technology. The professional claims to exclusive control of technology are strengthened insofar as its associated values and presuppositions drop out of sight (16, p. 12).

This professional control is also reinforced by the perception that technology should always be used when it can be used: the so-called technological imperative. Barbara Koenig suggests that the technological imperative acquires a certain moral force as the technology becomes habitual or routine (1). Her field research focused on therapeutic plasma exchange, a procedure that involves the removal and then replacement of blood plasma. Koenig identified three steps in the process by which plasma exchange became a routine therapy for certain conditions. The first step was a transformation in roles and responsibilities. The physician–nurse relationship shifted from egalitarian to hierarchical while, at the same time, the physicians moved from being closely involved to delegating many of the routine tasks to nurses (1). Similarly the physicians involved in a home ventilation program maintain close control while shifting many of the routine tasks from in-hospital nurses and respiratory therapists to parents and visiting home nurses. The second step was the use of treatment rituals that appeared to reduce uncertainty, anxiety, and disorder and thus established the meaning of the technology as standard therapy for both patients and staff (1). When a patient is placed on a home ventilator after a tracheostomy, there is an orderly and nearly invariable process of parental training, arranging for nursing services, equipment purchase, and so forth, that must take place prior to discharge from the hospital. Any deviation from this process generally results in uncertainty, inefficiency, omissions, and the like. The third step that Koenig identified was the generation of research data. Noting the enthusiasm with which the physicians engaged in plasma exchange collected clinical data as part of their ongoing research, she speculated that the machine's capability

of producing research data supported the physician's tendency to use the technology (1). Although this may be true with therapeutic plasma exchange, it does not appear that the development of home ventilation programs was driven by a research imperative. The physician's use of home ventilator technology is more likely driven by such factors as the need to find alternative placements for children who otherwise would survive intensive care but remain dependent on ventilator technology. Also the immediate efficacy of the ventilator when compared to plasma exchange is obvious, for otherwise the child would die. Despite these differences Koenig's conclusion remains essentially correct. The technological imperative is transformed into a moral imperative through the development of a "sense of social certainty experienced by health professionals" (1, pp. 485–486). The technology simply begins to feel routine, and hence both appropriate and necessary.

The decision to perform a tracheostomy and then to transfer a child to a unit where the use of chronic home ventilation is considered routine is governed by a similar moral imperative. The unit is organized so that home ventilator technology is accepted as standard therapy. Within this social context, it becomes difficult, if not impossible, to question whether this technical standard of care ought to be used for any particular child. The moral question of what is in a particular child's "best interest" thus receives an axiomatic answer applicable to all children: "Given these circumstances, we should provide the standard technology." The moral meaning of our medical technology is thus created and sustained by the professional culture of the hospital. Since the statutory definition of medical neglect in Wisconsin, for example, is simply failure to provide necessary medical care, the technological and moral imperative experienced by medical and nursing professionals clearly has "the potential to wrest control of decisions about the use of technology" from parents and patients (1, p. 489).

Andrew Feenberg proposes that one of the assumptions behind our modern image of technology is that social institutions must adapt to the technological imperative. Noting that "the economic significance of technical change often pales besides its wider human implications in framing a way of life," Feenberg encourages us to study the "social role of the technical object and the lifestyles it makes possible" through defining "major portions of the social environment, such as ... medical activities and expectations" (16, p. 9, 16). This assumption that we must adapt to technology is readily apparent over the past two decades with the development of home care programs for so-called technology-dependent children (17). The family is explicitly expected to change in response to the demands of caring for a child who is to be discharged from the hospital on a home ventilator. The only other available option is foster care, which is problematic for two reasons. While the child is in foster care, a parent may lose control over any decisions to either withhold additional medical treatment or withdraw existing medical treatment. In addition there is often an unspoken assumption that to choose foster care reflects poorly on the ability of a parent to provide for his or her child. Although many parents choose to take their ventilator-dependent child home out of a sincere concern for their continued life and well-being, the normative pressures against choosing otherwise are enormous once the child is within the context of the home ventilation program. This assumption then that social institutions such as the family must adapt to the technological imperative is another manifestation of the extension of professional power implicit in the ideology of value-free technology.

The apparent inevitability of the technological imperative is rejected by both Koenig and Feenberg. Consistent with Koenig's thesis, Feenberg asserts that "technology is just another dependent social variable" and the "scene of social struggle" (16, p. 8). Contrary to the claim that technology itself requires professional control, Feenberg argues that technology has been used to block the extension of public or democratic control to "technically mediated domains of social life" (16, p. 20). Thus the professional medical culture seeks to reinforce the image of technology as both value-neutral and complex in order to maintain control despite the "routinization" process of placing that same technology into the home.

CONTESTING POWER OVER TECHNOLOGY

To insist that technology is "socially constructed" may give the impression that people ultimately hold complete power over its meanings and uses. The typical circumstance of one group having more control and another group having less control over technology results, in this view, solely from social and political considerations independent of the constraints of particular devices. To deny that technology is "socially constructed" may imply the opposite extreme: that we have no power over technology and that our moral and cultural response is determined by its concrete and independent reality. We argue against both extremes. To assume that technology is neutral — the core of the anticonstructionist position — reinforces the professional dominance of physicians. By failing to recognize the extent to which technical knowledge is constructed by and for the interests of a particular community, we are likely to ratify this group's power and authority. At the same time the strong constructionist position ignores the material effects of this technology and the way it constrains moral deliberation. By use of this technology, a child's breathing becomes, ineradicably and by definition, assisted breathing. As a result the object of clinical decision making has become altered. It is no longer the child as such but the hybrid object of the ventilator/child. Once this massive technological intervention has taken place, it is not clear who gives life to whom: the ventilator to the child, or the child to the ventilator? In the face of this ambiguity, medical workers substitute technical rationales for action (algorithms and expertise about ventilator use) for the search for mutual understanding about the child's experience and, possibly, suffering. In the end this process makes humans subservient to things (18). However, this result is produced by both the social power of the profession and the particular way this technology transforms the very objects of clinical decision-making.

Cleaving to either the strong constructionist or the strong anticonstructionist position misses this complex result.

Controlling the technical mediation of social activities such as medical care is a major source of public power within our society. The ability to manage or expand this technical mediation results in the concentration of power in an elite group of experts, the narrowing of acceptable options for public discussion, and an increase in the extent of administrative or professional control over aspects of daily life (15). Changes in the way medical technology is delivered or applied to a particular problem will require a shift in this expert control of technology. As Feenberg writes: "If authoritarian social hierarchy is truly a contingent dimension of technical progress, . . . and not a technical necessity, then there must be an alternative way of rationalizing society that democratizes rather then centralizes control" (16, p. 5). Is the link between the physician's social role and the control of medical technology necessary or contingent? For example, one approach to the issue of physician-assisted suicide is to allow for assisted suicide while preserving the traditional social role of the physician by making available to the general public the technical tools that to-date remain under the physician's prescriptive authority.

If we move the control of medical technology into the public domain, we will need to create an appropriate community of discourse to monitor development and application. Such a task may be difficult given the diversity of our current communities. Although the reform of technology is a better option than simply resistance, it is not clear that those (e.g., nurses and physicians) who have been socialized in the modern medical ethos could resist attempting to impose new forms of professional control (15). The creation of a community for the reform of medical technology should include those who anticipate needing or who may resist medical technology and thus will require abandoning the notion of professional expertise. In addition such a community of discourse must begin by questioning the assumption that technology is a value-free instrument — an assumption that serves to reinforce professional control and hinder rational debate. Or should we simply recognize the legitimate existence of disparate communities and thus reframe the question of the appropriate application of medical technology as a choice of which community to belong to?

BOUNDARY OF BODY

Let us now return to the question of whether the medical and nursing staff simply see a mechanically ventilated child with a tracheostomy differently than he is seen by his parent(s). As the disease progresses, a child with nemaline rod myopathy cannot move, cannot breath, cannot express emotion, indeed cannot make any facial expressions; communication at best may occur through the movement of an eye in response to a question. Consequently it may be difficult if not impossible to get any indication of what a physical sign such as "tearing" meant to the child. As the passive object of our application of ventilator technology, the child is reduced to either a resource for our instrumentalist projects or a mask for our dominant interest in maintaining control (6). Modern medical technology, as we have seen, clearly includes the feature of the technical control of some human beings by others. Donna Haraway attributes this modern tendency towards technical domination to the dualism between objective nature and subjective culture so that the projects or interests that shape our determination of natural objects are hidden from view. As an alternative, she offers us a view of "objectivity as positioned rationality" (6). To capture a notion of the object as active and not passive, Haraway asserts that "bodies as objects of knowledge . . . materialize in social interaction. Boundaries are drawn by mapping practices; 'objects' do not pre-exist as such" (6, p. 200–201). The issue then is the various positions from which each one of us, including the child's parents, view the ventilated child — a question that necessarily draws us back into an explicit discussion of the power of professional "mapping practices" in determining the boundaries of the ventilated child as the object of our attention.

How then are we to understand who the ventilated child is, this body attached to a ventilator? Through an exploration of the "semiotic use of the body" among the Kayapo of the Brazilian Amazon, Terence Turner illustrates how "the body is at once a material object and a living and acting organism possessing rudimentary forms of subjectivity that becomes, through a process of social appropriation, both a social identity and a cultural subject" (19, p. 145). For example, the Kayapo use various modifications of their body surface to define and redefine their social identity, as in the use of ear piercing to indicate age cohort, marital status, and other social identities. The individual Kayapo, as both a social body and an embodied subject, assumes the dual role as product and producer (19). In our case the body of the ventilated child as a material object of our technical interventions takes on the social identity of a patient in the home ventilation program. Although his parent(s) may try to resist this medical appropriation of the child's body, the tracheostomy and attached ventilator tubing are key modifications of his body that produce the child's social identity as a patient in the home ventilation program (19). The ventilator-infant as embodied subject appears to be the socially patterned product of our technical activity, rather than the producer of its own activity. Similar to the ideological consequences of the view of technology as value-neutral, the misrepresentation of the "cultural subject" of the ventilator-infant as an "objective (natural) feature existing independently" of our social production further reinforces the dominant power of the physician (19). In infancy, it is unclear that there is any content to the notion of the subject existing prior to and independently of the social production of the embodied subject by others. In other words, what meaning can we give to the notion of the "best interest" of the child apart from the specific interests of a particular embodied social subject? Once a child undergoes a tracheostomy and is placed on a chronic home ventilator, he is and will remain a patient in the home ventilation program. This much is visibly announced on his body. Thus we come full circle to the notion of the ventilator-infant as cyborg, the machine-human as "embodied subjectivity" rather than the machine as external to the body. The

social identity of the ventilator-infant/infant-ventilator is a product of being a machine-human hybrid, that is, the ventilator gives life to the body and the body gives life to the ventilator. To contemplate taking the patient off of the ventilator would be to contemplate amputation — a request that the medical and nursing staff cannot and will not honor.

CONCLUDING REMARKS

What have we learned from this story of the social production of the ventilator-infant as a patient in a home ventilation program? We have come to doubt the universality of the classic teaching of the symmetry between withholding and withdrawing technology. We have a renewed understanding of the insight that our medical technology is not value-neutral, and that it often serves to reinforce the professional dominance of physicians. While we acknowledge that specific hospital units have different cultures, the general impact of the organizational context on the ability of patients and parents to control the application of medical technology is greater, on reflection, than previously appreciated. This impact occurs not primarily through the imposition of a different set of moral values, but through fundamental shifts in our point of view, and thus how we see and come to know our patients. The fundamental conflict between medical staff and the parents of young, chronically ventilated children does not turn on a choice between competing ethical principles. As we have seen, medical staff often do not advance any explicit ethical principle in support of their action (i.e., refusing to withdraw the ventilator). The conflict turns rather on what counts as the proper object of concern: the child, the ventilated child, or the hybrid "ventilator-child." In this conflict, medical workers enjoy enormous power to make authoritative readings of the child's subjective experience and, more generally, what is admitted as knowledge in the medical setting.

Using the example of ventilator technology, this article demonstrates how the use of biotechnology constrains our subsequent moral choices concerning the application of that technology in a manner that belies the claim that the technology itself is morally neutral. Physicians and other medical staff are thus "technicians" in the following sense: They translate moral and political issues surrounding the application of biotechnology into the dominant technical discourse. To accomplish this, these technicians may constrain the freedom of expression and diversity of legitimate knowledge, or they may structure the community of discourse in such a way as to reinforce their own power and interests. Whatever strategy they ultimately follow, the ideology of technology as value-neutral reinforces their dominance in the clinical encounter and their authority to establish the local meanings of technology. The belief that biotechnology is a value-neutral means to certain ends selected according to entirely different criteria (moral, political, or physiological) thus perpetuates professional dominance over patients and their families. Moreover this belief obscures that the process is happening at all. Clinical actions, such as withdrawing or continuing ventilation, as well as knowledge claims about the child at the center of attention, are always underdetermined by available physiological evidence. Ethics and politics, even if hidden, play the crucial role in the outcome of conflicts between medical staff and parents. Elucidating that role, and restoring moral discourse where it has been banished, demands that we abandon the model of morally neutral biotechnology.

ACKNOWLEDGMENT

Portions of this article have been previously published in R.M. Nelson, The Ventilator-Infant as Cyborg: A Case Study in Technology and Medical Ethics, and P. Brodwin, Introduction, in P. Brodwin, ed., *Biotechnology, Culture and Body*, Indiana University Press, Bloomington, IN, 2000.

BIBLIOGRAPHY

1. B.A. Koenig, in M. Lock and D. Gordon, eds., *Biomedicine Examined*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988, pp. 466–490.
2. S.T. Iannacone and T. Guilfoile, *J. Child Neurol.* **3**, 30–32 (1988).
3. M.B. Connolly, E.H. Roland, and A. Hill, *Pediatric Neurol.* **8**(4), 285–288 (1992).
4. G.L. Downey, J. Dumit, and S. Williams, *Cult. Anthropol.* **10**(2), 264–269 (1995).
5. L.J. Nelson and R.M. Nelson, *Crit. Care Med.* **20**(3), 427–433 (1992).
6. D.J. Haraway, in *Simians, Cyborgs, and Women: The Reinvention of Nature*, Routledge, New York, 1991, pp. 193–201.
7. H.E. Longino, in A. Feenberg and A. Hannay, eds., *Technology and the Politics of Knowledge*, Indiana University Press, Bloomington, IN, 1995, pp. 200–203.
8. L. Winner, in A. Feenberg and A. Hannay, eds., *Technology and the Politics of Knowledge*, Indiana University Press, Bloomington, IN, 1995, pp. 75–77.
9. Committee on Bioethics, American Academy of Pediatrics, *Pediatrics* **99**(2), 279–281 (1997).
10. D. Archard, *Children: Rights and Childhood*, Routledge, New York, 1993.
11. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Deciding to Forego Life-Sustaining Treatment: A Report on the Ethical, Medical, and Legal Issues in Treatment Decisions*, U.S. Government Printing Office, Washington, DC, 1983, pp. 60–90.
12. A.J. Dyck, in S.J. Reiser, A.J. Dyck, and William J. Curran, eds., *Ethics in Medicine: Historical Perspectives and Contemporary Concerns*, MIT Press, Cambridge, MA, 1977, pp. 529–535.
13. P. Ramsey, *Ethics at the Edges of Life: Medical and Legal Intersections*, Yale University Press, New Haven, CT, 1978, pp. 145–188.
14. S. Hauerwas and D.B. Burrell, in S. Hauerwas, *Truthfulness and Tragedy: Further Investigations in Christian Ethics*, University of Notre Dame, Press, Notre Dame, IN, 1977, pp. 15–39.

15. R.B. Pippin, in A. Feenberg and A. Hannay, eds., *Technology and the Politics of Knowledge*, Indiana University Press, Bloomington, IN, 1995, pp. 43–51.
16. A. Feenberg, in A. Feenberg and A. Hannay, eds., *Technology and the Politics of Knowledge*, Indiana University Press, Bloomington, IN, 1995, pp. 5–20.
17. J. Arras, ed., *Bringing the Hospital Home: Ethical and Social Implications of High-Tech Home Care*, Johns Hopkins University Press, Baltimore, MD, 1995.
18. S. Vogel, in A. Feenberg and A. Hannay, eds., *Technology and the Politics of Knowledge*, Indiana University Press, Bloomington, IN, 1995, pp. 23–37.
19. T. Turner, *Cult. Anthropol.* **10**(2), 145–169 (1995).

See also GENETIC DETERMINISM, GENETIC REDUCTIONISM, AND GENETIC ESSENTIALISM.

PUBLIC PERCEPTIONS: SURVEYS OF ATTITUDES TOWARD BIOTECHNOLOGY

DOROTHY C. WERTZ
The Shriver Center
Waltham, Massachusetts

OUTLINE

INTRODUCTION

Surveys can provide a useful baseline for policy planning, educational development, marketing, and ethical discussions. Surveys may include the general public or may seek to identify the views of various stakeholders, such as researchers, physicians, genetic counselors, or patients. Surveys also have limitations. This article will present the rationale for doing surveys, outline major types of survey methods, and describe some results of major national and international surveys of public and stakeholder views on biotechnology, including genetic testing, screening in the workplace, gene therapy (including germ-line gene therapy), and enhancement of average human characteristics.

RATIONALE FOR CONDUCTING SURVEYS

Uses of Surveys

Surveys can be useful in providing a baseline for public policy debates or ethical discussions, devising educational programs, marketing genetic services, and identifying existing or potential abuses of genetics. In the public policy arena, for example, it makes little sense to outlaw a test or procedure that most people would either use themselves or think others should have a right to use. Policies related to abortion for fetal "defects" are one example. The General Social Surveys (GSS) conducted by the University of Chicago's National Opinion Research Center on a random sample of the United States adult public have indicated that about 80 percent believe a woman should be able to get an abortion if there is a "serious defect" in the fetus. This percentage has varied by only a few points between 1971 (two years before *Roe v. Wade*, the Supreme Court decision legalizing abortion) and 1998 (1). Reports such as this can be useful to lawmakers, courts, advocacy organizations, and lobbyists. Surveys have the advantage of allowing a wide variety of people to express their views, including many who would not otherwise come forward to comment on proposed policies, regulations, or ethical guidelines.

In the arena of education, surveys can identify areas of ignorance that may require special educational programs. For example, a 1986–87 survey of 1473 physicians conducted by the U.S. Congress Office of Technology Assessment (OTA) found that 63 percent would reject a sperm donor with a family history of Huntington's disease (an autosomal dominant disorder with severe effects on the nervous system, transmitted to 50 percent of the offspring and not presymptomatically diagnosable in the donor at the time), while 61 and 49 percent respectively would reject a healthy donor with a family history of Duchenne muscular dystrophy or hemophilia A, both of which are transmissible only by females (2). The survey results indicated need for better education of physicians, greater oversight by sperm banks, and new guidelines from professional societies. A 1995 survey of 499 U.S. primary care physicians found that substantial percents held inaccurate knowledge about the life expectancy and functioning of people with common genetic disorders such as cystic fibrosis or Down syndrome (3).

Surveys also inform ethical discussion. A 1985 survey of 683 geneticists reported that 62 percent in the United States would either perform prenatal diagnosis at parental request solely to select the sex of the fetus (34 percent) or would offer a referral (28 percent) (4). Subsequent discussions among bioethicists and medical practitioners led to various international ethics guidelines rejecting this practice, notably those of the World Health Organization (WHO) (5). These discussions appear to have had little effect on attitudes of professionals in the United States; a 1995 survey of 1085 genetics professionals found 72 percent willing to do prenatal diagnosis for sex selection or refer (5).

Market researchers have always used surveys to identify markets for new tests and treatments. Many of these researchers work for nonprofit institutions or health departments, trying to predict demand for new tests or treatments. Surveys of potential uptake of presymptomatic tests for Huntington's disease (7) and carrier status for cystic fibrosis (8) were conducted before such tests become available, and results (indicating that most people at risk would take such tests) affected allocation of government and institutional funds. More recently, studies of willingness to take tests for mutations in BRCA1 and BRCA2 genes (partially predictive of breast cancer) have led to commercial outlays of plant and equipment (9). More general studies of public attitudes toward genetic testing in general and gene therapy (10,11) have enabled researchers, companies, and policy makers to identify trends toward increased public acceptance of genetic technologies, at least for some purposes. Surveys can also identify the extent of possible abuses of new technologies, such as "genetic discrimination" by insurers or employers (12–14), the testing of asymptomatic children for genetic mutations for adult-onset disorders (15), or perceptions of minority populations that they are being used as guinea pigs (16).

In sum, surveys of the public or stakeholders, including patients, professionals, or others (e.g., religious organizations) can help to prevent making policy decisions or ethical guidelines in a vacuum. Surveys provide a window on the opinions of those who will ultimately benefit or lose from research or policy. Surveys also identify differences among the various stakeholders. The public may think differently from the members of bioethics commissions; patients may (and do) think differently from the majority group (18); women may think differently from men (19). Surveys provide a means — albeit imperfect — for gauging some of these differences.

Limitations of Surveys

Most surveys involve hypothetical questions ("What would you do if. . .?"). In genetics, many surveys involve tests, treatments, or products not yet available. The reliability and validity of responses to hypothetical questions are always debatable. Reliability means that the respondent would answer the question the same way if given the survey instrument a second time. In other words, the survey measures the attitudes or intentions it was designed to measure. In real life it is almost impossible to go back to the same respondents and re-ask the questions, especially if a survey was anonymous. In ethics, particularly, the reliability of surveys is always in doubt. The more difficult the ethical, social, or emotional implications of a question, the more likely it is that respondents themselves will report that they might answer the same question differently another time. One possible response to the reliability problem is to present analogous situations (perhaps in different case vignettes) in different portions of a questionnaire and to measure whether there is a meaningful cluster of responses pointing to a possibly stable underlying attitude (one statistical procedure used is called *factor analysis*). Another response is to conduct longitudinal surveys on the same population (e.g., the public) and to look for wide swings of opinion in the short term. Underlying opinions on most ethical issues do not change rapidly unless there is some well-publicized new discovery or other precipitating event, so a major change in 6 to 12 months suggests that the initial survey may have been unreliable.

The problem of validity (predictive value) is perhaps more important. Validity means that a respondent will actually do what he or she says, if presented with the "real" situation. One limitation of surveys is that the questions posed usually cannot describe accurately the complex characteristics of a hypothetical genetic test or predict whether it will actually provide benefits once it becomes available. Time and again research has predicted an active uptake of new genetic tests, yet when such tests actually became available, few used them (9). There are two major reasons for this lower-than-predicted uptake: (*1*) the uncertainty inherent in many of the tests, and (*2*) the absence of proven or acceptable prevention or treatment (20). For example, BRCA mutation testing provides only a "risk" of breast cancer (about 80 percent for women with a family history of BRCA-related cancer, somewhat lower for women without such a history), not a certainty. "Treatment" involves removal of breasts and ovaries and even then does not provide absolute certainty of freedom from the disease. For Huntington's disease there is no prevention or treatment. For autosomal recessive disorders, such as sickle cell anemia or cystic fibrosis, carrier testing may lead to difficult reproductive decisions, including decisions about prenatal diagnosis and abortion, but does not lead to treatment. People who say in surveys that they will take future genetic tests perhaps implicitly believe that treatment will become available at the same time as testing (11). When no treatment materializes, they do not take the tests. Although the validity of surveys may be improved by detailed explanations of what is and is not possible, lengthy explanations tend to reduce response rates and are difficult to apply to technologies still under development. Future technologies often turn out differently than presented in surveys. For example, no one could have predicted that there would be over 800 different mutations for cystic fibrosis or that a few individuals could have two mutations without the classical symptoms (21). The most controversial future technologies — germ-line gene therapy, genetic enhancement, and human cloning — will

undoubtedly present as-yet-unforeseen scientific and technical possibilities posing new ethical problems.

SURVEY METHODS

Sampling

Most large-scale surveys use one of two methods to select samples: random probability sampling or quota sampling (22). Random probability sampling selects a proportion of the population to be sampled at random, to avoid researcher bias, and requires a comprehensive, accurate list of the elements to be sampled (census tracts, blocks, house numbers in the case of door-to-door or mail samples, telephone numbers in the case of random digit dialing). Usually a table of random numbers is used to select the first element in the survey, and interviewers count down the list and select elements at regular intervals, depending on the proportion of the population to be sampled. (For a 1 in 10 sample, every tenth telephone or house number is selected.) Interviewers continue to make visits or calls to the numbers selected until they contact someone. In budgeting for surveys, researchers usually allow for at least three or four visits or calls per selected household in order to obtain a response.

Stratified sampling is a modification of random probability sampling. It is used when researchers are especially interested in comparing several groups, for example, African-American and white women with and without a family history of breast cancer, in regard to their attitudes toward BRCA testing. In a stratified sample, the population would be organized into four groups (African-American women without a family history, African-American women with a family history, etc.) and an appropriate number selected at random from each group. Stratification ensures that each group of interest is included but limits the numbers in each group, thereby reducing the statistical "power" of finding real differences between groups, especially if these differences are small. Power always depends on the smallest group in a comparison. Usually this group must include at least 200 in order to provide an 80 percent likelihood of identifying a difference of 10 percent between groups.

Quota sampling differs from probability sampling. Quota sampling begins with a matrix based on relevant characteristics of the population to be sampled. Researchers first need to know the overall characteristics of the population, for example, sex, age, race, geographical area, income, or education. The population is divided into cells (e.g., white women aged 25–30, living in southern suburban areas, of Catholic background, with a college education and income of $30,000–40,000), and each cell is given a weight according to its percentage in the overall population. Interviewers seek to fill an assigned "quota" of respondents for each cell. They do not attempt repeated visits or calls to randomly preassigned addresses or telephone numbers.

The random probability sample is the "gold standard" of survey research, and is the method preferred by U.S. government agencies and used by the University of Chicago National Opinion Research Center's General Social Surveys (1), the U.S. Congress Office of Technology Assessment (10), and a 1995 survey of primary care physicians (3). The method is expensive, however, since a sampling frame (list of "units" — persons, house numbers, telephone numbers — to be sampled) must be assembled and interviewers have to make repeated attempts at contact. Quota sampling usually costs less than one-quarter the price of random probability sampling and can be completed much more quickly. Most political campaign polls use quota sampling. Some survey firms combine probability and quota sampling. Households are selected randomly, but individuals within each household are selected by quota. This is called *modified probability sampling*. At least one public survey of ethical views in genetics has also used modified probability sampling (23).

Other sampling methods include so-called convenience sampling (e.g., stopping people on a street corner or giving out questionnaires in a college cafeteria) and "snowball sampling" (letting the initial interviewees direct the researcher to other possible interviewees, who in turn suggest further interviewees). For precise statistical reporting of population views, neither method is adequate, though college students frequently use both. These methods can, however, be useful in anthropological studies of communities and in identifying questions for further large-scale research. The snowball method, by providing access to a cross section of willing interviewees sharing particular social characteristics, has enabled researchers to sample in depth the emotional lives of particular groups, for example working-class women (24). The convenience sample is sometimes the only sample readily available, and the researcher faces the choice of using this or not doing the research at all. Scholarly journals have published results of surveys using these methods, providing that authors acknowledge their limitations.

Survey Approaches: Interview, Telephone, Mailed Questionnaire

The various approaches are approximately equal in terms of reaching study populations. Most people now have telephones (although in some of the poorest rural or innercity areas, about 5 percent do not). People with unlisted telephones can be included in surveys by dialing from random number tables instead of phone books. People appear willing to answer phone questionnaires up to about 15 minutes in length (about the same length as in-person interviews) though some surveys have lasted up to half an hour. Mailed questionnaires have the advantage of allowing for greater length and complexity, but they require more tasks from the respondents and must be accompanied by stamped return envelopes. Sometimes written questionnaires are delivered by in-person interviewers, who wait while the respondent completes them. Many people feel more comfortable with methods that promise anonymity, such as mailed questionnaires or randomly dialed calls, especially if questions touch on deeply held personal beliefs (25). In-person and telephone methods now provide similar response rates. Questionnaires may produce lower response rates, but they can convey more complex questions and can use case vignettes.

Requirements for Consent

Participation in anonymous surveys does not require written informed consent of the type usually required in medical research. Anonymous studies usually fall under the educational exemption of the Federal Office for Protection from Research Risks (OPRR) rules, though researchers funded by the federal government or at institutions that receive federal funds must request this exemption from their Institutional Review Board. Answering the survey constitutes consent. Political pollsters and market researchers usually give a very brief description of the survey ("I'm going to ask you a few questions about. . ., which will take about . . . minutes") and ask if the person is willing to participate. Questionnaires usually have a brief introduction describing the purpose of the survey, naming the individual researchers and institution conducting the research and stating that participation is voluntary. If sensitive questions are involved, the introduction gives a warning to this effect (sometimes repeated at intervals in the questionnaire) and tells people that they do not have to answer questions that make them feel uncomfortable. (Similar warnings can be given in voice interviews.) The most sensitive question of all — even in ethics questionnaires — is usually about income. Many people are reluctant to disclose how much money they make. Repeating the word "optional" before this question is advisable.

Anonymity can be preserved even in "before-and-after" surveys by placing matching numbers on sets of questionnaires. In studies where respondents receive payments after completing sets of questionnaires, anonymity can still be preserved by preventing names from reaching researchers and by destroying code numbers once payment is made. Surveys where researchers can match individuals' names to questionnaires or interviews require a signed informed consent document, provided that the survey is sponsored by an agency of the U.S. government or conducted at an institution receiving federal funds. Most medical and educational institutions receive federal funds.

Developing Survey Instruments

Focus Groups. Focus groups are frequently used to help develop questionnaires or interviews. A focus group is somewhat like a group interview focused on a certain topic (26). Focus groups usually consist of 6 to 12 people, often selected by research companies specializing in focus groups. There is no attempt at random selection. Focus groups are not surveys. Ideally participants will be as different from each other as possible. Over an approximately two-hour period, a facilitator gently leads the group through a list of areas that the organizers wish covered and also coaxes out the quieter members of the group. Comments by one member may reverberate in others, leading to a fuller range of responses than would occur in a one-on-one interview. The group is tape-recorded and transcribed. Researchers continue to conduct new focus groups until new viewpoints cease to arise. Focus groups are not "town meetings" of people with special agendas, nor are they neighborhood gatherings. Participants are selected so as not to know others in the group, and usually do not include the most vocal activists. The transcripts help identify concerns and wordings that should be included in the survey instruments. Focus groups are also widely used by market researchers (to develop advertising) and politicians (to develop campaign strategies).

Wording. In developing the survey instruments, researchers will need to consult with some of the major interest groups, whether or not they have used focus groups. For example, in developing questions on relationships between biotechnology and people with disabilities, persons from groups in the disability community should be consulted. It is important to try to avoid language that some people may regard as insulting, such as "burden," "affected child," "defect," or phrases that put diseases ahead of persons, such as "Tay-Sachs baby" or Down syndrome case." The disability community prefers "person-first" language, such as "persons with mental retardation."

Choice of words can have strong effects on responses. For example, most people do not want to be "genetically engineered" or to eat food that is "genetically engineered." People are more receptive to "biotechnology" than to "genetic engineering" (27). Technical terms such as "germ-line gene therapy" are confusing to most people. An explanation, such as "correcting genes that would carry the disease to future generations" (11) may be better. Even "enhancement" is a questionable word for inclusion in a public survey. It is usually better to give an example, such as "increase athletic abilities" or "improve performance in school."

Sometimes researchers skew questions toward a particular type of response that they hope to find. For example, a researcher who hopes that people will agree to participate in biomedical research may say "Do you agree or disagree with the following statements: A. The benefits of the proposed research outweigh the risks. B. This study may lead to an important treatment. I would want to participate in this research myself." Both statements emphasize the benefits and ignore the risks of the research. The vast majority of people will agree with them. In most cases the surveyor constructs skewed questions inadvertently, rather than by design. Even bioethicists construct skewed questions. The usual methodological approach to counteract this is to construct an equal number of questions that may elicit the opposite response. For every question that may produce a positive response to biomedical research, a question phrased in a negative light (stressing risks or uncertainties) should be included. A balanced set of questions not only contributes to internal reliability of responses but is also useful in the data analysis.

Ideally surveys should be short. Exploring complex issues can require more lengthy instruments, however. Projects should not be rejected on the ground that "the public will never understand the issues." Researchers in developing nations have found that most people, including those with no formal education, understand complex concepts (including risk and the placebo effect) if these are explained adequately (28). Simplistic global questions,

such as "Do you think the Human Genome Project poses greater risks than benefits?" tend to produce answers of limited usefulness. Questions about self-perceived knowledge may measure self-confidence rather than actual knowledge, unless accompanied by actual knowledge tests.

Cross-Cultural Surveys. Survey research can be conducted crossculturally (29). Biotechnology will affect populations differently, but human beings face similar issues. For federally funded surveys requiring translation, two steps are ordinarily required. First, the survey is translated and tested on several members of the population who will receive it. Second, the translation is "back translated" into English by an independent translator who has never seen the original and compared with the original English version so that mistakes and nuances of language can be corrected.

Pilot Testing. All survey instruments are ordinarily field-tested on several people, revised for clarity, and then pilot-tested on a larger group. Observers watch people completing questionnaires or interviews and then debrief them regarding comprehensiveness, emphasis, and clarity. Usually an instrument undergoes several rounds of pilot testing.

Response Rates

Some authorities consider 50 percent an acceptable response rate (22). Some journals, such as the *American Journal of Public Health*, have preferred 60 percent but are willing to publish surveys with lower response rates if they include an especially hard-to-reach group, such as physicians. A 70 percent response rate is good by anybody's standards (22). In-person quota surveys usually have very high (over 95 percent) response rates. Questionnaire surveys usually have low response rates (about 30 percent) on the initial mailing, but these are increased by successive mailings and telephone reminders. Most responses to the initial mailing arrive within two weeks. At that time, a second mailing, sometimes in the form of a postcard reminder, is sent, followed by a third and final mailing (sometimes a complete questionnaire in case the first one has been lost) in a week or two. Usually successive mailings produce about half the number of responses received in the previous mailing.

Physicians are a notoriously difficult group to survey and are difficult to reach even for a telephone reminder. Some survey firms specialize in physicians, charging approximately $100 per completed interview. A small payment of $25 for completing a questionnaire or interview can increase response rates dramatically, even for physicians.

Including Members of Minority Groups

Members of minority groups may hold considerably different views from those of the majority with regard to the ethical conduct of medicine and research, in view of past experiences where they were used as guinea pigs without their knowledge (31,32). Unfortunately, most survey research on ethical, legal, and social issues in genetics, like most medical research (33), does not adequately reflect the views of minorities. Two possible reasons are (*1*) underrepresentation of minorities in some groups, notably service providers, and (*2*) extra costs of including adequate numbers of minorities in public and patient surveys, costs that may exceed a funding agency's customary limits.

According to the American Society of Human Genetics, of 4810 members based in the United States, 4031 (84 percent) are white, 604 (13 percent) are Asian, 57 (1 percent) are African-American, 34 (0.7 percent) are Hispanic, 30 (0.6 percent) are Native American, and 54 (1 percent) are "other." These percentages apply to the entire membership, which includes researchers as well as service providers and also includes many people who are not board certified. An examination of names (where sex can be attributed) suggests that about half of board-certified members in the United States may be women. The National Society of Genetic Counselors (the professional association for Master's level counselors) reports that 96 percent of its members are women, 93 percent are white, 4 percent are Asian, 1 percent are Hispanic, 1 percent are African-American, and 1 percent are "other" (34). Women are better represented in genetics than in some other medical specialties; according to American Medical Association data, 44 percent of pediatricians, 29 percent of obstetricians, and 23 percent of family practitioners are women (35). According to U.S. Census Bureau data, 21 percent of all physicians were women in 1990 (36). Some minorities, however, are underrepresented in genetics when compared with other medical specialties and with their proportions of the U.S. population. According to 1990 U.S. Census data on 587,675 physicians, 80 percent were non-Hispanic white, 11 percent were Asian, 5 percent were Hispanic, 4 percent were African-American, and 0.1 percent were Native American (37). In the 1990 census, African-Americans constituted 12 percent of the U.S. population, and Hispanics constituted 9 percent. Today Hispanics are almost 12 percent of the population, excluding Puerto Rico (38). These groups are underrepresented in medicine generally, but especially underrepresented in genetics. There are no national data on race, ethnicity, or sex of individuals or families receiving genetics services.

Surveys of minority groups may cost several times the fee for surveys of the general population. For example, because African-Americans represent only 12 percent of the adult U.S. population, it is necessary to contact successfully and to screen more than 4000 households in order to locate 500 that include an eligible African-American respondent for a door-to-door or telephone survey. Surveying Hispanic populations requires translation and use of bilingual interviewers. Spanish-speaking groups originating from different areas (e.g., Puerto Rico, Mexico, Cuba, Peru) hold different views and their responses cannot be lumped together under the general label "Hispanic" or "Latino." Each group must be surveyed separately. Reaching members of "low-incidence" groups such as Southeast Asians or Native Americans requires construction of special sampling frames, something that most

survey research organizations are unable or unwilling to do.

It is not sufficient simply to conduct a general population survey and hope that minorities will be adequately represented. According to Roper Starch Worldwide, a survey research firm, the average English-speaking population sample achieved is 80 percent white, 10 percent African American, and 10 percent other or unknown. Telephone surveys may result in inclusion of even fewer African-Americans than door-to-door surveys, since fewer households have phones. In order to represent minority views, each group must be "oversampled." In survey research language, "oversampling" means that to get an adequate sample of a minority group, which represents its proportion of the population, it is necessary to solicit participation from proportionately more members of the minority group then are represented in the population as a whole. The low incidence and difficulty of reaching minority groups may quadruple the cost of a survey. For example, a survey of 1000 members of the U.S. public may cost around $25,000, while a comparable survey of only 500 African-Americans may cost a minimum of $50,000. A survey of Native Americans or Asians will probably be unobtainable from most survey research firms. Some researchers consider a survey of 500 persons inadequate to represent the views of an entire minority. The sampling error is plus or minus four percentage points, and many persons will never be reached, either door-to-door or by telephone. To represent all of America's various minorities in adequate numbers for analysis could be economically prohibitive by the usual funding standards for survey research.

Data Analysis and Presentation of Survey Results

Survey data are usually coded (put into numeric form) and entered into statistical programs in computers. The most common statistical packages are SAS (Statistical Analysis System) and SPSS (Statistical Package for the Social Sciences). Qualitative data (people's "write-in comments" on questionnaires or verbal responses to open-ended interviews that ask them to reply in their own words) can be quantified for purposes of statistical analysis. This requires development of categories to organize the responses. Each piece of numerical data is entered by two independent operators, whose work is compared for corrections, a procedure known as "punch and verify."

Survey results are usually expressed as one figure. However, if another randomly selected group of similar size were to be sampled, another figure might emerge. The results of separate surveys would fall on a bell curve, with 95 percent of these results falling within a certain range. Thus, if 80 percent of 1000 people surveyed give answer A, a statistician would say that the true figure is in the range of 77 to 83 percent with 95 percent confidence. In other words, we are likely to be right that the true value falls in this range 95 percent of the time, but there is a 1 in 20 chance that the answer is outside this range. The range is sometimes reported in tables as a "confidence interval"; the 95 percent is the "confidence level" for this interval.

VIEWS ON BIOTECHNOLOGY

Major Surveys

The most careful and comprehensive surveys were conducted for the U.S. Congress Office of Technology Assessment, which is no longer extant. These include a mail survey of the 500 largest U.S. Corporations and 11 labor unions, reported in *The Role of Genetic Testing in the Prevention of Occupational Disease* (1983); a survey of surrogate mother matching services, reported in *Infertility: The Medical and Social Choices* (1988) (40); a survey of 1575 physicians, 1213 fertility specialists, and 30 commercial sperm banks, reported in *Artificial Insemination: Practice in the United States* (1988) (2), and, most important, a 1000-member telephone Survey of Public Attitudes toward Biotechnology, Science, and Engineering, conducted in 1986 and reported in *Public Perceptions of Biotechnology* (1987) (10). The National Center for Human Genome Resources (NCGR), a private, nonprofit research organization in Santa Fe, New Mexico, commissioned a partial repeat of the OTA Survey of Public Attitudes in 1996, by the same survey researchers, in order to examine possible trends (11). In addition, the NCGR study surveyed 521 primary care physicians, 100 leaders of patient organizations, 102 research and development directors from the biotechnology industry, 76 genetic researchers, 50 science journalists, 50 religious leaders, 70 medical directors of insurance companies, and 79 federal and state policy makers. In the United States, Singer (41) has conducted public surveys on the ethics of utilizing prenatal diagnosis, including complex issues such as sex selection, and has examined the views of labor unions and religious groups.

In Canada, the Royal Commission on New Reproductive Technologies surveyed over 9000 people by mail questionnaire and telephone, regarding opinions about use of fetal tissue and treatments for infertility (42).

In Europe, the Eurobarometer surveys in 1991, 1993, and 1996, commissioned by the European Union (EU) and conducted by INRA, a European network of market and public opinion research agencies, are the most comprehensive (27). In the most recent, Eurobarometer 46.1 (1996) public surveys were conducted in all 15 member states of the EU, plus Norway and Switzerland, using random probability sampling, for a total of 16,246 face-to-face interviews about public perceptions of risks and benefits of various aspects of biotechnology.

A series of large-scale surveys of portions of the Japanese public by Macer also looks at global attitudes, but response rates were about 24 percent, well below standards of acceptability in the West, although Macer claims it is average for Japan.

Two worldwide surveys of genetics professionals by Wertz and Fletcher, a 19-nation survey of 683 in 1985 (4) and a 36-nation survey of 2901 in 1995 (44) together with surveys of 499 primary care physicians, 476 genetics patients, and 988 members of the general public in the United States, concentrate largely on ethical issues in genetics services, rather than research, but some questions touch on new technologies.

Overall Views on Biotechnology

In the United States, the public's self-perceptions of adequacy of knowledge increased between 1986 and 1996 (11). The percent who thought they knew what a gene is increased from 85 to 91 percent; those who thought they understood the meaning of "human gene therapy" increased from 29 to 49 percent. In 1996, 68 percent knew that scientists were trying to map the human genome, 48 percent thought this effort would have at least a moderate effect on themselves or their families, and 53 percent knew that genetic tests were available. About 7 in 10 approved of gene mapping. Among the various leadership groups, 87 percent approved mapping the human genome, and 92 percent expected improvement in the early diagnosis of disease in the next ten years; 88 percent of patient group leaders, 78 percent of industry representatives, and 68 percent of scientists thought there would be at least moderate improvement in treatment of chronic diseases, though notably fewer scientists (13 percent) thought there would be "a lot" of improvement, as compared to 42 percent of patient organizations and 39 percent of policy makers. A majority of all leaders (62 percent) said they followed the scientific literature fairly closely, but only 17 percent followed it very closely, including 17 percent of patient organizations, 8 percent of policy makers, 7 percent of insurers, and 0 percent of the media. All groups thought that funding would present the biggest frustration for scientists in biotechnology, though 27 percent of the biotech industry also cited government regulations. Overall, few leaders saw research confidentiality (3 percent), patient confidentiality (2 percent), or discrimination (1 percent) as the biggest source of frustration. Four out of five (81 percent) in the leadership sample expected current efforts to identify markers of genetic disease to have at least a moderate effect on society. Nine out of 10 (92 percent) expected society to benefit from medical applications of biotechnology in the next 10 years, including 51 percent who expected "a lot" of benefit. Notably fewer religious leaders (20 percent) expected a lot of benefit than patient organizations, scientists, and industry (65–67 percent). About two-thirds (65 percent) thought that medical applications of biotechnology would pose some risk to society, ranging from 86 percent of religious leaders and 80 percent of policy makers to 49 percent of scientists and 53 percent of industry. Only 15 percent thought there would be "a lot" of risk. Most (85 percent) thought the benefits of biotechnology would outweigh the risks, while 8 percent thought risks would outweigh benefits. Leadership groups were divided with regard to government regulation. About 40 percent would like to see regulations left as they are now, including 57 percent of scientists, 47 percent of industry, and 27 percent of insurers; 27 percent would like more stringent regulation, including 44 percent of religious leaders, 37 percent of policy makers, 30 percent of insurers, and 17 percent of scientists and industry; 18 percent would like less stringent regulation, including 32 percent of industry but only 5 percent of policy makers.

In the 1996 Eurobarometer survey (27), 50 percent thought that biotechnology would improve life, with Finland the most optimistic and Greece, Norway, and Germany the least optimistic; 15 percent thought it would make things worse, with Austria, Norway, and the Netherlands the most pessimistic. When questions used "genetic engineering" instead of biotechnology, optimism fell to 39 percent and pessimism rose to 27 percent. Although 54 percent thought biotechnology would lead to cures for most genetic diseases in the next 20 years, 70 percent thought it would create dangerous new diseases. Overall optimism about telecommunications, computers, solar energy, and new materials exceeded optimism about biotechnology. Optimism about biotechnology decreased slightly, but significantly, between 1991 and 1996, from 51 to 48 percent. About half of respondents had heard about and talked with someone about biotechnology in the three months preceding the survey. "Textbook knowledge" did not increase significantly between 1993 and 1996; there was wide variation among countries, and generally greater pessimism in countries with lower levels of textbook knowledge, such as Austria. There was great national variation in beliefs about heritability of human characteristics. In all, 62 percent, ranging from 78 percent in Ireland and 77 percent in the Netherlands to 39 percent in France, thought that human intelligence was mainly inherited; 17 percent, ranging from 32 percent in Ireland to 9 percent in Switzerland, believed that attitudes toward work were mainly inherited; 15 percent, ranging from 24 percent in Italy, and 23 percent in Austria, to 10 percent in Denmark and 9 percent in Sweden, thought criminality was mainly inherited; 25 percent, ranging from 39 percent in the Netherlands and 34 percent in Germany to 13 percent in France and 17 percent in Portugal, thought homosexual tendencies were mainly inherited. There are no simple cultural explanations for these findings.

Europeans thought that international organizations such as the WHO (35 percent), or scientific organizations (22 percent) were better able to regulate biotechnology than national governments (17 percent) or the EU (6 percent) (27). Only 24 percent thought that existing regulations were sufficient. Respondents placed the greatest confidence in consumer organizations to tell the truth about biotechnology, followed by schools and universities. Public authorities were way down the list. Over half of respondents (54 percent) thought that "irrespective of regulations, biotechnologists will do what they like" especially in Denmark (71 percent), Switzerland (65 percent), France, Germany, and the United Kingdom (all 60 percent). The Japanese surveys, which were also distributed in New Zealand for comparison, used the words "genetic engineering," "genetic manipulation," and "genetically modified organisms," and the framing of questions makes comparison with European and U.S. surveys difficult. Nevertheless, they suggest a considerable amount of fear about and concern over new developments in biotechnology, as well as need for more education (43).

Views on Genetic Testing

In 1996, 53 percent of the United States public said that they were aware that genetic tests were available,

including 42 percent of high school graduates and 68 percent of college graduates (11). More than 9 out of 10 (93 percent) approved of the use of genetic information for early diagnosis of disease; 85 percent approved of presymptomatic genetic testing for diseases that occur later in life, including 43 percent who approved strongly. About three-quarters (73 percent) thought it was "a good thing for a healthy person to be able to find out how likely they are to get a serious disease in the future," while 17 percent thought it was "bad" to know. More women (21 percent) than men (13 percent) thought it was bad to know. Those who thought knowledge was good cited the possibility of prevention (57 percent), changing lifestyles (11 percent), finding a cure (10 percent), making decisions about having children (7 percent), or early treatment (6 percent), but 23 percent simply said they wanted to know in advance. Those who thought foreknowledge was bad cited worry (34 percent), negative influence on life plans (27 percent), depression (11 percent), fatalism (no need to know) (13 percent), negative self-image, or emotional effects (12 percent). Perspectives on why knowledge is good or bad may explain one of the most persistent problems in survey research on genetics: why the majority of people say that they would take a test and then do not take it once it becomes available. People who think foreknowledge is good implicitly associate the knowledge gained through testing with prevention or eventual cure. If a test becomes available without acceptable means of prevention or cure, most people are not interested (20).

Nine out of 10 (88 percent) approved the use of genetic tests to find out whether future children are likely to have a serious genetic disease. The 1996 NCGR survey did not specify what kind of testing this meant (prenatal diagnostic or parental carrier testing). A 1995 survey of 476 genetics patients (44) (mostly working-class Catholic white women bringing children in for evaluation) found that 64 percent agreed that "before marriage, responsible people should find out whether they could pass on serious diseases or disabilities to their children" (64 percent also agreed in a 1994 public survey); 81 percent thought that "a woman should have tests on the unborn baby if she is at risk of having a child with a serious disease or disability" (62 percent agreed in the public survey); and 80 percent thought that "tests on unborn babies should be available to all women who request them." Fewer patients (21 percent) thought that "a woman should have an abortion if tests say the unborn baby has a serious disease or disability"; 43 percent were neutral on this question.

In Europe, 81 percent thought it was "useful for society to use genetic testing to detect diseases that we might have inherited from our parents," and 72 percent thought that people should be encouraged to take such tests (11). In Japan, most thought prenatal testing (76 percent) and presymptomatic testing (73 percent) should be available under national health insurance (43).

The level of certainty provided by a test might be expected to affect public acceptance of it (20). This was not the case in the 1996 NCGR survey. Most people were willing to accept some uncertainty (11). Only 27 percent said tests should be made available only if they predicted with certainty that someone would develop a disease. Most of the rest were willing to settle for tests that could indicate "a high risk." ("High" was not defined in the survey.) Treatability of the condition had less effect than might be expected on the degree of risk that people would tolerate. For serious, untreatable conditions 27 percent thought tests should be available even if they indicate only a slightly increased risk, 20 percent would require a moderately increased risk, and 40 percent thought tests should be available only if they indicate highly increased risk. Slightly higher percents would approve tests for slightly (36 percent) or moderately (26 percent) increased risk if a condition were treatable. Most (94 percent) thought doctors should "advise" people with a family history of cancer to take a genetic test for cancer, and 48 percent thought doctors should advise everybody to take such a test.

Reports of personal willingness to take genetic tests are of dubious validity, tending to exaggerate greatly the numbers who would take such tests. In 1996, 65 percent in the NCGR survey (11) said they would take a test indicating whether they would develop a fatal disease later in life, almost the same percent as in 1986. There was a statistically significant decline between 1986 and 1996 from (83 to 76 percent) in the number who said they would take a genetic test, before having children, that would indicate whether their children would inherit a fatal genetic disease. Nevertheless, the percent who would be tested on behalf of future children exceeds those who would be tested on their own behalf. A review of recent similar studies of willingness to be tested indicated high percentages of acceptance of BRCA testing, as in earlier surveys, but little evidence that people were actually requesting tests (45).

In Japan, 63 percent of the public, when asked the same question, said they would have tests before conceiving children; 53 percent would have presymptomatic tests for themselves, and 76 percent would have prenatal diagnosis (43). Percents were not substantially different among students, general public, university staff, and scientists.

Prenatal Testing

A survey of labor unions and Protestant religious groups found that most either supported genetic testing and prenatal diagnosis or (in the case of unions) had no position (41). A survey of public attitudes toward using prenatal diagnosis to select the sex of the child found that only 5 percent approved this use; however, when presented with the "hard-luck case" of a couple with four children of the same sex, a substantial minority (38 percent) supported sex selection (41). In a 1995 survey, 72 percent of U.S. genetics professionals, 68 percent of U.S. primary care physicians, 59 percent of patients, and 38 percent of the general public thought the doctor should do prenatal diagnosis in this case (percentages for professionals include those who would offer a referral) (44). This points to another potential weakness of surveys: People may answer one way on a general question and another way in response to a more concrete situation. In medical practice the exigencies of the concrete situation

usually win. This is why some ethics surveys employ case vignettes.

Surveys of women having prenatal diagnosis for genetic disorders generally document increased anxiety, need for sympathetic and accurate counseling, reduction in anxiety after receiving favorable results, and willingness to undergo the procedure in a future pregnancy (46). For women whose tests indicate presence of a genetic condition, studies document the difficulty of decisions (46). Women who already have a child with a genetic condition are often reluctant to abort a fetus with the same condition (47). Most women who choose abortion recover psychologically within months, but a small minority continue to have strongly negative emotions (48). In general, Americans appear reluctant to use selective abortion. In a 1994 public survey describing eight fetal conditions, the majority would not abort for any condition listed, including severe mental retardation, with the child "unable to speak or understand" combined with death in the first few months of life; 48 percent would abort in this situation, 47 percent would abort for severe mental retardation accompanied by a nearly normal lifespan, and 41 percent would abort for paralysis from the neck down, with no retardation and a normal lifespan. Smaller percents would abort for mild retardation ("child could live independently") (17 percent), severe incurable mental disease appearing at age 40 (21 percent), moderate retardation ("could communicate but not live independently") (22 percent), "gross overweight" (16 percent), and "child not of the sex desired" (7 percent). For all of these conditions, except sex selection, majorities believed that abortion should be legal for others.

Testing Children

The genetic testing of asymptomatic children for disorders that occur later in life has occasioned much discussion. A survey of U.S. laboratories found that most had no comprehensive policies and many had performed such tests (15). Other surveys found that geneticists in English-speaking nations and Northern/Western Europe rejected the practice, but majorities of those in other parts of the world would test children for mutations for Huntington's disease or Alzheimer's disease at parents' request (49).

Perspectives on Confidentiality

Access to an individual's genetic information by spouses and genetically related family members is one of the most controversial areas in bioethics. Although surveys of genetics professionals indicate that most believe that spouses should not have access to information without the individual's consent (44), surveys suggest patients and the public think otherwise. In one set of surveys, 20 percent of the public thought spouses should have general access without consent; among patients, substantial minorities thought spouses should have access to their partner's genetic information without consent if there were a risk that a child could inherit mental retardation (43 percent), if the spouse/partner had mutations predisposing to mental illness (31 percent), or alcoholism (23 percent), or if the spouse/partner was a carrier of cystic fibrosis (30 percent). In the NCGR Survey, 75 percent thought spouses should have access, but the question did not specify whether this was with or without consent (11). There was far less support (9 to 12 percent) in any group for access for blood relatives, at least in general questions. When the situation was presented as a case vignette describing Huntington's disease, however, 75 percent of patients and 38 percent of U.S. genetics professionals thought the doctor should tell the relatives, against the patient's wishes (50). These results point to divisions within the professional community and differences among professionals, patients, and the public. Beliefs about privacy are different outside English-speaking nations and Northern/Western Europe, where the unit of privacy is the individual. In most of the world, the unit of privacy is the family (50–52). Surveys point to an overwhelming consensus against access for employers and insurers without a person's consent; many respondents believed that these institutions should have no access at all, even with consent.

Eugenics

An international survey of geneticists found little support anywhere for state-mandated testing or sterilization (53). However, substantial percents of geneticists in developing nations (especially China and India) and Eastern Europe believed that "reducing the number of deleterious genes in the population" was "an important goal of genetics." Except in the English-speaking world, majorities would offer purposely pessimistically slanted information after prenatal diagnosis, so that people would abort without the professional suggesting it directly. Overall, genetics professionals held a pessimistic view about disability (53).

Gene Therapy

The NCGR survey suggests a trend toward public acceptance of gene therapy. In 1996, 87 percent approved of "correcting genes that cause serious illness" (11). The percentage who believed that "changing the genetic makeup of human cells is morally wrong" decreased from 42 percent in 1986 to 22 percent in 1996. In 1986, 83 percent approved of "changing the makeup of human cells to cure a usually fatal genetic disease"; in 1996, 85 percent approved (not a statistically significant increase). There was a significant increase (77 to 84 percent) in those who approved changing human cells "to reduce the risk of developing a fatal disease later in life." There has been a significant decline in approval of gene therapy for enhancement ("to improve the physical characteristics children would inherit.") In 1986, 44 percent approved this use; in 1996, 35 percent approved this use. Approval was highest among those with less than a high school degree (61 percent), lowest among college graduates (28 percent). Among leadership groups, 25 percent of patient organizations but only 12 percent of scientists approved this use.

The majority of the public approved both somatic cell and germ-line gene therapy. In 1996, 68 percent thought doctors should be allowed to correct both the gene affecting the disease in the patient, *and* the gene

that would carry the disease to future generations, an increase from 62 percent who approved in 1986. Four of five (83 percent) primary care physicians thought that doctors should be allowed to correct genes carrying disease to future generations. Majorities in all leadership groups, including 79 percent of patient organizations, 74 percent of religious leaders, 79 percent of insurers, 72 percent of policy makers, 66 percent of industry, and 55 percent of scientists (who may be more aware of the risks) approved of both somatic cell and germ-line gene therapy.

Genetics in the Workplace

The OTA survey of industry suggested that most companies were not using genetic testing, though some had used it in the past (mainly for sickle cell trait, which is irrelevant to occupationally-related disease) and some expected to use it in the future (39). One study found that while employers supported genetic testing to identify presumably susceptible workers and move them to less hazardous jobs before damage occurs, unions preferred genetic monitoring, which means testing to see whether cellular or molecular damage is actually occurring (54). The unions' argument was that genetic testing could have low predictive value for disease and would lead to unfair discrimination against many workers. In a 1995 international survey there was agreement among genetics professionals everywhere that testing should be voluntary, the worker should have access to the results, and no one else should have access without the worker's consent (44). Half of U.S. patients, however, thought that testing should be required, apparently because they thought testing would protect the worker. In the United States, the Americans with Disabilities Act (ADA) now prevents employers from refusing to hire because of family history or presymptomatic tests, provided that the person is able to do the job.

Genetic Discrimination

Although many people believe firmly that insurers and employers are using genetic information to deny insurance or employment, surveys have found little evidence that companies are singling out genetic information for special treatment. What surveys have found is considerable fear of discrimination. A survey of educated members of genetic consumer groups found that most were afraid that information would be used against them (12). The majority of Americans thought that health insurers (85 percent) and employers (59 percent) will probably ask applicants in the future to take genetic tests (11). In Europe, 41 percent of the public thought insurance companies would use genetic tests to set premiums within the next 20 years (43). In the United States, surveys of medical directors of insurance companies and state insurance commissioners suggest that insurance companies intend to rely primarily on family histories, as they have always done (55–57). In an attempt to assess the prevalence of genetic discrimination, one group of researchers sent out over 30,000 questionnaires to members of support groups for families with Huntington's disease, hemochromatosis, and sickle cell anemia in anticipation of identifying asymptomatic people who had been discriminated against solely on the basis of genotype (13). The response rate was about 3 percent, of whom about half reported some form of discrimination. In another set of surveys (14), 1084 genetics services providers reported 76 clients refused employment and 474 refused life insurance because of genetic predisposition or carrier status, for a total of 550 persons refused employment or insurance. The 499 primary care physicians in the survey, with a median of 14 years in practice and 100 to 150 patients per week, reported 29 patients refused health insurance on the basis of genetic predisposition. Patients were asked "because of a genetic disability or disease, have you or a member of your family been refused employment or health or life insurance?" Two percent reported being denied or let go from a job, 3 percent were refused health insurance, 7 percent were refused coverage for some services, 5 percent were refused life insurance, and 1 percent were refused school admission. Most patient descriptions fell within the scope of employment and insurance practice generally and were only indirectly related to genetics and not at all to genetic testing.

Views of Minority Groups

A recent two-wave survey of 500 African-Americans and 500 members of the public before and after President Clinton's May 16, 1998, apology for the unethical treatment of African-Americans in the Tuskegee Syphilis Study (31) showed dramatic differences between African-Americans and the general public on many questions related to medical research in general (16). For example, almost three-quarters of African-Americans thought they were very likely (36 percent) or somewhat likely (38 percent) "to be used as guinea pigs without their consent," as compared with 16 and 34 percent of the general public (16). A 1995 survey on ethics and genetics, while reaching too few African-Americans for proper analysis, found significant differences in responses to half the ethical questions (18).

Forensics

Survey responses have shown overwhelming approval of DNA identification, especially for law enforcement agencies. Over 90 percent of genetics professionals and 95 percent of patients believed that persons convicted of serious crimes (not only sex crimes) should be required to have DNA fingerprinting and that the DNA should be kept on permanent file, like regular fingerprinting (44). Majorities of both professionals (59 percent) and patients (72 percent) would also require DNA fingerprinting for persons charged with, but not convicted of, serious crimes. Most professionals (80 percent) and patients (86 percent) favored DNA fingerprinting of members of the armed forces, to identify casualties. Most patients (73 percent), but fewer professionals (37 percent), would DNA fingerprint newborns to prevent mixups in the nursery. Half the patients, but only 20 percent of professionals, would require DNA fingerprinting for passport applicants. Almost half the patients (47 percent) would require it for people receiving welfare, to prevent fraud, and 34 percent

would require it for credit applicants. In the Eurobarometer surveys, 70 percent thought that genetic fingerprinting would lead to solving more crimes in the next 20 years (27).

Genetically Modified Organisms

The largest surveys of public attitudes have been done in Europe (27). In all, there is greater support for introducing human genes into bacteria to produce medicines or vaccines and for introducing genes into crop plants to make them more resistant to insect pests than there is for using modern biotechnology in food production, for developing genetically modified animals for cancer research, or for xenotransplantation (27). Majorities regarded all these technologies as useful for society, but majorities also found food production (60 percent), genetically modified animals for laboratory research (52 percent), and xenotransplantation (60 percent) risky. Only minorities of respondents believed it morally acceptable to develop genetically modified animals for laboratory research (44 percent), to use biotechnology in food production (48 percent), or to produce animals for xenotransplantation (35 percent). Japanese surveys indicated considerable public concern about the health effects of eating genetically modified foods (43).

CONCLUSION

Surveys are useful in gauging public optimism or pessimism about biotechnology, identifying sources of concern, and pointing to differences of opinion among stakeholder groups. Most surveys have shown that publics, especially in the United States, are generally optimistic about biotechnology and believe that benefits outweigh risks. Majorities support genetic research, testing and gene therapy, including germ-line gene therapy. In the United States, fears center on possible misuses of information by insurers and employers. There appear to be some substantial differences among views of professionals, patients, and the general public. The views of minority groups on issues specific to biotechnology remain largely unknown. In some areas of ethical concern, such as views on research uses of biological samples, potentially useful surveys are still lacking.

BIBLIOGRAPHY

1. National Opinion Research Center, *General Social Survey, 1972–1997, Cumulative Codebook*, University of Chicago, Chicago, IL, 1998.
2. United States Congress, Office of Technology Assessment, *Artificial Insemination: Practice in the United States: Summary of a 1987 Survey—Background Paper*, No. OTA-BP-BA-48, U.S. Government Printing Office, Washington, DC, 1988.
3. D.C. Wertz, Primary Care Physicians' Knowledge of Genetics, *Am. J. Hum. Genet.* **61**(4), A 56 (1997).
4. D.C. Wertz and J.C. Fletcher, *Ethics and Human Genetics: A Cross-cultural Perspective*, Springer-Verlag, New York, 1989.
5. World Health Organization, Hereditary Diseases Programme, *Proposed International Guidelines on Ethical Issues in Medical Genetics and the Provision of Genetics Services*, WHO, Geneva, Switzerland, 1998.
6. D.C. Wertz and J.C. Fletcher, *Soc. Sci. Med.* **46**(2), 255–273 (1998).
7. G.J. Meissen and R.L. Berchek, *Am. J. Med. Genet.* **26**, 283–293 (1987).
8. M.M. Kaback, in D. Lawson, ed., *Cystic Fibrosis: Horizons*, Wiley, New York, 1984.
9. *Science* **275**, 782 (1997).
10. United States Congress, Office of Technology Assessment, *New Developments in Biotechnology. 2: Public Perceptions of Biotechnology-Background Paper*, No. OTA-BP-BA-350, U.S. Government Printing Office, Washington, DC, 1987.
11. J.M. Boyle, *National Survey of Public and Stakeholders' Attitudes and Awareness of Genetic Issues*, National Center for Genome Research, Santa Fe, NM, 1997.
12. E.V. Lapham, C. Kozma, and J. Weiss, *Science* **274**, 621–624 (1996).
13. L.N. Geller et al., *Sci. Eng. Ethics* **2**, 71–78 (1996).
14. D.C. Wertz, *Health Law Rev.* **7**(3), 7–8 (1999).
15. D.C. Wertz and P.R. Reilly, *Am. J. Hum. Genet.* **61**, 1163–1168 (1997).
16. S.B. Thomas, *N.Y. Times*, April 27 (1997).
17. D.C. Wertz, *Health Law J.* **6**, 1–42 (1998).
18. D.C. Wertz, *Comm. Genet.* **1**(3), 175–179 (1998).
19. D.C. Wertz, *J. Am. Med. Wom. Assn.* **52**(1), 33–38 (1997).
20. S. Shiloh et al., *Psychol. Health* **13**, 1071–1086 (1998).
21. United States Department of Health and Human Services Secretary's Advisory Committee on Genetic Testing, *A Public Consultation on Oversight of Genetic Tests*, U.S. Government Printing Office, Washington, DC, 1999.
22. E.R. Babbie, *The Practice of Social Research, Second Edition*. Wadsworth, Belmont, CA, 1979.
23. D.C. Wertz, *Ethical Issues in Genetics Study*, Roper-Starch Worldwide, New York, 1994.
24. L.B. Rubin, *Worlds of Pain: Life in the Working-Class Family*, Basic Books, New York, 1976.
25. R.M. Lee, *Doing Research on Sensitive Topics*, Sage Publications, Thousand Oaks, CA, 1993.
26. D.L. Morgan, *Planning Focus Groups*, Sage Publications, Thousand Oaks, CA, 1998.
27. G. Gaskell, M.W. Bauer, and J. Durant, in G. Gaskell, M.W. Bauer, and J. Durant, eds., *Biotechnology in the Public Sphere: A European Sourcebook*, Science Museum, London, 1998, pp. 189–216, 231–298.
28. Nuffield Council on Bioethics, *The Ethics of Clinical Research in Developing Countries*, The Nuffield Council, London, 1999.
29. D.C. Wertz, in R. DeVries and J. Subedi, eds., *Bioethics and Society: Constructing the Cultural Enterprise*, Prentice-Hall, Upper Saddle River, NJ, 1998, pp. 145–164.
30. E.S. Tambor et al., *Am. J. Pub. Health* **83**(11), 1599–1603 (1993).
31. J.H. Jones, *Bad Blood: The Tuskegee Syphilis Experiment—A Tragedy of Race and Medicine*, Free Press, New York, 1981.
32. P.S. Eichstaedt, *If You Poison Us: Uranium and Native Americans*, Red Crane Books, Santa Fe, NM, 1994.
33. Advisory Committee on Human Radiation Experiments, *The Human Radiation Experiments*, Oxford University Press, New York, 1996.
34. S. Goering, in A. Dula and S. Goering, eds., *It Just Ain't Fair: The Ethics of Health Care for African Americans*, Praeger, Westport, CT, 1994, pp. 182–192.

35. K.A. Schneider and K.J. Kalkbrenner, *Perspect. Genet. Couns* **20**, S1–S8 (1998).
36. American Medical Association, *Physician Characteristics and Distribution in the United States*, 1996, American Medical Association, Chicago, IL, 1996.
37. United States, Department of Commerce, Economics and Statistics Administration, *Bureau of the Census: 1990 Census Population, Supplementary Reports, Detailed Occupation and Other Characteristics from EEO File for United States*, U.S. Government Printing Office, Washington, DC, 1992, p. 2.
38. United States, Bureau of the Census, *Estimates and Projections 1998*, U.S. Government Printing Office, Washington, DC, 1998.
39. United States Congress, Office of Technology Assessment, *The Role of Genetic Testing in the Prevention of Occupational Disease* U.S. Government Printing Office, Washington, DC, 1983.
40. United States Congress, Office of Technology Assessment, *Infertility: Medical and Social Choices*, No. OTA-BA-358, U.S. Government Printing Office, Washington, DC, 1988.
41. E. Singer, *Pop. Res. Pol. Rev.* **10**, 235–255 (1991).
42. Royal Commission on New Reproductive Technologies, *Social Values and Attitudes Surrounding New Reproductive Technologies*, Final Report, vol. 2, Royal Commission, Ottawa, Canada, 1993.
43. D.R.J. Macer, *Attitudes to Genetic Engineering: Japanese and International Comparisons*. Eubios Ethics Institute, Tskuba, Japan, 1992.

See other entries BEHAVIORAL GENETICS, HUMAN; EDUCATION AND TRAINING, PUBLIC EDUCATION ABOUT GENETIC TECHNOLOGY; GENETIC DETERMINISM, GENETIC REDUCTIONISM, AND GENETIC ESSENTIALISM; MEDIA COVERAGE OF BIOTECHNOLOGY.

RECOMBINANT DNA, POLICY, ASILOMAR CONFERENCE

CHARLES WEINER
Massachusetts Institute of Technology
Cambridge, Massachusetts

OUTLINE

INTRODUCTION

The development and laboratory use of recombinant DNA research techniques in the early 1970s made it possible to manipulate genes, opening the path to genetic engineering. The research and its potential applications were embroiled in controversy from the beginning and became one of the major ethical and public policy issues in the 1970s and 1980s. Unresolved questions persist through the present and the earlier experiences of scientists, citizens and policy makers continue to influence perceptions and actions today. In the 1970s the concern was about the potential health and environmental hazards of laboratory use of these novel research techniques. The researchers, their institutions, and their funding agencies developed a system of self-regulation to avoid hazards and to forestall legislative control. They focused on the means not the ends, on the tools of genetic engineering rather than on the moral limits. Rapid and pervasive commercialization of academic research in the field in the 1980s provoked continuing controversies on the role and nature of universities, the effects of corporate ties on education and on research goals and communication processes, and the threat of potential conflicts of interest of commercially involved academic researchers whose expert advice was sought on ethical and public policy issues. Commercial applications of recombinant DNA and related genetic engineering techniques continue to raise public concern about environmental and health hazards in agriculture and medicine. And the moral limits to applications of human genetic engineering are deeply controversial (1,2).

SCIENTISTS' ORIGINAL CONCERN ABOUT HAZARDS

At the Gordon Research Conference on Nucleic Acids in July 1973, invited specialists on DNA research heard fascinating reports of new techniques for manipulating and moving genetic material. The use of the newly discovered restriction enzymes made it possible to cut strands of DNA at specified precise points and to insert them into the DNA of other organisms, combining the hereditary material of animals and bacteria. These recombinant organisms could be replicated in billions of copies through cloning. It was apparent to the involved scientists that they now had a tool for studying the structure and functions of genes and to probe the details of DNA and its transcription in cells of higher organisms. Biologists recognized that this would open up a new field of work, enabling the posing of fundamental research questions that would not have been feasible before. They expected that the answers to these questions would help solve problems at the forefront of knowledge with important applications.

Amid the excitement about the potential of the new recombinant DNA technique some of the conference participants were alarmed over its possible immediate hazards in their own laboratories. They were concerned that using some of these hybrid DNA molecules might cause unforeseen hazards to human health and the environment. There was a possibility that harmless microbes could be unintentionally changed into human pathogens through introduction of antibiotic resistance, which was part of the technique; through the production of dangerous toxins, which was a possible outcome; or through the transformation into cancer-causing agents of materials that previously were benign. In this relatively new field there was a great deal of uncertainty and little information about the hazards.

The Gordon Conference participants asked for a special discussion of these larger questions. At that brief special session, they decided to write a letter to ask the National Academy of Sciences to study the potential hazards and to devise a plan to do something about them. They voted by a large majority to compose the letter and they approved the content of it. They also voted, this time by a slim majority, to send a copy of the letter to be published in *Science*. The reluctance of many of the participating scientists to call public attention to the problem was an indication of a continuing conflict. They were concerned about a possible public health problem, and yet they feared that talking about it publicly might bring intrusion, as they saw it, into the scientific process. The Gordon Conference letter, replete with technical language, was intended for other scientists (1, pp. 70–80). It was published in *Science* in 1973 and did not generate much public attention (3).

THE BERG COMMITTEE LETTER

The National Academy asked Paul Berg, a distinguished biochemist and a principal researcher in the field, to

organize a group of scientists to consider the issues. They met at MIT on April 17, 1974, and planned a conference for February 1975 to evaluate the hazards of the research and ways of dealing with them. Feeling a sense of urgency, they also drafted a letter to alert the larger community of biologists. Two months after the MIT meeting, Berg described these actions and the group's motivations in a letter to a colleague in England:

> We met at MIT for a day and settled on the idea of calling a conference next February of those scientists working on methods of joining DNA molecules and particularly those involved in constructing hybrid DNAs. It was our plan that one of the major purposes of the Conference, besides a report on the scientific progress, would be a wide ranging discussion of potential hazards growing out of these types of experiments. Were there any experiments that should not be done? How could such a moratorium be proposed or enforced? In short, we expected a frank and searching review of what people were doing or wanted to do, particularly from the point of view of whether they should be done. But as we talked we realized that the pace of events might not wait for February and that some of the experiments many people would agree could be hazardous would be done by then (e.g., attempts to fuse portions of Herpes DNA to appropriate plasmids for cloning in *E. coli* were imminent). Since the technology for constructing hybrids has become ridiculously simple, that fear was well founded.
>
> Consequently we decided to devise a letter to be submitted to *Science* and *Nature* calling on scientists to defer certain kinds of experiments until these potential hazards could be better evaluated and certainly until there was an opportunity to discuss the issues at the February meeting (4).

Drafts of the Berg committee letter were circulated privately among the relevant scientists, and in July 1974 the final version was published in *Science* and in *Nature* (5). Why did the letter go public? Because the committee felt it was the quickest way to bring the potential hazards to the attention of the community of researchers who would be likely to use the new recombinant DNA techniques. They felt that the situation was urgent, because of pending experiments and because the power and fruitfulness of these research tools rapidly would attract many scientists to the field who were not experienced in handling pathogenic organisms. The letter called for a voluntary moratorium, a temporary deferral of those experiments which at the time were thought to be potentially hazardous. This appeal for self-restraint was linked to an end point, the conference scheduled for February 1975.

The response to the letter by the relevant scientific community was generally favorable. When the draft was read at the Cold Spring Harbor meeting in June, 12 of the European scientists in attendance immediately drafted a letter to John Kendrew, director of the European Molecular Biology Organization, requesting urgent consideration of the matter. They felt that it was essential that research utilizing the new recombinant DNA techniques be made possible in Europe by providing appropriate special risk laboratories.

When the Berg et al. "moratorium letter" was published in *Nature*, it was followed by responses from leading British biologists. Michael Stoker, head of the Imperial Cancer Research Fund wrote:

> No doubt a good many dirty tricks have been attempted and discarded by nature in the course of evolution, but the disquiet arises from the utterly novel associations of genetic material which are now possible. The potential benefits should, therefore, be delayed, not for ever, but until consequences can be assessed, and preliminary experiments carried out under conditions of maximum security. . . .[I]t is encouraging that the very leaders in the field have taken the initiative and have been supported by the [National Academy of Sciences]. It is now to be hoped that academics and learned societies in other countries will add their weight, and that international organizations such as the European Molecular Biology Organization will lend support. . . .For many it will be a test of self denial and social responsibility in the face of strong intellectual temptation (6, p. 278).

Kenneth Murray of the University of Edinburgh stated:

> The NAS request is both reasonable and responsible and deserves to be universally respected. It recognizes both the difficulty in evaluating real or potential hazards that may be involved in such work, as well as the obvious criticism that these will remain obscure in the absence of experimental study; urgent consideration of the latter is explicitly recommended. Fears that the proposed limitations to experiments will seriously obstruct research in vital areas of biology seem unfounded. The NAS initiative, by focusing attention on the hazards involved, could well promote rather than hinder work on in vitro recombination in animal viral systems, an area believed by many to hold the key to gene therapy in its broadest terms. . . .[I]f we follow the moderate tone set by the NAS we shall be careful not to oversell the social benefits devolving from recent experiments (7, p. 279).

THE 1975 ASILOMAR CONFERENCE

The February 1975 meeting at the Asilomar Conference Center in California evaluated knowledge in the field and its potential for research. It was the equivalent of an international review conference which ordinarily would be held well into the development of a research field and not at such a very early stage. The detailed review enabled the conference participants, who were the researchers and the potential researchers in the field, to consider the potential risks and ways to control them. The motive from the start was to avoid public interference and to demonstrate that scientists on their own could protect laboratory workers, the public, and the environment. Of course, there is that contradiction again: They were dealing with a public health issue and simultaneously attempting to keep the public out of it.

Initially it was not clear whether any media representatives would be allowed to attend the publically funded conference, but later the organizers decided to limit press attendance to eight invited reporters. A deal was struck with 16 journalists, most of whom were invited, that they would not report on the conference until it was over, because things would be too much in flux. That pleased the reporters because they did not have to call in stories to their editors every day. Instead, the telephone booths were jammed with scientists calling their laboratories in Europe and the United States about the need to tool up for this very exciting new research. The conference

gave them an opportunity to learn as much as possible about the recombinant DNA techniques, and it stimulated the growth of the field while producing a framework for pursuing it safely.

Several technical working groups met independently over a period of months in preparation for the conference. The most active was the Plasmid Working Group, focusing on the circular pieces of DNA which were the main tools for this new technique. They scoured the literature and their own knowledge, talked with other people in the field, and produced a detailed technical document (8). Reports of the working groups were presented and discussed at the meeting and one session was devoted to presentations of lawyers on policy and liability issues. Participants paid special attention to their legal responsibility for damage resulting from their laboratory work.

The narrow technical focus of the conference was evident in the opening remarks of David Baltimore, one of the organizers. He first acknowledged that the techniques that were developed could have applications in a number of areas, including biological warfare, and that it had larger societal implications, but that such issues would be excluded, since there was a full agenda of technical issues:

> The issue that ... [brings] us here is that a new technique of molecular biology appears to have allowed us to outdo the standard events of evolution by making combinations of genes which could be immediate natural history. These pose special potential hazards while they offer enormous benefits. We are here in a sense to balance the benefits and hazards right now and to design a strategy which will maximize the benefits and minimize the hazards for the future (9).

What happened at Asilomar? The recombinant DNA issue was defined as a technical problem to be solved by technical means, a technical fix. Larger ethical issues regarding the purposes and the long-term goals of the research were excluded, despite the rich discussions that had occurred among geneticists and other biologists in the 1960s about where to draw the line when it became possible to do genetic engineering. The 1960s discussions led to congressional proposals for anticipatory study of the ethical limits of genetic engineering, which were resisted as premature by several leading biologists (10). Instead of those longer-term issues, the focus at Asilomar in 1975 was on safety of the newly developed technical tools for genetic engineering, on the means not the ends.

The Asilomar participants adopted provisional safety guidelines based on a two-part system of physical and biological containment of potentially hazardous recombinant organisms (11). The extent of physical containment was graded according to the anticipated level of hazard an organism might present if it escaped the laboratory, ranging from good laboratory technique for those experiments deemed to be of low hazard, to hooded glove boxes, negative pressure, showers and clothes changes for laboratory workers dealing with organisms thought to be especially dangerous. Biological containment would introduce mutations in the organisms that were to be used in the experiments so that if they escaped they could not survive in the environment beyond the laboratory.

RECOMBINANT DNA ADVISORY COMMITTEE

In November 1974 the NIH had established the Recombinant DNA Advisory Committee (RAC), advisory to the director of NIH. The first meeting was held immediately after the Asilomar conference at the end of February 1975. RAC appointees were knowledgable researchers in the field, who were asked to develop and extend the Asilomar provisional safety guidelines to control all recombinant DNA work at institutions receiving NIH funding of any kind. They were designing safety protocols that had the potential for restricting their own work. These controls were to be administered by the NIH, which funded and encouraged the research and therefore was itself in a position of conflict of interest. NIH officials acknowledged the potential conflict, and maintained that although NIH was not a regulatory agency, it had the best expertise in the field and needed to act in the absence of any other government group playing a role. Similar efforts were also underway in other countries.

During 1975 and 1976 scientists on the RAC argued about whether the proposed guidelines were too strict or too permissive, and the document went through many drafts (12). All of this occurred in the absence of risk assessment experiments. At the same time scientists at laboratories throughout the country were tooling up to use the new technique and were impatiently waiting for the green light that would allow them to proceed as rapidly as possible. They exerted a great deal of pressure on RAC and NIH. The process of establishing safety rules involved a series of compromises aimed at achieving a consensus within that portion of the scientific community affected by the guidelines while providing assurances to the public that they would be protected from possible hazards.

LOCAL AND NATIONAL POLITICAL RESPONSES

The long expected NIH safety guidelines for recombinant DNA were approved by the director of NIH on June 23, 1976. On that day when the green light flashed, an extraordinary event took place in Cambridge, Massachusetts. Scientists from MIT and Harvard and representatives of NIH appeared at a special City Council hearing. They had been invited to explain to the citizens of Cambridge why the scientists themselves had been arguing about the safety of recombinant DNA and whether the guidelines were adequate to protect the communities in which the research was to be done. Was there any danger to citizens? Who was going to monitor and enforce the safety standards? Could the scientists and their universities be trusted to regulate themselves? Testimony by several biologists that recombinant DNA techniques posed few risks and that they could be contained by the new guidelines was countered by testimony from other biologists who argued that the guidelines were inadequate and that they were formulated by self-interested advocates of the research. After a second hearing in July 1976 the City Council established a citizens' review board to examine the problem and, pending the outcome of the board's deliberations, placed a temporary ban on experiments classified in the guidelines as posing moderate to major hazards.

The nine member Cambridge Experimentation Review Board met twice weekly for a total of more than 100 hours over a four month period. About one half of the time was used for testimony by scientists on both sides of the issue. The board presented its report to the city council on January 5, 1977, recommending the creation of a city biohazards committee to oversee adherence to the NIH guidelines for all recombinant DNA work in the city whether funded by NIH or not, and several additional safeguards on experimental procedures, containment and testing of organisms (1, pp. 302–307; 13). These community confidence building measures were incorporated in a city council ordinance passed in February 1977, which was the first recombinant DNA legislation in the United States and was interpreted as a qualified public endorsement of the NIH guidelines (14).

A major fear of the recombinant DNA scientists was that their own early concern about laboratory safety had initiated public scrutiny of the new research. This was emphasized by the events in Cambridge and in other communities such as Ann Arbor, Cambridge, San Diego, New Haven, and Princeton where academic biologists were tooling up to use recombinant DNA techniques. By 1978, 16 separate bills had been introduced in Congress to regulate recombinant DNA safety standards by making the NIH guidelines mandatory for both publically and privately funded research and providing enforcement and punishment provisions for any violations. Research universities and scientific organizations saw this local and national activity as public "overreaction" threatening their control of laboratory safety procedures and their research funding. They vigorously lobbied to oppose or influence legislation. Several prominent biologists who had shared the early concern about possible safety hazards of the research publicly recanted, and a resolution to Congress signed by most of the participants in a 1977 Gordon Conference stated that they previously had overstated the risks and now could provide reassurance that the work was safe (15). In the end no legislation was passed by Congress.

EVOLUTION OF RESEARCH GUIDELINES

By 1979 the NIH Recombinant DNA Guidelines had been made far more permissive than the original 1976 version. More than 90 percent of U.S. research in the field was either no longer covered by the guidelines or was subject to only minimal controls equivalent to standard laboratory practice. By 1982 most experiments subject to the guidelines were controlled at the local level through institutional biosafety committees and RAC reviewed only research that had the potential for special safety problems. No demonstrated harm had been caused by the research as conducted under the guidelines. A limited amount of risk assessment research had been done during that period, and several small consensus workshops of scientists in the field were held to review existing knowledge and to refute the earlier concerns (1,2). NIH's approach to the guidelines was that they would be flexible enough to respond to new scientific knowledge. That also opened them up to flexible response to pressures from researchers and their interests, pressures from industry, and pressures from national policy priorities and political interests.

BEGINNINGS OF COMMERCIALIZATION

Downgrading of the guidelines coincided with rapid commercialization of the field and the involvement of academic scientists in biotechnology companies. In November 1974 during the moratorium period, a patent for the recombinant DNA technique was filed by Stanford University and the University of California on behalf of two of the scientists who developed the technique. The patent was granted in 1980 after the Supreme Court decision allowing patenting of human-made organisms (16). Biologists and their universities became involved in what soon became almost a complete commercialization of the work. In the 1980s political climate of deregulation the U.S. biotechnology industry was promoted as a national priority. Emphasis was on government, industry, and media claims of medical, practical, and economic benefits of the research and the need to develop the industry. Critical questions about the health and environmental safety of research techniques and products were met by arguments that if the United States did not move forward rapidly in biotechnology, the country would lose out in international competition. The "gene gap" argument was deployed to resist special regulation of the field.

ENVIRONMENTAL RELEASE OF MODIFIED ORGANISMS

As the guidelines faded away for most laboratory work, attention shifted from the accidental escape of genetically engineered microorganisms to the intentional release of these organisms into the environment for agricultural purposes. The U.S. Department of Agriculture (USDA) and the U.S. Environmental Protection Agency (EPA), the agencies who would ordinarily become involved, initially claimed that they did not yet have the expertise to evaluate the possible hazards and they urged NIH to provide safety oversight for these applications through the RAC. Evaluation by the RAC seemed like a very comfortable approach for scientists and companies who had been working with it. In the absence of federal legislation for recombinant DNA, industry had been in voluntary compliance with the NIH guidelines. It was not until 1984 that EPA issued an interim policy statement on field testing of genetically engineered microbial pesticides. By that time NIH had approved proposals for small scale field testing of a genetically modified organism that was to be sprayed on strawberry and potato plants to prevent frost damage. The "ice-minus" controversy of the mid-1980s involved approvals by NIH, EPA, and California agencies, legal challenges by genetic engineering critic Jeremy Rifkin, congressional hearings, and protests and demonstrations by citizens in the community where field testing was to occur. As in Cambridge several years earlier the citizens asked, "Why are we the last to know?" The test plot was definitely in their backyard, but they were not informed of its exact location. They were also concerned

about unresolved safety questions raised by ecologists. By the time the tests were finally conducted in 1987, RAC's role in approval of environmental release of genetically modified organisms had been superceded by EPA (17).

HUMAN GENE TRANSFER EXPERIMENTS

The RAC also played a transitional role in the oversight of experiments in human gene transfer, generally referred to as gene "therapy" to reflect the as yet unrealized hopes of its advocates. In 1983 the RAC responded to the report of the President's Commission on Bioethics' study of genetic engineering which considered several approaches to the oversight of future human genetic engineering. The commission's study was initiated after the leaders of the three major U.S. religious groups wrote a letter to the President stimulated by the 1980 Supreme Court decision permitting patenting of genetically engineered organisms. They called for study of the ethical issues associated with genetic engineering and observed that "no government agency or committee was currently exercising adequate oversight or control, nor addressing the fundamental ethical issues in a major way" (18). RAC's response to the Commission's report was to establish a Working Group to consider whether it would review proposals for human gene transfer. In 1985 RAC's "Points to Consider in the Design and Submission of Human Somatic Cell Therapy Protocols" was issued by NIH. The RAC said it would be willing to review proposals for human gene transfer protocols for somatic cells but would not "at present entertain proposals for germ line alterations" (19,20). When pressed by a public interest group, the Council for Responsible Genetics, to specifically ban human germ-line engineering, the committee refused. Leroy Walters, the bioethicist who had been for many years a member of RAC and was the head of its human gene therapy subcommittee, subsequently argued that in his view voluntary programs of germ-line genetic intervention were "ethically acceptable in principle" (19).

Gene therapy became the primary task of the group in the late 1980s, and since then it has dealt with the scientific validity of proposals as well as risks for human subjects, the adequacy of informed consent, the role of local institutional review boards, and the liability of researchers. RAC nurtured the development of human gene therapy by applying the clinical standards of biomedical ethics, but bypassed the larger ethical issue of whether it should be done at all. The role of RAC remained as advisory to the director of NIH. In 1995 the Food and Drug Administration (FDA) became the regulatory oversight agency for human genetic engineering, with RAC playing an advisory role in reviewing proposals involving novel techniques or applications (21–24). The adequacy of this approach was questioned by Congress and government agencies when revelations and allegations about violations of regulations were reported in the media in September 1999. These included abuses of informed consent procedures, failure to report adverse effects and harm to human subjects, and possible commercial conflicts of interest. Several clinical trials were shut down by FDA and the situation was under intense review in 2000.

LIMITATIONS OF SELF-REGULATION

Throughout RAC's history — from its creation in 1974 to deal with initial concerns about laboratory safety to its current role in human gene transfer experiments — it has been friendly to researchers and dominated by their interests. At the same time the work of the committee has been relatively open and visible. NIH made efforts to create a full public record of RAC deliberations and documents in addition to the announcements of meetings, proposed changes in the guidelines and decisions required to be published in the *Federal Register*. However, very few citizens read that relatively inaccessible, small print publication. Nor do many people have the opportunity to travel to Bethesda, Maryland, to sit in on committee meetings. The RAC minutes list the noncommittee members who attended the meetings. As the commercialization of genetic engineering increased from 1980 on, the record shows that representatives of companies were consistently present to follow the deliberations and look after their interests.

Public participation on the RAC was broadened in 1978 in response to complaints that it was dominated by self-interested researchers. Yet there were built-in limits and constraints to this participation because most of the issues placed before the committee were technical and often beyond the expertise of the nonscientists. Another problem was that RAC was increasingly asked to review industry proposals. Biotechnology companies were in voluntary compliance with the guidelines and sought NIH approval for their recombinant DNA work with the condition that proprietary information would be kept confidential, as was the practice with federal regulatory agencies, even though NIH was a research-supporting agency. As a result, public representatives on the committee frequently were not able to report to the public about information relevant to environmental and public health.

The development of genetic engineering clearly involves more than the safety issues that have been the major focus of RAC's mandate and activities. The larger ethical concerns about where to draw the line in applications of genetic engineering were occasionally discussed when raised by some members of the committee or at the request of outside groups. RAC, however, resisted taking a stand against the use of recombinant DNA techniques for biological warfare and refused to recommend an unambiguous ban on the review of proposals for human germ line intervention (25). Instead, RAC's emphasis was to develop safe procedures for the research, focusing on *how* to do it rather than *whether* it should be done. As Leon Kass observed in 1997, "the piecemeal formation of public policy tends to grind down large questions of morals into small questions of procedure" (26). Recombinant DNA research was safer as a result of the NIH guidelines developed by RAC. The biologists at Asilomar in 1975 and the subsequent generations of RAC members raised important safety issues and set standards for good laboratory practice.

Despite the success in improving the safety of research, the quasi self-regulation model developed in the recombinant DNA controversy is not adequate for expressing and enforcing societal and moral limits for

potential genetic engineering applications such as human cloning or human germ-line interventions. These potential applications are not inevitable, and they raise profound issues beyond laboratory and environmental safety and patients' rights. They occur in a context of increasing genetic determinism, pervasive commercialization, and aggressive efforts to sell genetic intervention as a cure-all for medical and even social problems. Separation of the technical issues from the ethical issues, and the narrowing of ethical concerns to clinical biomedical ethics, limits meaningful public involvement and obscures the larger picture.

ACKNOWLEDGMENT

This historical summary draws on my observations and documentation of the recombinant DNA controversy from 1975 to the present, utilizing archival materials and interviews collected from 1975 to 1979 in a project under my direction and deposited in the Recombinant DNA History Collection available for study at the Institute Archives and Special Collections, Massachusetts Institute of Technology. Portions of this account are included in C. Cranor, ed., *Are Genes Us? The Social Consequences of the New Genetics*, 1994, pp. 31–51 and C. Weiner, *Health Matrix* **9**, 289–302 (1999).

BIBLIOGRAPHY

1. S. Krimsky, *Genetic Alchemy: The Social History of the Recombinant DNA Controversy*, MIT Press, Cambridge, MA, 1982.
2. S. Wright, *Molecular Politics: Developing American and British Regulatory Policy for Genetic Engineering*, 1972–1982, University of Chicago Press, Chicago, IL, 1994.
3. M. Singer and D. Soll, *Science* **181**, 1114 (1973).
4. P. Berg to H. Kornberg, June 18, 1974, Recombinant DNA History Collection, Massachusetts Institute of Technology, Cambridge, MA.
5. P. Berg et al., *Science* **185**, 303 (1974); *Nature* **250**, 175 (1974).
6. M. Stoker, *Nature* **250**, 278 (1974).
7. K. Murray, *Nature* **250**, 279 (1974).
8. Plasmid Working Group, *Proposed Guidelines on Potential Hazards Associated with Experiments Involving Genetically Altered Microorganisms*, Asilomar Conference, February 24, 1975. Recombinant DNA History Collection, Massachusetts Institute of Technology, Cambridge, MA.
9. Audiotape of the International Conference on Recombinant DNA Molecules, Asilomar, February 24, 1975. Recombinant DNA History Collection, Massachusetts Institute of Technology, Cambridge, MA.
10. U.S. Senate, 1968. *Hearings before the Subcommittee on Government Research of the Committee on Government Operations*, 90th Congr., 2nd Sess., on S.J. Res. 145, a joint resolution for the establishment of the National Commission on Health, Science & Society, March 7, 8, 21, 22, 27, 28 and April 2, 1968, p. 5.
11. P. Berg et al., *Science* **188**, 994 (1975).
12. National Institutes of Health, Office of Recombinant DNA Activities, *Guidelines for Research Involving Recombinant DNA Molecules*, National Institutes of Health, Bethesda, MD, 1976, in press.
13. Cambridge Experimentation Review Board, *Guidelines for the Use of Recombinant DNA Molecule Technology in the City of Cambridge*, January 5, 1977.
14. R. Goodell, *Sci. Technol. Hum. Values* (Spring), 36–43 (1979).
15. W. Gilbert, *Science* **197**, 208 (1977).
16. *Diamond, Commissioner of Patents and Trademarks v. Chakrabarty*, No. 79–136. Argued March 17, 1980, decided June 16, 1980.
17. S. Krimsky and A. Plough, *Environmental Hazards: Communicating Risks as a Social Process*, Auburn House, Dover, MA, 1988, pp. 75–121.
18. U.S. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Splicing Life: A Report on the Social and Ethical Issues of Genetic Engineering with Human Beings*, U.S. Government Printing Office, Washington, DC, 1982, pp. 95–96.
19. L. Walters and J.G. Palmer, *The Ethics of Human Gene Therapy*, Oxford University Press, New York, NY, 1997, pp. 148–151.
20. I.H. Carmen, *Am. J. Hum. Genet.* **50**, 245–260 (1992).
21. H. Miller, *Wall Street J.* A18 (May 10, 1994).
22. L.E. Post, *Hum. Gene Therapy* **5**, 1311–1312 (1994).
23. S. Krimsky, *Hum. Gene Therapy* **5**, 1313–1314 (1994).
24. D.T. Zallen, *Hum. Gene Therapy* **7**, 795–797 (1996).
25. National Institutes of Health, *Recombinant DNA Advisory Committee, Minutes of Meeting, January 30, 1989*, National Institutes of Health, Bethesda, MD, 1989, pp. 19–37.
26. L. Kass, *New Republic* No. 18, 17–26, (June 2, 1997).

See other GENE THERAPY entries.

RELIGIOUS VIEWS ON BIOTECHNOLOGY, BUDDHISM

RONALD Y. NAKASONE
Graduate Theological Union
Berkeley, California

OUTLINE

INTRODUCTION

The modifications and improvements of living organisms, including human beings, and the development of microorganisms through biotechnology to produce or modify products to improve plants or animals challenge Buddhists to reexamine their doctrines, sharpen their interpretative insights, and expand their moral imagination. Buddhist views on biotechnology require an understanding of the spiritual goals that govern the Buddhist life, its doctrines, and the practical demonstration of its ideals. This article begins with a review of Buddhism's origins and spread, continues with a description of the major doctrines, and surveys selected Buddhist responses to biotechnology. Since modern science and technology and population increases have resulted in problems neither confronted nor anticipated by the Buddha and, until very recently, by Buddhist thinkers, this article also describes the Buddhist attitude toward change and new knowledge. Buddhist attitudes toward scientific and technological manipulation of human beings and nature, however, are often tempered by indigenous traditions such as Confucianism and Shinto, by modernization, and by Western culture.

Buddhist interest in biotechnology is relatively recent, and its impact has been felt most keenly in technically advanced countries, such as Japan, and technologically advanced regions of China and other countries. Buddhist thinkers of Third World countries and their devotees who struggle to survive from day to day are concerned with agricultural yields, clean water, political stability, and basic medical care. The dearth of reflections on biotechnology from developing countries such as Cambodia, Laos, and Vietnam indicate their preoccupation with meeting basic survival needs. For these Buddhists exotic high tech innovations, such as mapping of the human DNA or organ transplants, are inconceivable luxuries. While Buddhist thinkers in Japan are reflecting on biotechnology, their thinking is dominated by modern Western bioethical paradigms. They are, however, beginning to reshape questions into Buddhist categories and envisioning Buddhist solutions. The speeds at which change is taking place and new possibilities are emerging, outpace, at least for the moment, religious and moral thinking.

ORIGIN AND SPREAD

Siddhārtha Gautama (563–483 B.C.E.) (1) founded the Buddhist faith and community in what is now northeast India and Nepal. The community remained a single unit until about a hundred years after the founder's death. Disagreements centering on the status of the *arhat*, the archetype personality, the status of Śākyamuni, the historical Buddha, and the *vinaya*, the monastic rules of conduct, split the early community into the Sthavira and the more liberal Mahāsaṅghika. The Sthavira argued for the perfectibility of human-nature in the guise of the *arhat*, the humanity of the historical Buddha, and strict observance of the *vinaya* outlined by the Buddha. Present-day Theravāda, an offshoot of the Sthavira, claims to observe the faith established by its founder. The Mahāsaṅghika, on the other hand, believed that the *arhat* is not completely free from impurities. Its devotees understood Śākyamuni to be a manifestation of a transcendental Buddha who is pure, infinite, and eternal. The tradition stressed the spirit of the *vinaya*, not its letter. These initial disagreements led to further splintering. Buddhist documents mention the existence of as many as 34 monastic sects between the second and fourth centuries after the Buddha's passing.

The origins of Mahāyāna (great vehicle) Buddhism is obscure. Many scholars believe that Mahāyāna (2) emerged during the two centuries between 100 B.C.E. and 100 C.E. from ideas advanced by the Mahāsaṅgika and related sects, and as a reaction against the aloofness of the monastic sects. Others, notably Akira Hirakawa, a Japanese Buddhologist, argue that Mahāyāna began as a lay movement that appeared immediately after the death of the historical Buddha. These devotees who honored the memory of the Buddha at the *stūpas*, memorials that housed his relics, evolved over time their own liturgies, doctrines, and institutions (3). Still others argue that Buddhism's encounter with non-Indian peoples and ideas and the increased influence of the laity stimulated the monastic tradition to redefine their goals, iconography, and doctrines to be more inclusive. These early Mahāyānists referred to themselves as *bodhisattvas*, "beings who aspire for wisdom." The bodhisattva, an outgrowth of the idealization of the historical Buddha, vows to save all beings before he or she achieves full enlightenment.

After subduing his adversaries, the Gupta king, Aśoka (circa 274–236 B.C.E.) embraced the Buddhist faith and sent missions throughout India, Sri Lanka, North Africa, Macedonia, and Central and South Asia. These missions initiated the gradual spread of Buddhist and Indian culture. Sri Lankan chronicles report the establishment of Theravāda Buddhism in the later half of the third century B.C.E. By the eighth century C.E. Indian culture, including Buddhism, stretched from the east coast of the Indian subcontinent all the way to Vietnam, and Bali in the Indonesian archipelago. Theravāda eclipsed Mahāyāna and remains dominant today. Buddhist missions did not leave a lasting impact in North Africa and the Near East. Meanwhile, to the north, Buddhism fared better. It had a well-established presence in Central Asia by the second century B.C.E. Among the many schools, Mahāyāna and the Sarvāstivāda, an influential branch of the Sthavira linage, were the most strongly represented. Buddhist culture established itself in Khotan, Kucha, Turfan, and other city states that straddled the caravan routes to China. Buddhism entered China sometime during the first century B.C.E. and the beginning of the common era (4), and gradually became an integral part of the national life. Distinct Chinese forms of Buddhism emerged between 500 and 800, its most prosperous and creative period. Buddhism continued its eastward advance into Manchuria, Korea, and Japan. It officially entered Korea in 372, and by about 525 had penetrated the entire peninsula. Buddhism arrived in Japan in 552. After six hundred years, the Japanese evolved forms of Buddhism that suited and reflected their temperament.

Mahāyāna in the form of Vajrayāna, later called Tantric Buddhism, traveled to Tibet in the seventh century.

It conceived the Buddha as a cosmic body and as the substance of all things and all beings. By harnessing the forces that pervade the universe, the devotee can achieve Buddhahood in this very life and this very body. Tibetan monks carried their faith to Mongolia in 1261 and again in 1577. The fourteenth Dalai Lama (1935–), the spiritual and secular leader of Tibet, fled his occupied country in 1959 with tens of thousands of other Tibetans. The Chinese claim that Tibet has always been part of China.

Since the mid-nineteenth century when Chinese laborers joined the California gold rush, Asian immigrants have carried their Buddhist faith to Hawaii, North and South America, and other parts of the globe. The 1893 World's Parliament of Religions in Chicago introduced Buddhism and other non-Christian faith traditions to the West. The United States and other Western nations today have sizable ethnic Buddhist communities. D.T. Suzuki (1870–1966) and Alan Watts (1915–1973) popularized Zen Buddhism at midcentury and spawned a still small, but vital American–Buddhist community. Political refugees from Southeast Asia introduced Theravāda Buddhism during the latter half of the century. At present, Theravāda Buddhism is dominant in Sri Lanka and the Southeast Asian countries of Burma, Kampuchea, Laos, and Thailand. Mahāyāna exists in North and East Asia. Tibet, Mongolia, China, and Japan have substantial Buddhist populations. In a once Buddhist country, South Korean Buddhists constitute about one-fourth of the current population of 50 million. The social reformer and former Minister of Law, Bhimrao Ramji Ambedkar's (1891–1956) conversion to Buddhism in 1956 generated a Buddhist rival in India after 700 years of its disappearance from the land of its origin.

BELIEFS AND DOCTRINES

Siddhārtha Gautama began his spiritual journey with the question of human suffering that accompanies old age, sickness, and death. After six years of spiritual exercises Gautama realized the Dharma, the truth of *pratītyasamutpāda* (dependent co-arising or interdependence) and became the Buddha, "the Enlightened One." The Buddha awakened to the truth that all things and all beings are mutually related and mutually dependent. *Pratītyasamutpāda* represents the ideological content of the enlightenment and is the common theme throughout Buddhist thought and practice (5). The history of Buddhist thought can be understood to be simply an unfolding of the implications inherent in this central idea.

This overview begins with a discussion of the temporal and relational aspects intrinsic to *pratītyasamutpāda*. It proceeds to explain the notions of *karma*, *samsara*, *nirvana*, *anātman*, and other key Buddhist ideas, and the Four Noble Truths within the context of *pratītyasamutpāda*. The following section, The Buddhist Posture, discusses the implications for our interest in biotechnology.

Pratītyasamutpāda can be understood to be an extension of karma, the law of cause and effect. The idea of karma, literally "action," appeared approximately two or three centuries before the birth of Siddhārtha Gautama and is closely associated with the notions of samsara, literally "passage," and personal responsibility. There are three classes of karma: good, bad, or morally neutral. An individual's present station in life has been determined by the moral quality of action, or karma, generated in the past. Deeds committed in the present life affects one's status in the next. Buddhists divide karma into three categories: mental, verbal, and physical. Early Buddhists exerted considerable effort debating whether the essential nature of karma is mental or physical. Theravāda concluded that the mental is the essence of karma. Volition, which is mental activity, generates verbal and physical action. It is thus essential, if one wishes to realize nirvana or spiritual peace, to quicken thoughts that generate behavior that lead to that end. The notion that it is not possible to escape the consequences of one's deeds is intrenched in present day Theravāda Buddhist cultures. Sri Lankan Buddhists, for example, explain the death of an impaired infant as the result of the working of karma (6). Present-day Thai women considering abortion of a HIV-afflicted fetus or an unwanted pregnancy from rape or forced prostitution weigh the consequences of an unfavorable rebirth from their poverty as they struggle to be faithful to Buddhist teachings (7). While the idea of karmic retribution is also very strong among Mahāyāna devotees, the inexorable consequence of karmic action is mitigated by the compassion of Buddhas and Bodhisattvas.

The ideas of personal karmic retribution and successive lives are part of the fabric of popular Indian thought and played a key role in the development of Indian Buddhist doctrine. Rebirth, however, is not a necessary tenet of the Buddha's teaching and was not central to the development of East Asian Buddhist thought. Chinese, Korean, and Japanese beliefs in spirits and soul were not based on rebirth (8). Moreover the Buddha maintained that claims of rebirth, like questions of whether life continues after death, are not empirically verifiable and cautioned against such speculations. Does the Enlightened One exist after death? Or not exist? are two of 10 questions that the Buddha refused edification. Hakuin, (1685–1768), the Japanese cleric, offered a similar response when a wealthy parishioner queried about the nature of death: "Why ask me?" The parishioner replied, "Because you are a Buddhist monk." Hakuin retorted, "But not a dead one." Rather than engaging in endless speculation, the Buddha proposed that we deal directly with those problems that will ease human suffering and lead to spiritual ease.

In addition to the temporal understanding of karma, *pratītyasamutpāda* also describes the simultaneous presence of cause and effect. The individual threads of the warp and woof of a piece of fabric illustrate this expanded notion of karma. Individual threads constitute the entire fabric; the fabric in turn defines each thread in relation to all other threads. The metaphor illustrates the mutual dependency of cause and result and, by extension, the mutuality of all things. In a mutually dependent universe each individual does not simply exist in the world. By being involved in the world, he or she helps to create the world through the manner in which he or she thinks, speaks, and lives. This understanding of *pratītyasamutpāda* provides a vision of identity and responsibility to all beings and all things, and

quickens a sense of gratitude for all things and beings. *pratītyasamutpāda* dissolves the preoccupation with the self and gives rise to sentiments of compassion and service to others.

Mahāyāna documents that appeared during the first century of the common era interpret *pratītyasamutpāda* to be compassionate and morally purposeful. The *Larger Sukhāvatīvyūha sūtra* casts the doctrine of *pratītyasamutpāda* in the myth of the Bodhisattva Dharmākara, a spiritual hero who vows to forgo supreme enlightenment until all beings enter the Pure Land, the realm of spiritual ease. Dharmākara and other spiritual heroes accomplish this monumental task by transferring the vast store of merits they have accumulated over innumerable eons to the spiritually impoverished. *Pariṇāma*, or the transference of merit, is a soteriological idea based on the belief that an individual's life is irrevocably linked with all beings and things. Bodhisattvas and Buddhas do not literally withdraw merits from their merit repositories and deposit them in another's. Merits are "transferred" in the sense that we benefit from their spiritual exercises. *Pariṇāma* softens the harsh and uncompromising individualism of karma. In contrast, the rigidly individualistic view of karma dissolved society into isolated individuals, and it fails to acknowledge the mutuality among all beings and the complexity of the human experience. Also in an interdependent world, countless karmic forces intersect to often thwart our noblest intentions and propel us to violate our deepest instincts.

Though Theravāda and Mahāyāna Buddhism emphasize differing aspects of *pratītyasamutpāda*, they agree that change is the nature of reality, that suffering is endemic to the human condition, and that *nirvana* or spiritual bliss is a transcendent reality. For Theravāda, *nirvana* means the transcendence of *samsara*, the realm of suffering that is associated with change. Mahāyāna, on the other hand, identifies *nirvana* with *samsara* and thus speaks of spiritual release in the world of change. Both traditions assert the doctrine of *anātman* (non-self) which allows a person to identify and empathize with others. Early Buddhists explained *anātman* by analyzing the body and mind through five aggregates: form, sensation, perception, mental formations, and consciousness. A person, composed of these five constantly changing and mutually dependent aggregates, is devoid of a substantial and abiding self. Further investigation by the early Buddhist thinkers expanded these 5 categories into 75 dharmas or elements of existence. Eventually form was interpreted to include all material things. Mahāyāna accepted the view of the non-self but advocated a more radical view that dharmas themselves are without substantial and enduring reality. The Mahāyāna view is supported by the *prajñapāramitā hṛdaya sūtra* or Heart Sutra which proclaims: "Form is emptiness [*śūnyatā*] and emptiness [*śūnyatā*] form." Later Nāgārjuna (circa 150–250), a Mahāyāna thinker, asserted "*Pratītyasamutpāda* is *śūnyatā* [emptiness]." The empirical person, the result of the coming together of countless dharmic elements and conditions correlates with "form." "Emptiness" of the self in early Buddhism evolved to designate the intrinsic reality of an individual devoid of all accidental characteristics or *śūnyatā*. *Śūnyatā* thus refers to the "suchness" or "thusness" of the person.

The Buddhist notion of *anātman* came under stiff attack. Critics asked: "If there is no-self, how does one account for the need for rebirth? Who is the agent responsible for action and change? Who becomes enlightened?" These questions arose from a misunderstanding of *anātman*. While the Buddha spoke against the psychological and nonrelational reality of an independent self, he never denied the ontological self (9). The *anātman* doctrine describes a relationship whereby any given person or thing derives its being and meaning, not from itself but from its relations with others. Simply, each individual is defined by his or her role in society and is affirmed by his or her interactions with others. Partly as a reaction against its critics and to affirm the great reverence for the "empty" and ontological self, Buddhists postulated such notions as Buddha-nature, *ālayavijñana* (storehouse consciousness), and *tathāgathagarbha* (womb of the *Tathāgata*, one who has touched the shore of *nirvana*) that posit an underlying reality on which the affectations occur. The significance of an ontological self is further emphasized by the observation in the prologue to the *Tri-śarana-gamana* or Threefold Refugees recited daily by Buddhists, that the appearance of an individual in the world is a rare event. The Japanese cleric, Dōgen (1200–1253), and the aesthetician and art critic, Yanagi Sōetsu (1898–1961) expanded the notion of the intrinsic value sentient beings to include things. Dōgen asserted that even inanimate objects are Buddha-nature (10). Yanagi spoke of enlightened things (11).

While the doctrine of *pratītyasamutpāda* describes the Buddhist understanding of reality and is the rationale for karmic interaction, the Four Noble Truths crystallize this doctrine's existential import. According to Buddhist lore, the lesson of the Four Noble Truths is the first the Buddha shared after the enlightenment. It relates directly to the doctrine of *pratītyasamutpāda* as a moral principle based on a reworking of the law of karma. The Four Noble Truths, an empirical–rational methodology that is closely associated with ancient Aryadevic medicine, parallels the steps — diagnosis, etiology, recovery, and therapeutics — that summarize the medical treatment of a disease. The Four Truths profile the condition of our lives, explain the cause of suffering, and the means by which we, residing in a samsaric world, can extract ourselves and realize an abiding spiritual reality. The Four Truths are (*1*) the Noble Truth of Suffering, (*2*) the Noble Truth of the cause of suffering is illusion and desire, (*3*) the Noble Truth of Nirvana, a realm free from suffering, and (*4*) the Truth of the Noble Eightfold Path is the way to enlightenment or nirvana. The Eightfold Path consists of Right View, Right Thought, Right Speech, Right Action, Right Livelihood, Right Effort, Right Mindfulness, and Right Meditation. In the First Truth, the Buddha acknowledges that spiritual suffering, though the most serious, was just one of many ills. The cause or etiology of this suffering, the Second Truth, stems from illusion and desire. Illusion is the belief in a substantial self and an unchanging world, and desire refers to wishing for unattainable things. The Fourth Truth is the Eightfold Noble Path that releases the individual from ignorance and delusion. The Eightfold

Path is the medicine or spiritual therapy that leads to the Third Truth, wisdom or nirvana.

The Four Noble Truths outlines a method to transcend, not escape suffering through understanding. In the *Vissudhimagga*, Buddhaghosa (circa 400 C.E.), the great Theravāda commentator, correlated the four phased method, for spiritual health systematized in the Four Noble Truths, with the treatment of a disease. "The truth of suffering is like a disease, the truth of origin is like the cause of the disease, the truth of cessation is like the cure of the disease, the truth of the path is like the medicine" (12). Modern Buddhists have applied the four-step method outlined in the Four Truths to remedy social and economic problems. The Sarvodaya Shramadana Movement, a rural self-help program initiated by Ahangamage Tutor Ariyaratnes of Sri Lanka, is the most celebrated example. Since the 1960s Ariyaratnes and his method have empowered poor villages to recognize their problems, discern their causes, envision solutions, and devise remedies.

THE BUDDHIST POSTURE

A vision of an interdependent world affirms the reality that we live in a complex, ever-changing web of interrelationships. This vision allows for important implications when thinking about ethical, legal, and policy issues concerning biotechnology. This brief section begins with a sketch of the ideal relationship between the whole and part explicated by *pratttyasamutpāda*, proceeds to discuss some of its inherent difficulties, and concludes with a some thoughts of the virtues of deliberation in the context of an interdependent world.

Ethical, legal, and policy deliberation and exercise that posit an interdependent world consider a microscopic view along side a macroscopic one. The symphony offers an analogy. To grasp the beauty in a Beethoven symphony, for example, it is not enough to listen to the individual instruments sequentially. One must hear the instruments together, as each musician modulates his or her instrument and timing in response to the other instruments. Each instrument participates in creating a greater whole. The whole in turn gives value to the sounds from each instrument. The microscopic and the macroscopic resonate together.

Like a symphonic composition, in the Buddhist vision of the world, we are parts of a larger whole. The whole in turn gives each individual value and worth. The idea of interdependence links our individual lives to each thing and each being in the universe. To act according to this vision is to work to nurture the lives and relationships that enrich and sustain the life of individuals and the whole universe (13). Such an interpretation of life dictates the various virtues and ends that Buddhists should consider when reflecting on ethical questions, legal issues, and formulating policies. A sense of fair play, compassion, gratitude, humility, and patience are some virtues to nurture and consider when reflecting on such questions as health care, allocation of resources, relating to the most vulnerable in our society, the environment, biodiversity, and the manipulation of life.

This harmonious ideal is difficult and almost always impossible to realize. In a world where lives are inexorably intertwined, like ice and water, our concern should be extended to all beings. In a practical sense, our energies should focus where we can make a difference, before extending our energies to embrace all lives and relationships. The Japanese cleric, Shinran (1173–1262) reiterates this attitude in his fourteenth letter of the *Mattōshō*, "the one who first attains nirvana vows without fail to save those who were close to him first and leads those with whom he is karmically bound, his relatives, and his friends" (14). Additionally our life experiences we are often beset by conflicting demands and responsibilities. Our wishes are continually frustrated by the demands of others and by events beyond our control. While some forces nurture our lives, others demean. In the Thirteenth Chapter of the *Tan'nishō*, Shinran underscores the reality that we are often swept up in events that thwart our best intentions. In his conversation with Yuienbō, a fellow devotee, Shinran says, "It is not that you keep from killing because you are good. A person may wish not to harm anyone and yet end up killing a hundred or even a thousand people" (15). Since our lives are intimately linked with the karmic tide of others, to society, and even the whims of nature, we may be propelled to violate our deepest moral instincts. Under those circumstances we yield — mournfully and perhaps, even justifiably — to the dictates of more powerful karmic forces. In such a world the best that we can hope for, according to Thich Nhat Hahn, the Vietnamese monk, is to be determined to go in the direction of compassion and try to reduce suffering to a minimum (16). The exercise of compassionate aspirations, no matter how insignificant, is based on the belief that an act of kindness resonates throughout the farthest reaches of the universe.

A vision of an interdependent world recognizes the complexity of even the most common event. The quiet unfolding of the morning glory is supported by the entire universe. Fa tsang (643–712) articulates the complexity of this singular event in his Ten Subtle Principles of the Unobstructed Fusion of *Pratttyasmutpāda* (Shih-hsüan-yüan-ch'i-mu-ai-fa-mên). He reasoned that in an interdependent world no dharma (thing or event) is independently established and thus all dharmas are mutually supportive and mutually dependent within the *dharmadhātu*, the realm of dharmas; each dharma is thus of equal importance. However, when a dharma is arbitrarily singled out for consideration, that particular dharma becomes the principal dharma and the remaining dharmas take on a secondary role. Each dharma has the potential of alternately assuming the principal role or a secondary role. The role a dharma assumes is determined by what is weighed to be important at any given moment. Moreover each cause and condition offers a different perspective of how a thing or event arises. No thing or event is ever the locus of attention for everyone. The construction of a much needed new bridge, for example, requires the approval of many public agencies and private interests. The necessity of efficient and reliable thoroughfares must balance commercial, environmental, engineering, aesthetic, and other needs

of the community. The environmentalists concern for biodiversity relegates commercial interests to a secondary concern. The engineer is concerned first with structural integrity, rather than aesthetics. Knowing that we live in an interdependent world means that we may never resolve issues to everyone's satisfaction.

Predicaments also arise from competing interests as well as conflicting perspectives. We may never know for certain how a specific event transpired, as Akira Kurosawa (1910–1998) dramatizes in *Rashamon*. In this murder mystery we hear the testimonies of the bandit, the police agent, the woman who was raped, and through a shaman, the murdered husband. Each individual relates a slightly different version of the circumstances that surround the murder. The film ends without a resolution. Often we must deliberate, make decisions, and act knowing we are unable to reconcile or understand completely differing perspectives. Appreciating the complexity of an issue permits us to see many points of view and is perhaps the most productive way of ethical deliberation and action. We nurture humility and patience knowing that others may not approve or follow our example. Living with and appreciating alternative points of view is reminiscent of a Cubist painter who renders an object from different viewpoints. In contrast, perspective, a visual rendering technique perfected during the European Renaissance, renders an object from a single-fixed point.

Acknowledging the validity of other points of view, the Buddha urged his would be followers not to accept any of his teachings without first critically examining them. Only if any of his teachings lead to spiritual ease, should they be observed and accepted (17). Other paths may be more suitable for one's particular temperament. The Buddha insisted that he was a guide, not an authority. The Buddha's critical attitude toward religious authority, even his own, is seen in Thich Nhat Hahn, a Vietnamese monk whose experiences of the Vietnam War forced him to reinterpret and condense the traditional Buddhist precepts into Fourteen Precepts for Engaged Buddhism. The first Precept reads: Do not be idolatrous about our bound to any doctrine, theory, or ideology, even Buddhist ones. Buddhist systems of thought are guiding means; they are not absolute truth. Explaining this precept, Thich Nhat Hanh writes that clinging to our views can cause us to lose the opportunity to a higher and more profound view of reality. By being open to other points of view, we expand the frontiers of our knowledge and our understanding of the world (16).

BUDDHIST MEDICINE AND HEALTH CARE

Health and health care are metaphors common to Buddhist thought and practice. The Buddha, the great physician, dispenses the Dharma, the medicine that heals humanity's suffering and brings spiritual ease. To the terminally ill he administers the teaching of impermanence and to others meditative exercises. Though spiritual suffering was of paramount concern, the Buddha understood that spiritual well-being necessarily involved physical health, which in turn, is dependent on a wholesome community and its sound management of economic and other resources. The Buddha and his devotees attended to their spiritual ills through self-cultivation, and served as nurses to the sick by dispensing medicine and compassionate deeds. While illness testifies to the frailness and transiency of the human condition, caring for the sick is also an opportunity for spiritual quickening. The sick, in turn, have an opportunity for abundant giving. Spiritual health means to realize and to live with gratitude and responsibility to all things and beings. This section begins with an overview of health and health care within the context of the early Buddhist theory of medicine. It then proceeds to discuss caregiving and the relationship between the caregiver and the patient.

Medical Theory

Grounded in the belief of an interdependent world, Buddhist medical theory understands mind and body to be a single unit. Illness of the body is illness of the mind, and mental illness is directly related to the illness of the body. Health requires balance and reciprocity among all the four elements: earth or the solid element, water or the wet element, fire or the hot element, and wind or the mobile element, and the three peccant humors of wind, phlegm, and bile that constitute the human body. Illness arises when one or more of these elements experience an abnormal augmentation or diminution. Medicine and medical therapies provide the means to restore and maintain a healthy physical balance. By contrast, present biomedical diagnostics understand the body to be made up of distinct divisible parts, whose organs and functions can be isolated and treated. Modern etiology seeks the sources of disorder from external pathogens, rather than internal disorders. While Buddhist physicians traced the etiology of disease to empirical causes, karma or past action is also a category of medical etiology. Past deeds relate to present mental suffering caused by greed, hatred, and doubt.

In addition to the disequilibrium among the four elements and the three humors, Buddhist medical theorists understood that external and societal conditions affected the internal working of the body and cause of disease. Diet, daily regimens, alteration of the seasons, stress from unusual physical activities, and past actions affect one's physical and mental well-being. In keeping with the belief that prevention is the best guarantee against illness and disease, the Buddha urged moderation in spiritual exercises and in all life activities. His monastic rules emphasized personal hygiene and public health. Straining water served to purify it and to prevent consuming water-dwelling organisms. Living quarters and privies were to be kept clean. Even today in Zen monasteries in Japan, certain days are set aside for washing and mending.

The Buddha traced much of human illnesses to poverty. The poor, he reasoned, had limited access to material and nonmaterial resources that ensured basic necessities: food, clothing, shelter, medicine, and education (i.e., spiritual development). A Buddhist state would have the responsibility to provide a wholesome living environment by safeguarding the natural environment, by ensuring

the equitable distribution of resources, and protecting the poor and dispossessed against exploitation. These measures would optimize psychological and physical well-being, the basis on which to nurture spiritual development. The Buddha placed a high value on physical fitness and freedom from illness as a basis for mental and spiritual development. "When sentient beings have sick bodies, their minds cannot be at peace ... the bodhisattva who would cultivate awakening should first minister to the illness of the body" (18). This holistic approach to health and healing resonated within Taoist- and Confucian-based Chinese medicine. Both Buddhist and Chinese medicine do not distinguish between mind and body. In addition to the yin and yang theory which is based on harmony of the universe, the Chinese included the connections between a well-functioning social system and a healthy body.

Caregiving

Medicine and caregiving were integral parts of the early Buddhist community. The model for caregiving is the Buddha himself. On an occasion, the Buddha chanced on an unattended sick monk wallowing in his own excrement. He said, "O monks, if you do not nurse one another, whoever will nurse you?" Thereupon the Buddha bathed the monk, changed his garments, and laid a bed for his ailing comrade. This experience led the Buddha to declare, "Anyone who wishes to make offerings to me, let him make offerings to the sick." Initially the early community focused medical treatment and nursing activity on the care of monks and nuns by fellow cenobites or by pious lay devotees. Caring for their compatriots became part of the monastic code. From around the mid-third century B.C.E. medical care was extended to the population at large. Later the medical arts became part of the curriculum at Nalanda and other monastic universities.

Buddhist documents abound with medical injunctions and prohibitions. The *Aṅguttara Nikāya* lists the five qualities of those who tend to the sick. It also lists the five faults of a patient that impede his or her recovery. The competent caregiver (*1*) possesses knowledge of medicaments and their application, (*2*) tends to the sick with amity of mind and without thought of personal nourishment and profit, (*3*) is not lazy, (*4*) or prone to annoyances and does not loathe removing excrement, urine, sweat, or vomit, and (*5*) since the early Buddhists linked illness with mental states gone awry, the competent caregiver should delight in sharing the Buddha's Dharma and conversing with the sick (19). Conversing requires the art of listening which is personified in Avalokiteśvara, the Bodhisattva of Compassion. "Avalokiteśvara" or "Kanzeon," the Japanese rendering, clearly captures the Bodhisattva's special talent. "Kan" means "to hear," "ze" means "world," and "on" means "sound." Avalokiteśvara who hears the pleas of the world, is the ideal caregiver. Listening to the outbursts of anger, despair, and hurt, the Bodhisattva attends to the journeying spirit, not the momentary stammer. When we engage in conversation and allow another to speak, we act as a midwife who helps to bring new life into the world. By listening, we allow for self-discovery that gives birth to a new being.

Patient Responsibility

The vision of an interdependent world requires a patient's involvement in his or her care. The *Aṅguttara Nikāya* lists five faults of the patient that discourage healing. The patient impedes his or her healing (*1*) by not being selective of what he or she eats or drink, (*2*) fails to take nourishment at the proper times, (*3*) refuses to take medicaments, (*4*) abandons him or herself to melancholy, merriment, and annoyance, and (*5*) is pitiless toward the sick nurse (20). In addition to cooperating with nurses to hasten health, one can use his or her illness as the occasion for spiritual exploration and abundant giving. Vimalakīrti, the most famous invalid in the Buddhist canon, describes the responsibility of the sick:

> Through one's own experience, a Bodhisattva should have sympathy with the sick person. Let them know of the pains they suffered from the infinite past, but encourage them by advising them to endeavor and become the Buddha, the Great King of Medicine and cure the illness of all people ... (21).
>
> Though the body may be in the world of delusion and diseased, if one gives abundantly and tirelessly, that is called expedient means. The body may not be freed from the disease and the disease may not be freed from the body, but if the disease and body are seen neither as new nor old, that is called wisdom. Though the body may be diseased, if one does not forsake this world and does not intend to enter Nirvana, that is called expedient means (22).

Should we choose to heed Vimalakīrti's injunction, illness is an occasion for service and spiritual quickening. Through their experience of illness, caregivers can sympathize with others who are ill. They can urge the sick not to succumb to their pains, but encourage them to relieve the suffering of all beings. Even if one is ill, one can give abundantly. "Expedient means," a rendering of *upāya*, ordinarily refers to the wisdom to convey the Dharma according to the needs and capacities of the listener. But here, Vimalakīrti defines *upāya* to mean the efforts the devotees of the Dharma should expend to relieve suffering.

Abundant giving takes on many forms. The invalid and sick often inspire and instill faith in the human experience to those who attend to them. "*Dana*" or selfless giving is the first of the Six Paramitas or Perfections a Bodhisattva observes. The selfless gift is "*Dana* ... most profound among all joys ... that which is found through witnessing and experiencing the joy of others [so that] ... the joy of others becomes one's own joy" (23).

In an interdependent and ever-changing world, health and healing are possible with the advent of new knowledge and with new relationships. Change assures that illusion can be transformed into enlightenment, illness can be cured, and social decay arrested.

MANIPULATION OF LIFE

The discussion of the manipulation of life includes the environment and attended concerns of ecology, biodiversity, and agriculture, and modifications of living organisms that include human beings, microorganism, and animals. While Buddhist documents offer insight into

dealing with the environment, we can only extrapolate what might be the Buddha's attitude toward the manipulation of life, organ transplants, and cloning.

Environmental Issues

Buddhist responses to ecological issues, biodiversity, and agricultural development are rooted in the sense of responsibility and gratitude, intrinsic to the doctrine of *pratītyasmutpāda*. As it was noted above, for the Buddha, personal health is established within the context of a wholesome society and environment. The *Abhidharmakośa śāstra*, an influential primer composed by Vasubandhu (circa 400 C.E.), defines the world to include both sentient beings and the container world, the realm that supports sentient life (24). Just as mind and body are considered to be a single unit, Buddhism understands the individual, society, and the natural world to constitute a single whole. The idea of "one is all and all is one" articulated in the *Avataṃsaka sūtra*, and the elaboration of this idea in such doctrines as the Ten Subtle Principles of Unobstructed Fusion of *Pratītyasmutpāda* and the Six Principles of the Causal Aspects of *Pratītyasmutpāda* (Yüan-ch'i-yin-mên-lu-i-fa) by Fa-tsang and other Hua-yen masters, articulate the ideological rationale for environmental concerns. These doctrines explain that human, plants, animals, and material entities do not simply exist in and by themselves. Their individual and separate existences are affirmed by and made possible through their relationship with others (25). Our individual well-being is dependent on the health of the world we live in. Our activities should thus be conducted with a sense of respect and reverence for all life. Conservation, harmonious coexistence, the need to care for and to restore the land (26), not exploitation, should be the hallmark of a Buddhist devotee. The destruction of the Amazon rain forests affects my well-being and the very life of the world. The Buddhist devotee should quicken feelings of gratitude to animals and plant life to whom he or she must depend on for life. Donald Swearer sums the modern Thai monk Buddhadasa's (1906–1993) rationale toward the environment:

> One cares for the forest because one empathizes with the forest, just as one cares for people ... [Empathy] is fundamentally linked with non-attachment or liberation from preoccupation with self, which is so central to Buddhasada's thought ... Caring in this deeper sense ... goes beyond the well-publicized strategies of the conservation monks to protect and conserve the forests ... [Empathy] translates as having at the very core of one's being the quality of caring for all things in the world and their natural conditions; that is to say, caring for them as they are in themselves rather than as I might benefit from them or as I might like them to be (27).

The modern environmental ethic that advocates minimum exploitation of and optimum utilization of natural resources and maximizes recycling finds much in common with the Buddhist idea of *pratītyasmutpāda*. Modern Buddhists actively work to preserve the environment and promote biodiversity. The Buddhist Peace Fellowship founded in 1978 under the leadership of Robert Aitken and others in Hawaii have galvanized Buddhists worldwide to work to protect the environment and issues of justice. In November 1997 villagers, Buddhist monks, and environmentalists gathered in Sai Yok National Park to protest the Thai government's decision to allow a pipeline to run through the park. To protest the deforestation and its animal inhabitants, and displacement of people, Buddhist monks ordained the trees (28). The voices of Buddhists thinkers and activists appeared in *Buddhism and Ecology, the Interconnection of Dharma and Deeds*, a collection of papers presented at the 1996 proceedings of Earth Charter, a project that set forth a vision of ethical principles for the twentyfirst century (29).

In the Buddhist countries of Asia, modernized agriculture resulted in increased yields, but it also brought about the large-scale depletion of natural resources. Deforestation for agricultural uses has destroyed much of the wild life, and use of chemical fertilizers and insecticides has killed off mudfishes and edible frogs that once thrived in the rice fields and served as a rich source of food (30). In response to the "destruction of human communities and nature in the name of globalization, of multinational corporations, governments, and local allies" more than 400 social activists and leaders from the Asia–Pacific region gathered in Kathmandu in 1996. The international gathering, "People's Convergence, Shaping Our Future" was the third event of the People's Plan for the twenty-first Century that began in 1989 in Japan to bring attention to environmental pollution. The Minamata Declaration called for a grassroots transborder participatory democracy to change the global structure (31). These grassroots nongovernmental organizations, including Buddhists and Buddhist organizations, attempt to live out the Buddha's insistence that material and environmental needs of the people be respected.

Brain Death and Organ Transplants

The question of brain death and the appropriateness of organ transplants generated great concern and a range of opinions in the Buddhist community. The controversy lies in part in the meaning of life and death, personal identity, and the belief in the inseparability of mind and body. In the United States and other countries where transplants are routine, death is defined as the absence of brain activity, which often occurs before a heart stops beating. Legally defining a patient, whose brain has ceased to function, dead is crucial in harvesting organs. Organs quickly deteriorate once the heart stops beating. Brain death is not satisfactory for Buddhists who subscribe to the traditional cardiopulmonary definition of death. More substantially, others object to the brain-death criterion of death because Buddhists have always associated life with sentience (32), which in its broadest sense includes feeling. Though the brain may have ceased to function, the individual with a beating heart may be pained by being cut, and having his or her organs removed. Doctrinally, death is defined as the dissolution of mind and body. Death dissolves the fortuitous interactions of karmic events that gave birth to and nurtured an individual. The separation of the mind from the body, however, is not death of the person, as we will presently see. Curiously, even with death defined as cessation of higher-brain functions, we commonly begin funeral preparations when the heart has

been removed or has stopped beating, and not when the brain is dead. The medical and legal definitions of brain death often conflict with social notions of death.

Another pervasive attitude against organ transplants is the assumption that life is impermanent. Since life is transient and death inevitable, there is no meaning to artificially extend life by receiving the organs of another. The extension of life by organ transplant disrupts the natural karmic life span. Rather than extending life through heroic measures, humane end-of-life care would be more in keeping with the spirit of the Dharma. Further, organ transplants are possible only at the expense of another's life, a violation of the precept to abstain from taking life. Consequently some Buddhists advocate the development and use of artificial organs. However, those who favor transplants argue that the gift of life is the greatest gift an individual can give. The body is, after all, transient and ultimately worthless (33). Buddhist lore is replete with legends that relate the sacrifice of limb and life by the Buddha. In Sri Lanka, Hudson Silva has used a legend of the Buddha with great success to persuade people to donate their eyes for corneal transplants (34). The legend even makes mention of an eye transplant.

The Buddhist misgiving toward defining death as the cessation of brain activity may be another reason for the ambivalence of the Japanese toward organ, especially heart transplants. After more than 30 years of debate, the Japanese Parliament on June 17, 1997, passed a law that allows a person, whose brain has stopped functioning, to be defined as dead; in cases where a patient has agreed to this definition of death, he or she can request that his or her heart and lungs be donated for transplants. The bill does not provide a legal definition of death; it does nonetheless allow the brain-death standard to be used for donors of hearts and lungs. Current Japanese law defines death as the moment the heart stops beating. The bill does not give the donor an absolute right to ask physicians to decide whether he or she is considered brain-dead. The donor's family has the ultimate right to veto the doctors' diagnosis of brain death and the patient's wishes. A patient's rights, as noted above, are not absolute. Ironically the new Japanese law allows for a greater measure of individual autonomy. Physicians are given permission by the prospective donor to declare himself or herself brain-dead for the purposes of organ donation. In contrast, in the United States brain death is defined by law, and physicians do not need the patient's permission to declare an individual brain-dead.

On February 28, 1999, 21 months after the approval of the new law, Japanese doctors performed their first heart transplant since Dr. Juro Wada attempted a heart transplant 32 years ago. Dr. Wada's patient died and his operation and motives are clouded with legal controversy.

The difference in the importance of individual autonomy differentiates the U.S. and Japanese approach to organ transplants. As we noted, in Japan an individual does not have an absolute right to self-determination; the Japanese Parliament allowed the family to void any prior directive a brain-dead individual may have made concerning the disposition of his organs. The Japanese approach to brain death reflects the manner in which a Buddhist would approach the problem. In an interdependent world where lives are intertwined with countless others, individuals do not have exclusive claim on their lives. We may have separate lives, but we live in resonance with others.

The hesitation of organ transplants among East Asian Buddhists can also be traced in part to the Chinese Confucian notion of filial piety. The opening lines of the *Hsiao Ching*, or Classic on Filial Piety, states, "Filial piety is the basis of virtue and the source of our teachings. We receive our body, our hair, and skin from our parents, and we dare not destroy them" (35). When Buddhism first entered China, the Chinese appealed to this passage to argue against their sons and daughters shaving their hair when entering the Buddhist order. A person should be buried with every part of his or her body. The donation of one's organs would thus constitute a most unfilial act. This attitude has prevented wide acceptance of organ transplants. While the Chinese value keeping the body intact after death, there is a countervailing attitude that the use of organs from executed prisoners can benefit social and public good (36). This attitude toward the asocial elements has its roots in imperial China when social order and individual health were closely linked (37). Korean Buddhists' wariness of organ transplants stems from a strong Confucian imprint. The indigenous shamanic belief that a person who is not buried with all of his or her body will suffer in the next phase of life also contributes to their hesitation.

The importance of family lineage and the reciprocity between the living and dead account for the reluctance of organ transplants among the Chinese, Koreans, and Japanese. In traditional East Asia, death of the physical body is not the death of the person. Incorporating this belief, Buddhist mortuary rites mark the transformation of the person from a physical to a spiritual being. The person matures or proceeds to ancestorhood with the aid of memorial observances sponsored by the living descendants. In return, the ancestor ensures health and prosperity for the family. This accounts for the complex and lengthy memorial cycle. The Japanese Buddhist mortuary rites are especially long. The memorial cycle begins immediately after death and continues for at least 33 years. The 49th day, 100th day, 1st year, and the 3rd, 7th, 13th, 17th, 25th, 33rd year observances are especially important. On the island of Okinawa, a living repository of Japanese culture and language, the 33rd year memorial service marks the complete transition of the individual to an ancestral spirit. After the completion of the service, the individual's memorial tablet is burned. Services are no longer dedicated to the memory of the deceased and the individual is honored collectively as an ancestor with all other ancestors. While the long memorial cycle ritualistically marks the transformation of a person's identity, it in fact reveals something of the nature of our memories. As years pass, our recollections of the deceased become less and less distinct and he or she gradually loses his or her individuality. Korean Buddhist mortuary rituals continue for up until three years, and thereafter the deceased is honored at an annual memorial service for ancestors. Thai and Burmese Buddhists' memorial rituals are seven years (38). In accordance with Confucian sentiments, Chinese Buddhists mourn for three years.

Cloning

In late February 1997, when Ian Wilmut, a scientific researcher, announced the first successful cloning of a sheep named "Dolly," the prospect of cloning a human being prompted celebration, caution, and concern. Buddhists have not raised objections to this scientific breakthrough, but they are concerned with the ends of and motivations for human cloning. The creation of new life should not be seen as a product but an end or value in itself. Such reverence appeals to the idea of Buddha-nature and that the appearance of a clone is a rare event. Cloning a human being to produce organs for use in transplantation, however, would be repugnant (39). In contrast, some appeal to the principles inherent in the doctrine of *pratītyasmutpāda* to celebrate the reality of change. The technique of cloning is a tool that further expands human scientific and technological promises (40) and new moral possibilities (41). Since change is the nature of reality, the present challenge of the cloning question is how to accommodate change, expand our notions of humanity, and our moral parameters.

Organ transplants and genetic manipulation raise questions of family continuity in those East Asian cultures where family lineage is valued. If a person is considered to be the unique repository of prior generations, receiving an organ from another person raises the question of identity. If genes are manipulated what is the relationship between ancestor and descendent?

The manipulation of life through genetic engineering, like other human-generated innovations, is consistent with the Buddhist belief in change. Countless causes and conditions propel us to this present moment, providing for new achievements. Our present thinking and activities interact with current concerns and with the natural order. As active participants and an integral part of the process of interdependence in the life of the world, human beings have the capacity to affect the subsequent course of events. This should give us pause to reflect on our responsibilities and present and future action. Often we are unaware of the consequences of our achievements. For example, an August 1998 article in *Nature* reports that climatologists have linked pollution emitted by factories and automobiles with rainy weekends (42).

CONCLUDING REMARKS

Buddhists, caught up in rapid scientific and technological changes, have been slow to reflect on the ethical, legal, and policy issues generated by recent advances in biotechnology. Since much of these changes occurred in the United States and Europe, technically the most advanced countries, and where the issues surrounding modern biotechnology initially appeared, Christian theologians have reflected on and activists voiced concerns over these advances. By contrast, Buddhism, which is dominant in the countries of South and East Asia where biotechnology has begun to have an impact, has only recently confronted these problems. Japan is the most notable exception. Since the successful cloning of twin calves in July 1998 by Japanese scientists under the direction of Yukio Tsunoda of Kinki University, the Japanese Ministry of Agriculture and Forestry reported that as of March 31, 1999, 57 calves have been cloned from somatic cells and another 461 from nuclear embryo transplantation (43). Buddhist thinkers in that country have been reflecting on cloning.

The notion of *pratītyasmutpāda* holds the key to the Buddhist approach to implications of biotechnological advances. *Pratītyasmutpāda* offers an understanding of change, humanity's place in the process of change, and a vision of human responsibility to all things and all beings and to the world. Nothing in the Buddhist documents suggests halting changes that new knowledge generates. Change is a cardinal Buddhist presupposition. While change may be the opportunity for expanding Buddhists' moral imagination, the idea that all things and all beings are mutually and irrevocably interdependent instills a sense of humility that is necessary for ensuring that all species and all things are accorded respect. Further the vision of an interdependent world quickens concerns for the safety of food from cloned animals and plants, and the long-term consequences of gene manipulation on the environment and all sentient life. These and other ramifications of biotechnological advances and policy decisions must carefully consider all aspects of suffering that change generates. The karmic energies of a single individual have wide repercussions. "A wise man should do things that are beneficial to living-beings" (44).

BIBLIOGRAPHY

1. The most widely accepted dates, 563–483 B.C.E., are based on the Sri Lankan historical chronicles, *Dīpavaṃsa* and *Mahāvaṃsa*. Hakuju Ui (1882–1963), a prominent Japanese scholar, disputed these dates based on Northern or Mahāyāna sources. Thus, according to Ui, the Buddha's dates are 466–386 B.C.E.
2. Followers of Mahāyāna (great vehicle) Buddhism coined the expression, "Hīnayāna" (lesser vehicle) as a pejorative for those who did not accept their documents and their doctrines. Mahāyānists accused Hīnayānists and their *arhat* ideal that was reserved for the select few. The Mahāyāna ferries all beings across the sea of samsara to nirvana, while the Hīnayāna transports only a few. No Buddhist group referred to itself as Hīnayāna. The devotees of Theravāda, a non-Mahāyāna tradition, object to being labeled "Hīnayānaists."
3. K. Mizuno et al., *Butten kaidai jiten [A Bibliographical Dictionary of Buddhist Texts]*, 2nd ed., Shunshu sha, Tokyo, 1978, pp. 15–16.
4. S. Kamata, proposed alternative routes and dates for the eastward advance of Buddhism in L. Lancaster and C.S. Yu, eds., *Introduction of Buddhism to Korea, New Cultural Patterns*, Asia Humanities Press, Berkeley, CA, 1989. See *The Transmission of Paekche Buddhism to Japan*, pp. 143–160.
5. R. Fukuhara, *Bukkyōgairon [Outline of Buddhism]*, Nagata shodo, Kyoto, Japan, 1995, p. 145.
6. K.N.S. Subramanian, *Hastings Cent. Rep.* **16**(8), 20–22 (1986).
7. P. Ratanakul, *Bridges* **3**, 4–5 (1997).
8. A. Hirakawa, *A History of Indian Buddhism, from Śākyamuni to Early Mahāyāna*, University of Hawaii Press, Honolulu, 1990, p. 6.

9. H. Nakamura, *Jiga to muga [Self and Non-self]*, Heirakuji, Kyoto, 1976, pp. 55–59, also J. Pérez-Remón, *Self and Non-Self in Early Buddhism*, Mouton, The Hague, Japan, 1980, pp. 301–305.
10. Dōgen, *"Busshō" [Buddha-nature] in Shōbōgenzō*, Iwanami, Tokyo, 1970, p. 49.
11. S. Yanagi, *The Eastern Buddhist, New Ser.* **13**(2), 16 (1978).
12. Buddhaghosa, *Visuddhimagga*, R. Semage, Colombo, Sri Lanka, 1956, p. 586 (trans. by Bhikkhu Nanamoli as *The Path of Purification*).
13. R. Nakasone, in Watanabe Takaō kyōjū kanreki kinen ronshū, ed., *Bukkyōshisō bunkashi ronsō [Essays on Buddhist Thought and Cultural History]*, Nagata shodo, Kyoto, Japan, 1997, pp. 302–330.
14. Shinran, *Mattōshō*, Nagata shodo, Kyoto, 1978, p. 47 (trans. by Yoshifumi Ueda as *Letters of Shinran, a translation of Mattōshō*).
15. Shinran, *Tan'nisho*, Ryukoku University, Kyoto, Japan, 1982, pp. 32–35 (trans. by D. Hirota as *Tan'nisho, a Primer*).
16. T. Nh'ât Hahn, *Interbeing, Commentaries on the Tiep Hien Precepts*, Parallax Press, Berkeley, CA, 1987, p. 27.
17. Saṅghabhadra, *Samantapāsādikā*, Bhandarkar Oriental Research Institute, Poona, India, 1970, pp. 171–172 (trans. by P.V. Bapat and A. Hirakawa).
18. J. Takakusu, ed., *Taishō Shinshū Daizōkyō hangyōkai, Tokyo* **10**(293), 710c–711a (1925–1931).
19. J. Takakusu, ed., *Taishō Shinshū Daizōkyō hangyōkai, Tokyo* **2**(125), 680c (1925–1931).
20. J. Takakusu, ed., *Taishō Shinshū Daizōkyō hangyōkai, Tokyo* **2**(125), 680b (1925–1931).
21. J. Takakusu, ed., *Taishō Shinshū Daizōkyō hangyōkai, Tokyo* **14**(475), 544c (1925–1931) (English trans. in *Buddha-Dharma*, Numata Center for Translation and Research, Berkeley, CA, 1984, pp. 319–320).
22. J. Takakusu, ed., *Taishō Shinshū Daizōkyō hangyōkai, Tokyo* **14**(475), 545b (1925–1931) (English trans. in *Buddha-Dharma*, Numata Center for Translation and Research, 1984, p. 320).
23. S. Makino, *The White Way* **34**(5), 3 (1990), quoted by J.C. Takamura, *Generations*, Fall/Winter (1991).
24. Vasubandhu, in J. Takakusu, ed., *Taishō Shinshō Daizōkyō hangyōkai, Tokyo* **30**(555), 57a (1925–1931).
25. Y. Matsunaga, *Bukkyō* [Buddhism] **6**, 121–130 (1991).
26. J. Takakusu, ed., *Taishō Shinshō Daizōkyō hangyōkai, Tokyo* **24**(1488), 1061a (1925–1931).
27. D.K. Swearer, in S. Sivaraksa, ed., *The Quest for a New Society*, Santi Pracha Dhamma Institute, Bangkok, Thailand, 1994, p. 7; quoted by R. Aitken in *Turning Wheel*, Summer, 1998, pp. 28–30.
28. *Turning Wheel* Winter, 7, 8 (1998).
29. M.E. Tucker and D.R. Williams, eds., *Buddhism and Ecology, the Interconnection of Dharma and Deeds*, Harvard University Center for the Study of World Religions, Cambridge, MA, 1997.
30. R. Bobilin, *Revolution from Below, Buddhist and Christian Movements for Justice in Asia*, University Press of America, Lanham, MD, 1988, p. 119.
31. I. Muto, *ANPO: Jpn.-Asia Q. Rev.* **27**(2), 40–45 (1996).
32. *Samantapāsādikā*, Bhandarkar Oriental Research Institute, 1970, p. 319 (trans. by P.V. Bapat and A. Hirakawa).
33. H. Masaki, *Seimei rinri* [Bioethics] **1**, 48–52 (1993).
34. M.V. van Andel, *Doc. Ophthalmol.* **74**(2), 141–150 (1990).
35. *Hsiao Ching*, St. John's University Press, New York, 1975 (trans. by Mary Lelia Makra as *The Hsiao Ching*, p. 4).
36. R. Qui and D. Jin, in B.A. Lustig, ed., *Bioethics Yearbook, Regional Developments in Bioethics: 1991–1993*, Kluwer Academic, Dordrecht, The Netherlands, 1995, pp. 339–365.
37. P.U. Unschuld, *Medical Ethics in Imperial China, a Study in Historical Anthropology*, University of California Press, Berkeley, CA, 1979, p. 27.
38. E. Namihira, in Kazumasa Hoshino, ed., *Japanese and Western Bioethics, Studies in Moral Diversity*, Kluwer Academic, Dordrecht, The Netherlands, 1997, pp. 61–69.
39. D. Keown, *Reflections*, Spec. Ed., May, 1997, p. 8.
40. Z. Kumatani, *Bukkyō* **42**(January), 94–106 (1998).
41. R. Nakasone, *Reflections*, Spec. Ed., May 1997, pp. 6–7.
42. R.S. Cerveny and R.C. Balling, *Nature (London)*, August 6, 561–563 (1998).
43. *Asahi Shinbun*, April 23 (1999).
44. *Samantapāsādikā*, Bhandarkar Oriental Research Institute, 1970, p. 331.

See other RELIGIOUS VIEWS ON BIOTECHNOLOGY entries.

RELIGIOUS VIEWS ON BIOTECHNOLOGY, JEWISH

ELLIOT N. DORFF
University of Judaism
Bel Air, California

OUTLINE

INTRODUCTION

Judaism, a religion tracing its roots to Abraham close to 4000 years ago and continuing through the Bible and rabbinic interpretations to our own day, has sought since its inception to use the world productively while yet preserving it, both seen as God's commands. This article describes

the theological foundations for Judaism's activist, and yet respectful, stance toward the world. It then describes how this stance is articulated in issues at the beginning and end of life and in environmental matters.

OVERALL CONTEXT OF JEWISH BELIEFS AND PRACTICES RELEVANT TO TECHNOLOGY

Adam and Eve are told in the Garden of Eden "to work it and to preserve it" (Genesis 2:15). Judaism has ever since tried to strike a *balance* between using the world for human purposes while still safeguarding and sustaining it. We are not supposed to desist from changing the world altogether: "Six days shall you do your work" is as much a commandment as "and on the seventh day you shall rest [literally, desist]" (Exodus 23:12).

In changing the world to accomplish our ends, though, we must take care to preserve the environment, whether we are practicing medicine, farming, traveling, or doing anything else. This balance is demanded because, in the end, we do not own the world; God does (1). We are but tenants in God's world, with a lease on life and on the world.

During the duration of that lease, we may and should act as God's agents to improve it. God, in fact, intended that we function in that way. This is probably most starkly stated in a rabbinic comment about, of all things, circumcision. If God wanted all Jewish boys circumcised, the rabbis ask, why did He not create them that way? The answer, according to the rabbis, is that God deliberately created the world in need of fixing so that human beings would have a divinely ordained task in life, thus giving human life purpose and meaning (2). We are, then, not only permitted but mandated to find ways to bend God's world to our purposes — as long, again, as we preserve God's world in the process.

Thus technology, in and of itself, is not good or bad: it depends on how we use it. If we employ it to assist us in bending the world to our ends while yet preserving the world, our use of technology is theologically approved and morally good; if we disregard our duty to preserve the world when using technological tools, we are engaged in a theologically and morally bad act.

FUNDAMENTAL BELIEFS RELATING TO HEALTH CARE

Three underlying principles regarding Judaism's positions on issues in health care emerge from Jewish sources:

1. *The body belongs to God*. Since God owns everything in the world (3), our bodies do not belong to us. Rather, God loans our bodies to us for the duration of our lives, and they are returned to God when we die.

The immediate implication of this principle is that neither men nor women have the right to govern their bodies as they will. Since God created our bodies and owns them, God can and does assert the right to govern the care and use of our bodies. Thus Jewish law requires us to safeguard our health and life (4), and, conversely, to avoid danger and injury (5). So, for example, Conservative, Reform, and some Orthodox authorities have prohibited smoking as an unacceptable risk to our God-owned bodies (6). Ultimately human beings do not, according to Judaism, have the right to dispose of their bodies at will (i.e., commit suicide), for that would be a total obliteration of that which does not belong to them but rather belongs to God (7).

2. *The body is morally neutral and potentially good*. For Judaism the body is as much the creation of God as the mind, the will, and the emotions are. Its energies, like those of our other faculties, are morally neutral, but they can and should be used for divine purposes as defined by Jewish law and tradition. Within that structure, the body's pleasures are God-given and are not to be shunned, for that would be an act of ingratitude toward our Creator (8). The body, in other words, can and should give us pleasure to the extent that that fits within its overriding purpose of enabling us to live a life of holiness.

The Jewish mode for attaining holiness is to use all of our faculties, including our bodily energies, to perform God's commandments. Maimonides states this well:

> He who regulates his life in accordance with the laws of medicine with the sole motive of maintaining a sound and vigorous physique and begetting children to do his work and labor for his benefit is not following the right course. A man should aim to maintain physical health and vigor in order that his soul may be upright, in a condition to know God. . . . Whoever throughout his life follows this course will be continually serving God, even while engaged in business and even during cohabitation, because his purpose in all that he does will be to satisfy his needs so as to have a sound body with which to serve God. Even when he sleeps and seeks repose to calm his mind and rest his body so as not to fall sick and be incapacitated from serving God, his sleep is service of the Almighty (9).

The medical and technological implications of this are clear. Jews have the obligation to maintain health not only to care for God's property but also so that they can accomplish their purpose in life, namely to live a life of holiness. Moreover, since pain is not perceived as a method of attaining holiness but is rather an impediment to acting according to God's law, it is our duty to relieve it. Thus perhaps the most pervasive corollary of Judaism's insistence on the divine source of our bodies is its positive attitude toward the body and medicine.

3. *Human beings are not only permitted but obliged to try to heal*. God's ownership of our bodies is also behind our obligation to help other people escape sickness, injury, and death (10). God is our ultimate healer, as the Bible asserts in many places (11), but God both authorizes us and commands us to aid in that process (12). In fact the duty of saving a life (*pikkuah nefesh*) takes precedence over all but three of the commandments in the Torah (13).

The Talmud reflects some ambivalence about the level of expertise of physicians of its time (most explicitly

in comments like "The best of physicians deserves to go to Hell!"), and some later Jewish authorities were particularly wary of physicians' abilities to practice internal medicine (in contrast to surgery and healing external wounds and diseases). In the end, though, the Talmud prohibits Jews from living in a community in which there is no physician (14). Here this third principle wraps back into the first, for if we were not within easy reach of a doctor, we could not as effectively carry out our fiduciary obligation to God to take care of our bodies.

Medical experts, in turn, have special obligations because of their expertise. Thus Rabbi Joseph Caro (1488–1575), the author of one of the most important Jewish codes, says this:

> The Torah gave permission to the physician to heal; moreover, this is a religious precept and is included in the category of saving life, and if the physician withholds his services, it is considered as shedding blood (15).

The following rabbinic story indicates that the rabbis recognized the theological issue involved in medical care and in the use of technology generally, but it also indicates the clear assertion of the Jewish tradition that the use of technology to assist in good purposes like producing food and preserving health is legitimate and, in fact, obligatory:

> It once happened that Rabbi Ishmael and Rabbi Akiva were strolling in the streets of Jerusalem accompanied by another person. They were met by a sick person. He said to them, "My masters, tell me by what means I may be cured." They told him, "Do thus and so until you are cured." The sick man asked them, "And who afflicted me?" They replied, "The Holy One, blessed be He." The sick man responded, "You have entered into a matter which does not pertain to you. God has afflicted, and you seek to cure! Are you not transgressing His will?"
>
> Rabbi Akiva and Rabbi Ishmael asked him, "What is your occupation?" The sick man answered, "I am a tiller of the soil, and here is the sickle in my hand." They asked him, "Who created the vineyard?" "The Holy One, blessed be He," he answered. Rabbi Akiva and Rabbi Ishmael said to him, "You enter into a matter which does not pertain to you! God created the vineyard, and you cut fruits from it."
>
> He said to them, "Do you not see the sickle in my hand? If I did not plow, sow, fertilize, and weed, nothing would sprout."
>
> Rabbi Akiva and Rabbi Ishmael said to him, "Foolish man!... Just as if one does not weed, fertilize, and plow, the trees will not produce fruit, and if fruit is produced but is not watered or fertilized, it will not live but die, so with regard to the body. Drugs and medicaments are the fertilizer, and the physician is the tiller of the soil (16).

The rabbis quite explicitly, then, understand God to depend upon us to aid in the process of healing. We are, in the talmudic phrase, God's partners in the ongoing act of creation (17).

TECHNOLOGY AFFECTING THE BEGINNING OF LIFE

Underlying Principles Regarding Family, Sexuality, and Procreation (18)

Marriage and children are the epitome of blessing in the Jewish view. "It is not good for man to live alone," the Torah declares, and so one goal of marriage is companionship, sexual and otherwise (19).

The second goal of marriage is procreation. Children figure prominently in the Bible's descriptions of life's chief goods (20), and so God's blessings of the Patriarchs promise numerous children (21). Procreation is not only a blessing; it is a commandment. Indeed, the very first commandment in the Bible is "Be fruitful and multiply" (Genesis 1:28). In rabbinic interpretation, for exegetical and probably economic reasons, it is the man who bears the responsibility to propagate, even though men obviously cannot do so without women. A man, then, fulfills the obligation to propagate when he fathers two children, and since we are supposed to model ourselves after God, the ideal is to have both a boy and a girl, thus creating both male and female, just as God did (Genesis 1:27) (22).

The family is important in Judaism not only because it is in that context that adults gain sexual fulfillment and the next generation is produced; it is also important because it is in the family that the tradition is passed on. Parents have a biblical obligation to teach the tradition to their children (23), and even after schools were established in the first century, parents remained ultimately responsible for the education of their children.

Preventing Conception

Contraception. With the importance of marriage and children in mind, one can understand that traditional Judaism looked askance at interruptions in the process of conception and birth. Normally one was supposed to marry and have children. Birth control, sterilization, and abortion were, both physically and ideologically, counterproductive.

Until very recently the use of birth control or even abortion for family planning purposes, so common in our day, was simply unknown to the tradition. Methods of birth control — either a cloth inserted in the vaginal cavity or a "cup of roots" taken orally — were unreliable, and abortion posed a major threat to the life of the woman. Moreover, if a couple wanted to have two or three children survive to adulthood, they had to produce six or seven. We must keep in mind this major distinction in context and purpose, then, when we examine and evaluate traditional Jewish sources on methods of preventing conception.

The rabbis state that the methods of contraception they had are permitted and even required under certain circumstances. Because the tradition understands the command to propagate to be the obligation of the male, male forms of contraception are generally forbidden. The specific conditions under which female contraception is permitted (and, in some cases, even required) depend on one's interpretation of a second-century rabbinic ruling describing three classes of women who "use" contraceptives — namely a minor (less than 12 years of age), a pregnant woman, and a nursing woman. The present tense of the verb is ambiguous in Hebrew, as it is in English. If it means that these women *must* use contraceptives to protect their life or health or that of their nursing infant, women in other circumstances then *may* use contraceptives. On the other hand, if these three categories of women may use contraceptives only to preserve life and health, then when that

is not a factor, women *may not* use contraceptives (24). In any case, because Judaism restricts the legitimacy of abortion to cases where the life or health of the mother is at stake, modern forms of contraception that prevent conception in the first place (e.g., the pill, the diaphragm) are preferred over those that abort the fertilized egg cell after the fact (e.g., RU486).

Sterilization. The same concerns govern the issue of sterilization, but there another issue arises, namely the prohibition against a person mutilating his or her body in light of the fact that the body is really God's property. Although the procedures are rather new, there are a few rabbinic rulings (responsa) available on the issues of vasectomies and tubal ligations. Both traditional and liberal respondents forbid male sterilization on the basis of the rabbinic interpretation and extension of Deuteronomy 23:2 ("No one whose testes are crushed ... shall be admitted into the congregation of the Lord") (25), or Leviticus 22:24 ("That which is mauled or crushed or torn or cut you shall not offer unto the Lord; nor shall you do this in your land") (26). They are more permissive about female sterilization, both because a woman does not come under those prohibitions and also because she is not legally obligated to procreate (27).

All sources agree, however, that even male sterilization is permitted and perhaps even required if the man's life or health makes it necessary as, for example, if he contracts testicular cancer. Moreover, even though I am not aware of any written opinion that would allow a vasectomy, I could imagine an argument consistent with Jewish law and principles that would permit a vasectomy when pregnancy would entail a severe risk to the man's wife. After all, that procedure is far easier and safer than tying a woman's tubes, and saving a person's life takes precedence over both the commandment to procreate and the prohibition of injuring oneself. Moreover, a vasectomy does not amount to castration or to crushing the testes, and so the biblical verses cited above are not directly violated by the operation. The question, though, would be whether pregnancy could be effectively prevented by other means that would not endanger the woman and would not even possibly violate the verses cited. If so, then such means would undoubtedly be preferable.

Most often, though, men contemplate vasectomies simply because they do not want to father any more children. In light of the strong bias of the Jewish tradition for having children, and in light of the major demographic crisis facing the Jewish community that I will describe below, rabbis have not endorsed vasectomies for family planning purposes, seeing it as a violation of Jewish law and values and a threat to the continuity of the Jewish people.

Abortion. There is a clear bias for life within the Jewish tradition. Indeed, it is considered sacred. Consequently, although abortion is permitted in some circumstances and actually required in others, it is not viewed as a morally neutral matter of individual desire or an acceptable form of *post facto* birth control. Contrary to what many contemporary Jews think, Judaism restricts the legitimacy of abortion to a narrow range of cases; it does not permit abortion at will.

Judaism does not see all abortion as murder, as Catholicism does, because biblical and rabbinic sources understand the process of gestation developmentally. Thus Exodus 21:22–25 makes a clear distinction between an assailant who causes miscarriage of a fetus, when only monetary fines are imposed, as opposed to one who causes the death of the mother, when the rule is "life for life." According to the Talmud, within the first 40 days after conception the zygote is "simply water" (28). Another talmudic source distinguishes the first trimester from the remainder of gestation (29). It is not a theory of ensoulment that determines these marking points; it is rather the physical development of the fetus.

The effect of these demarcations is to make abortion during the early periods permitted for more reasons than during the rest of pregnancy (30). Classifying the first 40 days of gestation as "simply water," though, does not amount to a blanket permission to abort. Thus the RU486 pill, advertised as a "morning after pill" for those couples who simply do not want to have a baby, would be forbidden as a *post facto* contraceptive. On the other hand, if the woman's life or health would be threatened by pregnancy, then use of the RU486 pill would be preferable to a later-term abortion, both because it poses less risk for the woman and because the fetus is further from becoming a full human being.

The fetus does not attain the full rights and protections of a human being until birth, specifically when the forehead emerges or, if it is a breech birth, when most of the body emerges (31). The mother, of course, has full human status. Consequently, if the fetus threatens the life or health of the mother, then it may and in some cases must be aborted, as the following Mishnah graphically stipulates:

> If a woman has (life-threatening) difficulty in childbirth, one dismembers the embryo in her, limb by limb, because her life takes precedence over its life. Once its head (or its "greater part") has emerged, it may not be touched, for we do not set aside one life for another (32).

While all Jewish sources would permit and even require abortion in order to preserve the life or organs of the mother (33), authorities differ widely on how much of a threat to a woman's health the fetus must pose to justify or require an abortion. Based on a responsum by Rabbi Israel Meir Mizrahi in the late seventeenth century, many modern authorities also permit an abortion to preserve the mother's mental health, and this has been variously construed in narrow or lenient terms in modern times (34). To the extent that Jewish law makes special provision for an unusually young or old mother, an unmarried mother, the victim of a rape, or the participant in an adulterous union, abortion is construed to preserve the mother's mental health (35).

There is no justification in the traditional sources for aborting a fetus for reasons having to do with the health of the fetus; only the mother's health is a consideration. As a result some people object to performing an amniocentesis at all, even when the intent is to determine whether to abort a malformed fetus (36). Others reason in precisely the opposite direction. They point out that the sources

could not have contemplated abortions due to the condition of the fetus because nobody could know anything about that until very recently through technologies like amniocentesis and sonograms. Now that we have those tools, most rabbis justify using them to aid in the delivery of a healthy baby. Moreover, when those technologies reveal fetal abnormalities, many rabbis justify abortion on the basis of preserving the mother's mental health where it is clear that the mother is not able to cope with the prospect of bearing or raising such a child (37).

Many Conservative and Reform rabbis, and even a few contemporary Orthodox rabbis, have handled the matter in a completely different way. Our new medical knowledge of the status of the fetus, they say, ought to establish the fetus' health as an independent consideration in determining when abortion is justified (38).

In practice much of this discussion is moot, for Jews engage in abortion as if it were a matter of individual choice. That is a particularly problematic phenomenon for the contemporary Jewish community because Jews constitute only 0.2 percent of the world's population (while Christians make up a full 33 percent). To make matters worse, Jews are barely reproducing themselves in Israel and are falling far short of that in North America, where the Jewish reproductive rate is approximately 1.6 or 1.7 children per couple. Consequently, even those rabbis who are liberal in their interpretation of Jewish abortion law are also calling for Jews to marry and to have children so that the Jewish people and Judaism can survive.

Generating Conception

Artificial Insemination. Since Judaism prizes children so much, it is no wonder that rabbinic authorities have permitted unusual ways of having them for couples who cannot have them otherwise. Nevertheless, there are objections, or at least precautions, connected to some of the procedures.

Rabbis have not objected to uniting a man's sperm with his wife's ovum artificially, whether through artificial insemination or through in vitro fertilization (IVF) (39). Because of Judaism's appreciation of medicine as an aid to God, there is no abhorrence of such means merely because they are artificial.

The matter becomes more complicated when the donor is not the husband. Some rabbis object to such procedures on grounds of adultery. For many, however, adultery takes place only when the penis of the man enters the vaginal cavity of the woman, and that is clearly not the case when insemination takes place artificially. Not only is the physical contact missing; the intent to have an illicit relationship is also absent (40).

More commonly the objection to donor insemination is based on the possibility of unintentional incest in the next generation — specifically, if the product of the artificial insemination later happens to fall in love with a person of the opposite sex who is the child of the semen donor conceived with his wife. Since their biological father is the same man, these two people would be each other's natural half-brother or half-sister. That is problematic for some because it represents a violation of the Torah's laws against incest. Even for those who would invoke the lack of intent to excuse the couple from those laws, there still remains a critical health concern — namely the increased likelihood among consanguineous unions of genetic diseases transferring from one generation to the next.

This issue dissolves if the semen donor is known or if the donor would not likely be a marital partner for someone in the Jewish community. It was on the latter basis that a prominent Orthodox rabbi, Rabbi Moshe Feinstein, ruled that donor insemination would be permissible if the donor were not Jewish, for in his community intermarriage between Jews and non-Jews was rare. Those Orthodox Jews who will use donor insemination will therefore often require that the donor be a non-Jew.

The Conservative Movement's Committee on Jewish Law and Standards has approved by rabbinic ruling, according to which donor insemination is permissible if either the identity of the donor is known or, lacking that, that enough is known about him so that the child can avoid unintentional incest in his or her sexual partners (married or not) and so that the child can know as much as possible about his or her family traits, both medically and characterologically. In view, however, of the psychological problems that may ensue for the child, the donor, and/or the parents who raise the child (the "social parents"), all parties to the insemination should seek and receive appropriate counseling (41).

Egg Donation. The considerations described above with regard to donor insemination apply as well to egg donation. If the identity of the egg donor remains confidential, the same problems arise with regard to possible unintentional incest in the next generation, and the same solutions by the various rabbinic authorities apply. Specifically, either the egg donor's identity should be shared with the couple who will raise the child and ultimately with the child him/herself, or the woman should be a non-Jew, or enough about the biological mother must be shared with the couple and child to enable the child to avoid such unintentional incest. The donor, in my view, must also share enough information about her talents and traits to help the child understand him/herself. Finally, psychological counseling is appropriate for all concerned both before the procedure and afterward.

Egg donation, though, raises some additional problems. Semen donors incur virtually no medical risks, but that is not true of egg donors. In order to procure as many eggs as possible during each attempt, the donor must be hyperovulated with drugs, and there is some evidence that repeated hyperovulation increases the risk of ovarian cancer (42). This is especially troubling since the donor herself will not, by hypothesis, be gaining a child of her own but will rather be helping another couple have a child. For all that Jewish law prizes procreation, it values the life and health of those already born even more. Consequently, while healthy women may undergo the procedure to donate eggs once or twice, they may not do so much more than that, unless new studies allay the fear of increased cancer risk.

Normally, a child is defined as Jewish in traditional Jewish law if born to a Jewish woman. In cases of egg donation, however, some rabbis have maintained that it

is the donor of the gametes who is the legal mother. Most, though, have ruled that even if the egg comes from some other woman, as it does in egg donation, it is the bearing mother whose religion determines whether the child is Jewish or not, and the Conservative Movement's Committee on Jewish Law and Standards has adopted that view (43).

In vitro Fertilization (IVF), Gamete Intrauterine Fallopian Transfer (GIFT), Zygote Intrauterine Fallopian Transfer (ZIFT), etc. When a couple cannot conceive a fetus through sexual intercourse, even when assisted by timing their intercourse, by stimulating ovulation, or by surgery to correct a problem in either the man or the woman, and when the couple prefers to use their own gametes rather than those of donors, they may try any of a number of new techniques, some of which are listed in the title of this subsection. Since the Jewish tradition does not frown upon the use of artificial means to enable people to attain permissible ends, much less sanctified ones like having a child, the mechanical nature of these techniques is not an issue. On the contrary, the important thing to note in recent Jewish rulings is that infertile couples are *not obligated* to use these means to fulfill the man's duty to procreate, even though they *may* (44).

When a woman is impregnated with more than three fetuses, either naturally or artificially, an abortion may be indicated in order to preserve both the life of the mother and the viability and health of the remaining fetuses. For that purpose, such abortions are permitted (and possibly even required). When it can be determined through genetic testing that some of the fetuses have a greater chance to survive and to be healthy than others do, then it is permissible selectively to abort those less likely to survive. This is the same criterion to be used for triage decisions made at the end of life. If all of the fetuses are equally viable, the abortions must be done on a random basis. To avoid the necessity of selective abortion as much as possible, the Conservative Movement's Committee on Jewish Law and Standards has ruled that only three zygotes should be implanted at one time (45).

Surrogate Motherhood. This is really two different forms of overcoming infertility: "traditional surrogacy" or "ovum-surrogacy," in which the surrogate mother's own egg is fertilized by the sperm of the man in the couple who are trying to have a baby (presumably not the husband of the surrogate), and "gestational surrogacy," in which both the egg and the sperm are those of the couple, and the surrogate mother's womb is used to carry and deliver the baby.

From a Jewish perspective, this method of overcoming infertility, or at least something much akin to it, is among the oldest ways recorded in the Jewish tradition. Sarai (later, Sarah), after all, gives her handmaid, Hagar, to Abram (later, Abraham) specifically to conceive a son who would be attributed to Sarai, and Rachel and Leah likewise have their handmaids conceive children with their husband, Jacob. Leah, in fact, had already borne four sons by the time that she uses a surrogate mother because "she stopped bearing"—although she herself was later to bear him two more sons and a daughter (46). These are all, in modern terminology, ovum-surrogates, and even so, because the handmaid belonged to the man's wife, the Bible attributes the child to the wife rather than the surrogate.

These precedents notwithstanding, though, surrogate motherhood raises difficult emotional and legal problems—although not technological problems beyond those of artificial insemination (in ovum surrogacy) or IVF (in gestational surrogacy). Thus rabbis raise some concerns about the way in which a surrogacy arrangement should be handled, but they do not ultimately prohibit it. Specifically, the couple must abide by civil law in their region and, in light of the recency of this matter in most systems of law, the couple must be informed of the possibility of legal challenges. Furthermore Jewish law would require that steps be taken to ensure that the surrogate mother has full and informed intent to abide by the agreement—perhaps, in ovum-surrogacy, at least, by giving her a period of time (usually 30 days) after birth to cancel the agreement. The surrogate mother must not have physical or other conditions that would make pregnancy dangerous for her beyond the risks normally associated with pregnancy. In ovum-surrogacy the child must either be told the identity of the woman whose gametes he or she inherited or at least be given enough information to be able to avoid incest in his or her own sexual relations and to know about his or her physical and characterological background. Within these parameters, the few rabbis who have written about this have generally permitted surrogacy (47).

Prenatal Diagnosis and Treatment. Both for their own good and for that of their fetuses, pregnant women should seek and get prenatal care. They should also take the preventive measures that modern medicine prescribes to ensure a healthy baby, including restrictions on alcohol, smoking, and some prescription drugs; avoidance of toxins (e.g., in paints) and people with diseases which have been shown to cause fetal damage (e.g., German measles); and adoption of generally health-promoting habits of eating, hygiene, exercise, and sleep.

If the age or genetic background of a couple puts the child at risk for a degenerative, fatal genetic disease (e.g., Tay-Sachs) or for being seriously malformed, the mother may—but not must—undergo prenatal testing, even though that puts the fetus at some risk. Moreover, if the tests reveal that the fetus suffers from such maladies, the mother may choose to abort it. If, however, techniques exist that can cure the child in utero or once born, she may, and probably should, choose to employ those techniques rather than abort the fetus.

According to all interpreters of Jewish law, it is generally not permissible to screen specifically for gender just because one wants a boy or a girl or to screen for any characteristic other than disease (e.g., height, intelligence). Similarly the new sperm-splitting machine (a flow cytometer) to enable couples to choose either a boy or girl would generally violate Judaism's appreciation of people of both genders as equally created by God in the divine image (48). At the same time, Jewish law, as noted earlier, requires a man to father at least two children, specifically a boy and a girl. While that could not be used

to justify aborting a fetus of the same gender as those already born, it might justify using the flow cytometer in families who have produced three or more children of one gender and none of the other.

Gene Therapy, Genetic Engineering, and Cloning. Gene therapy is very new and only available in limited areas, and genetic engineering is still only a theoretical possibility. For example, techniques of genetic therapy are already being used to cure hydrocephalus while the fetus is still within the womb of the mother. The Human Genome Project has already discovered the genetic roots of many diseases, and that holds out the hope that someday soon those diseases may be cured. Indeed, on April 27, 2000, scientists in France reported the first success of gene therapy, using it to save three babies with severe combined immune deficiency (SCID) (48a).

There is already general agreement among rabbis that the legitimacy of human intervention to effect cure extends to procedures within the womb as well (49). When used in this therapeutic way, genetic engineering is an unmitigated blessing. Some rabbis have reservations about changing the stem cells themselves and thus all future generations, claiming that our divine mandate is to heal individuals who are ill but not to alter the nature of future human beings. That, in their view, would be to arrogate too much power to ourselves. Thus for them only therapeutic changes in the somatic cells of a diseased individuals would be permitted. Others, though, maintain that if we can root out a genetic disease not only from those who now have it but from descendants of such people as well, we should definitely do so, for our religious mandate is not only to cure diseases but to prevent them, if possible. With respect to degenerative genetic diseases like Tay-Sachs, I myself fall within the latter group.

Genetic engineering also, though, holds out the possibility that someday we will be able to change the nature of the human being so as to avoid the diseases that kill most of us today—heart disease, cancer, and the like. That already raises the question of whether we are effectively trying to reverse God's decision in the Garden of Eden to prohibit us from becoming immortal. Presumably, even if the currently deadly diseases are cured, we will eventually die of something else. So the Jewish justification for human beings to engage in curing may also be applied to genetically engineering ourselves to resist heart disease and cancer.

When we gain the ability to do these arguably permissible (and maybe even mandatory) things, we will also have the ability to change other things about ourselves. In fact, once we can change not only the genes of a particular fetus but even its germ line, we will be able to screen out traits that are not manifestations of a disease at all but merely characteristics that are deemed undesirable by certain individuals or groups. Abortion to eliminate defective fetuses poses the danger of the slippery slope where the definition of "defective" is broadened to the point of allowing only "perfect" children to be born, thus creating a master race. For example, we might change the genetic traits of shortness, merely average intelligence, a particular skin color, a propensity to alcoholism, and, perhaps, homosexuality. Moreover genetic engineering will create a new organism, and that poses real risks to human beings and to the environment.

There are thus some uses of genetic engineering that are clearly legitimate or illegitimate, but there are many where it is, and will be, difficult to tell. How do we determine when we are using genetic engineering appropriately to aid God in ongoing, divine acts of cure and creation and when, on the other hand, we are usurping the proper prerogatives of God to determine the nature of creation? More bluntly, when do we cease to act as the servants of God and pretend instead to be God?

Although cloning has been much more thoroughly discussed in the media, it actually presents fewer moral problems for Jews than genetic engineering does. Cloning, after all, does not introduce into the environment any new organism; it just replicates an organism that already exists, thus posing lesser risks. If cloning is used to overcome infertility, to aid in the research of diseases, or, in plants and animals, to produce food for starving people, it will be a very positive thing. On the other hand, cloning to avoid the intimacy of sexual intercourse, to gain immortality (as if that were possible through this technique), or to replicate oneself without any admixture of someone else's genes would be illegitimate uses of the technique. They smack of self-idolization and of the denial of human mortality; they thus make the moral and theological error of confusing human beings for God.

Our moral doubts about genetic engineering and cloning do not mean that research into these techniques should stop; the potential benefits to our life and health are enormous. They should prompt us, however, to exercise care in how we use our new capabilities. The problems are not just medical and technological; they are moral and theological, requiring us to reexamine the very ways we understand ourselves as human beings, our relationships to others and to God, and the limits inherent in being human.

Care of Severely Handicapped Newborns. Once a child is born, the child is a full-fledged human being and is to be treated in its health care like all other human beings. That is true for disabled newborns (or adults, for that matter) just as much as it is for those with no disabilities. The image of God in each one of us does not depend on one's abilities or skills; in this way the Jewish way of evaluating life is distinctly at odds with the utilitarian view common in Western societies.

If the child is born with severe disabilities that threaten his or her life, however, heroic measures need not be employed to keep the child alive. Here the same rules that govern the withholding and removal of life-support systems of any human being apply to newborns, with all of the diversity of opinion among rabbis noted in that section below. Some rabbis, however, are more lenient with respect to the treatment of newborns than they are regarding people dying later on in life because of the possibility, noted in Jewish law, that the child was born prematurely. Specifically, until the child is 30 days old, he or she is not considered to be a person whose life is confirmed (a *bar kayyma*). Therefore, while we may not

do anything actively to hasten the child's death, we may, according to these authorities, do less to sustain it than we would be called upon to do with regard to people who had lived beyond 30 days. Thus some who would insist on artificial nutrition and hydration for most dying people would not require it for life-imperiled infants less than 30 days old — except of course if the intervention holds out significant promise of curing the infant of the disease or condition. Some would require incubators, but most would not require surgery or medications beyond those necessary to relieve the child of pain (50).

Stem Cell Research

Jewish Views of Genetic Materials. Since human embryonic stem cells can be procured from aborted fetuses, the status of abortion within Judaism immediately arises. As we have seen, sometimes abortion is required by Jewish law and sometimes it is permitted, but mostly it is forbidden. The upshot of the Jewish stance on abortion, then, is that *if* a fetus was aborted for legitimate reasons under Jewish law, the aborted fetus may be used to advance our efforts to preserve the life and health of others.

In general, when a person dies, we must show honor to God's body by burying it as soon after death as possible. To benefit the lives of others, though, autopsies may be performed when the cause of death is not fully understood, and organ transplants are allowed to enable other people to live (51). The fetus, though, does not have the status of a full-fledged human being. Therefore, if we can use the bodies of human beings to enable others to live, how much the more so may we use a part of a body — in this case, the "water" or "thigh" that constitutes the fetus — for that purpose. This all presumes, though, that the fetus was aborted for good and sufficient reason within the parameters of Jewish law.

Stem cells for research purposes can also be procured from donated sperm and eggs mixed together in a petri dish and cultured there. Genetic materials outside the uterus have no legal status in Jewish law, for they are not even a part of a human being until implanted in a woman's womb, and even then, as we have noted, during the first 40 days of gestation their status is "as if they were simply water" (52). Abortion is still prohibited during that time except for therapeutic purposes, for in the uterus such gametes have the potential of growing into a human being, but outside the womb, at least as of now, they have no such potential. As a result frozen embryos may be discarded or used for reasonable purposes, and so may stem cells procured from them.

Other Factors in Stem Cell Research. Given that the materials for stem cell research can be procured in permissible ways, the technology itself is morally neutral. It gains its moral valence on the basis of what we do with it. The question, then, reduces to a risk-benefit analysis of stem cell research. The articles in a recent *Hastings Center Report* (53) raise some questions to be considered in such an analysis, and I will not rehearse them here. I want to note only two things about them from a Jewish perspective.

First, the Jewish tradition sees the provision of health care as a communal responsibility, and so the justice arguments in the *Hastings Center Report* have a special resonance for me as a Jew. Especially since much of the basic science in this area was funded by the government, the government has the right to require private companies to provide their applications of that science to those who cannot afford them at reduced rates or, if necessary, even for free. At the same time, the Jewish tradition does not demand socialism, and for many good reasons we, in the United States, have adopted a modified, capitalistic system of economics. The trick, then, will be to balance access to applications of the new technology with the legitimate right of a private company to make a profit on its efforts to develop and market applications of stem cell research.

Second, the potential of stem cell research for creating organs for transplant and cures for diseases is, at least in theory, both awesome and hopeful. Indeed, in light of our divine mandate to seek to maintain life and health, one might even argue that from a Jewish perspective we have a *duty* to proceed with that research. As difficult as it may be, though, we must draw a clear line between uses of this or any other technology for cure, which are to be applauded, as against uses of this technology for enhancement, which must be approached with extreme caution. Jews have been the brunt of campaigns of positive eugenics both here, in the United States, and in Nazi Germany (54), and so we are especially sensitive to creating a model human being that is to be replicated through the genetic engineering that stem cell applications will involve. Moreover, when Jews see a disabled human being, we are not to recoil from the *disability* or count our blessings for not being disabled in that way; we are rather commanded to recite a blessing thanking God for making people different (55). In light, then, of the Jewish view that all human beings are created in the image of God, regardless of their levels of ability or disability, it is imperative from a Jewish perspective that the applications of stem cell research be used for cure and not for enhancement.

We thus should take the steps necessary to advance stem cell research and its applications in an effort to take advantage of its great potential for good. We should do so, though, with restrictions to enable access to its applications to all Americans who need it and to prohibit applications intended to make all human beings into any particular model of human excellence. Instead, through this technology and all others, we should seek to cure diseases while simultaneously retaining our appreciation for the variety of God's creatures.

TECHNOLOGY AT THE END OF LIFE

Care of the Dying

General Concepts and Categories. Judaism prohibits murder, and it views all forms of active euthanasia as the equivalent of murder (56). That is true even if the patient asks to be killed. Because each person's body belongs to God, the patient does not have the right either to commit suicide or to enlist the aid of others in the act. Those who assist someone in a suicide violate Jewish law; the specific

nature and severity of the violation depend on how the aid is proffered. No human being has the right to destroy or even damage God's property (57).

The patient does have the right, however, to pray to God to permit death to come (58), for God, unlike human beings, has the right to destroy His own property. Moreover Judaism does permit passive euthanasia in specific circumstances.

Point When Passive Euthanasia Is Permissible. When does the Jewish obligation to cure end, and when does the permission (or, according to some, the obligation) to let nature take its course begin?

Authorities differ. All agree that one may allow nature to take its course once the person becomes a *goses*, a moribund person. But when does that state begin? The most restrictive position is that of Rabbi J. David Bleich, who limits it to situations when all possible medical means are being used in an effort to save the patient and the physicians assume that he or she will nevertheless die within 72 hours (59). Others define the state of *goses* more flexibly and therefore apply the permission to withhold or withdraw machines and medications during that time more broadly, in some cases up to a year or more (60).

In a rabbinic ruling approved by the Conservative Movement's Committee on Jewish Law and Standards (61), I noted that classical Jewish sources describe a *goses* as if the person were "a flickering candle," so that he or she may not even be moved for fear of inducing death. That description and that medical therapy only apply to people within the last hours of life (not even the last three days). Consequently, I argued, the appropriate Jewish legal category to describe people with terminal, incurable diseases, who may live for months and even years, is, instead, *terefah*. Permission to withhold or withdraw medications and machines would then apply to people as soon as they are in the state of being a *terefah*, that is, as soon as they are diagnosed with a terminal, incurable illness. In judging a disease to be incurable, we are not responsible for knowing whether a cure is imminent, for we are not God; the attending physicians must just use their best judgment.

Artificial Nutrition and Hydration. While intravenous cannulation to provide nutrition and fluids is appropriately used in people where there is reasonable hope for recovery, where no such hope exists, may one remove such tubes? Some rabbis have said no, reasoning that since artificial nutrition and hydration supply the liquids and nutrients that all of us need to survive, they cannot be classified as medications, which are used only when specific people need them (62). Others, however, noting that the Talmud specifically defines "food" as that which is ingested through the mouth and swallowed, classify artificial nutrition and hydration as medicine and permit removing or withholding them when recovery is not anticipated. Those attending the patient, though, must still go through the motions of bringing in a normal food tray at regular meal times in fulfillment of our duty to feed the starving (63).

Curing the Patient, Not the Disease. The important thing to note, however, is that there is general agreement that a Jew need not use heroic measures to maintain his or her life but only those medicines and procedures that are commonly available in the person's time and place. We are, after all, commanded to *cure* based on the verse in Exodus 21:19, "and he shall surely cure him." We are not commanded to sustain life *per se* (64). Thus, on the one hand, as long as there is some hope of cure, heroic measures and untested drugs *may* be employed, even though this involves an elevated level of risk. On the other hand, physicians, patients, and families who are making such critical care decisions are *not* duty-bound by Jewish law to invoke such therapies and may instead follow a course of hospice care. Indeed, because hospice care involves the support of family and friends and subjects the patient to the least amount of physical invasion possible, it often is preferable to more technologically sophisticated forms of treatment.

Pain medication may be administered as needed. Even in the last stages of life, when the dosage needed may actually hasten the patient's death, it may be used so long as the intent is not to kill the person but rather to alleviate his or her pain (65).

Moreover our duty to cure the patient rather than any specific disease means that if a person who is suffering from multiple, incurable, terminal illnesses develops pneumonia, doctors may refrain from treating the pneumonia if that will enable the patient to die less painfully. This would be in line with the strain in Jewish law that does not automatically and mechanically assume that preservation of life trumps all other considerations but rather judges according to the best interests of the patient (66). That principle does not extend so far as to permit mercy killing (active euthanasia), but it does make it permissible to refrain from administering the antibiotic so that the patient can die of his or her other diseases.

A person may volunteer to undergo an experimental procedure that holds out no hope to improve his/her own health but may increase medical knowledge and thereby help others only if it subjects the person to minimal or no risk. One's duty to preserve one's own life takes precedence over one's obligation to help other people preserve theirs.

Care of the Deceased: Autopsies and Organ Transplants

General Principles. The treatment of these topics in Jewish law depends on two primary principles. The general tenet that governs treatment of the body after death is *kavod ha'met*, namely, that we should render honor to the dead body as a sign of respect for both the deceased person and for God's property. Honor of the corpse, then, underlies Jewish burial customs.

The other principle that affects the topics of this section is that of *pikkuah nefesh*. When interpreting Leviticus 18:5, which says that we should obey God's commands "and live by them," the rabbis deduce that this means that we should not die as a result of observing them. The tenet that emerges is *pikkuah nefesh*, the obligation to save people's lives. This tenet is so deeply embedded in Jewish law that, according to the rabbis, it takes precedence over all other commandments except murder, idolatry, and incestuous or adulterous sexual intercourse (67).

Jews are commanded not only to do virtually anything necessary to save their own lives; they are also bound by the positive obligation to take steps to save the lives of others. The imperative to do so is derived from the biblical command, "Do not stand idly by the blood of your neighbor" (Leviticus 19:16). This means, for example, that if you see someone drowning, you may not ignore him or her but must do what you can to save that person's life (68).

What happens, when you can only save your life or someone else's? Whose life takes precedence? Since the Torah says that one should not exact interest from a fellow Jew "so that your brother may live *with you*" (Leviticus 25:36), and since you must therefore be alive at the time that you care for your brother, "your life takes precedence (*hayyekha kodemim*)" (69).

Autopsies. The two procedures that may interrupt the normal Jewish burial process are autopsies and organ transplants. Even though autopsies require invading the body of the deceased, in 1949 Israeli Chief Rabbi Isaac Herzog enunciated what has come to be the generally accepted position among Jews — namely, that while autopsies may not be done routinely, they are permissible if required by civil law, if the cause of death cannot otherwise be ascertained, if three physicians attest that the autopsy might help save the lives of others suffering from an illness similar to that from which the patient had died, or if a hereditary illness was involved so that performing the autopsy might safeguard surviving relatives. In all these cases we honor the dead by using the body to save lives.

Organ Transplantation. The overriding principles of honoring the dead (*kavod ha-met*) and saving people's lives (*pikkuah nefesh*) also work in tandem in organ transplantation. So the default assumption is that a person would be honored to help another live through organ donation.

Living Donors. Because one's own life takes precedence over helping someone else live, contemporary rabbis have generally permitted, but not required, donations from living donors when their life or health is not thereby subjected to major risk (70). If a family member suffers from leukemia and no appropriate bone marrow match is available, a married couple may seek to have another child in an attempt to find such a match, but only if they will not abort the child even if it becomes clear that the child is not the match they seek.

Cadaveric Donors. Since a dead person incurs no health risk, cadaveric donations are not only generally held to be permissible but, according to a responsum approved by the Conservative Movement's Committee on Jewish Law and Standards, actually a positive obligation so as to prevent the need of living persons incurring such risks while also saving the recipient's life. While traditional Jewish sources define the moment of death as the cessation of heartbeat and breath (71), even the chief rabbinate of the State of Israel in 1987 approved heart transplants, thereby accepting evidence of full brain death (including the brain stem) as fulfilling those requirements (72).

If a fetus has been aborted for reasons approved by Jewish law — namely to save the life or health of the mother or because the fetus suffers from Tay-Sachs or some similar fatal illness — the fetus may be used for purposes of transplant or experimentation.

Use of Animal or Artificial Organs; Animal Experimentation. While Judaism seeks to minimize pain to animals (73), it permits their use for food, for work, and, certainly, for saving a life. This would include medical research based on animal trials and the use of animal parts for transplantation, if that proves successful.

ENVIRONMENTAL USES OF TECHNOLOGY

The Jewish tradition, from the Torah on, was concerned with preserving God's world, leading to a series of ecological laws (74). Although classical Jewish law could not contemplate all the opportunities and problems produced by modern technology, it already prohibits wasting natural resources, even if one owns them, and it makes people responsible for the air and water pollution they cause. Judaism's appreciation of the world as belonging to God would additionally require us, in modern times, to create less waste than we moderns do, especially in technologically sophisticated societies, to recycle, and to use our new technology to reduce and, if possible, prevent pollution.

One application of biotechnology that, on the face of it, might cause special problems for Jews is the use of technology to produce new foods. Jewish dietary laws (*kashrut*, or "keeping kosher") restrict the fish, fowl, and animals that Jews may eat and the way that they are killed and their meat prepared and served. The Torah also forbids mixing seeds (*kilayim*) (75), but if non-Jews, who are not subject to this law, create hybrids, Jews may use them. An established principle in Jewish law, though, is that if a substance is chemically changed so that it cannot be reconstituted in its original form, it is "a new thing" (*davar hadash*) and, as such, loses any characteristics of its origins (76). Therefore bioengineered foods, such as cloned tomatoes, may certainly be eaten in accordance with Jewish dietary laws if the original substances are kosher and, if there is sufficient chemical change, even if the original substances are not kosher. Similarly Jews may engage in bioengineering new foods without violating the laws against mixing seeds if all (or all but one of) the materials to be combined are already so chemically changed as to constitute a new substance.

BIBLIOGRAPHY

1. See, for example, Deuteronomy 10:14; Psalms 24:1. See also Genesis 14:19, 22 (where the Hebrew word for "Creator" [*koneh*] also means "Possessor," and where "heaven and earth" is a merism for those and everything in between); Exodus 20:11; Leviticus 25:23, 42, 55; Deuteronomy 4:35, 39; 32:6.
2. *Genesis Rabbah* 11:6; *Pesikta Rabbati* 22:4.
3. See, for example, Deuteronomy 10:14; Psalms 24:1. See also Genesis 14:19, 22; Exodus 20:11; Leviticus 25:23, 42, 55; Deuteronomy 4:35, 39; 32:6. For these three and four other foundational principles of Jewish medical ethics, see E.N. Dorff, *Matters of Life and Death: A Jewish Approach to Modern Medical Ethics*, Jewish Publication Society, Philadelphia, PA, 1998, Ch. Two.

4. Thus, for example, bathing is a commandment, according to Hillel: *Leviticus Rabbah* 34:3. Maimonides includes rules requiring proper care of the body in his code of Jewish law as *a positive obligation* (not just advice for feeling good or living a long life), parallel to the positive duty to aid the poor: Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Ethics (De'ot)*, Chs. 3–5.
5. Babylonian Talmud (edited circa 500 C.E.) *Shabbat* 32a; Babylonian Talmud (edited circa 500 C.E.) *Bava Kamma* 15b, 80a, 91b; Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Murder* 11:4–5; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Yoreh De'ah* 116:5 gloss; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Hoshen Mishpat* 427:8–10. Jewish law views endangering one's health as worse than violating a ritual prohibition: Babylonian Talmud (edited circa 500 C.E.) *Hullin* 10a; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Orah Hayyim* 173:2; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Yoreh De'ah* 116:5 gloss.
6. See J.D. Bleich, *Tradition* **16**(4), 130–133 (1977).

   S. Freehof, *Reform Responsa for Our Time*, Hebrew Union College Press, Cincinnati, OH, 1977, Ch. 11; *Proc. Rabbinical Assembly* **44**, 182 (1983). All of the above are reprinted in E.N. Dorff and A. Rosett, *A Living Tree: The Roots and Growth of Jewish Law*, State University of New York Press, Albany, 1988, pp. 337–362.
7. Genesis 9:5; Mishnah (edited circa 200 C.E.); *Semahot* 2:2; Babylonian Talmud (edited circa 500 C.E.) *Bava Kamma* 91b; *Genesis Rabbah* 34:19 states that the ban against suicide includes not only cases where blood was shed but also self-inflicted death through strangulation, and the like; Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Murder* 2:3; Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Injury and Damage* 5:1; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Yoreh De'ah* 345:1–3. Cf. J.D. Bleich, *Judaism and Healing*, Ktav, New York, 1981, Ch. 26.
8. The rabbis note that the Nazarite, who takes an oath to avoid such pleasures, must, according to Numbers 6:11, bring a sin offering after the time specified by his oath, and they derive from that law that abstinence is prohibited: Babylonian Talmud (edited circa 500 C.E.) *Ta'anit* 11a. Cf. also Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Ethics (De'ot)* 3:1.
9. Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Ethics (De'ot)* 3:3.
10. *Sifra* on Leviticus 19:16; Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 73a; Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Murder* 1:14; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Hoshen Mishpat* 426.
11. E.g., Exodus 15:26; Deuteronomy 32:39; Isaiah 19:22; 57:18–19; Jeremiah 30:17; 33:6; Hosea 6:1; Psalms 103:2–3; 107:20; Job 5:18.
12. The permission and duty to heal: Babylonian Talmud (edited circa 500 C.E.) *Bava Kamma* 85a, 81b; Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 73a, 84b (with Rashi's commentary there). See also *Sifrei Deuteronomy* on Deuteronomy 22:2 and *Leviticus Rabbah* 34:3. Nahmanides, *Kitvei Haramban*, Bernard Chavel, ed., (Jerusalem: Mosad Harav Kook, 1963 [Hebrew]), vol. 2, p. 43, bases the duty for the community to provide health care on Leviticus 19:18, "You shall love your neighbor as yourself"; this passage comes from Nahmanides' *Torat Ha'adam (The Instruction of Man), Sh'ar Sakkanah (Section on Danger)* on Babylonian Talmud (edited circa 500 C.E.) *Bava Kamma*, Ch. 8, and is cited by Joseph Karo in his commentary to the *Tur, Bet Yosef, Yoreh De'ah* 336. Nahmanides bases himself on similar reasoning in Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 84b.
13. Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 74a.
14. The best of physicians deserves to go to Hell: Babylonian Talmud (edited circa 500 C.E.) *Kiddushin* 82a. Abraham ibn Ezra, Bahya ibn Pakuda, and Jonathan Eybeschuetz all restricted the physician's mandate to external injuries: See Ibn Ezra's commentary on Exodus 21:19 and cf. his comments on Exodus 15:26 and 23:25, where he cites Job 5:18 and II Chronicles 16:12 in support of his view; Bahya's commentary on Exodus 21:19; and Eybeschuetz, *Kereti U'pleti* on Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Yoreh De'ah* 188:5. See I. Jakobovits, *Jewish Medical Ethics*, Bloch, New York, 1959, 1975, 5–6. That a Jew may not live in a city without a physician: Jerusalem Talmud (edited circa 400 C.E.) *Kiddushin* 66d; see also Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 17b, where this requirement is applied only to "the students of the Sages."
15. Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Yoreh De'ah* 336:1.
16. *Midrash Temurrah* as cited in J.D. Eisenstein, ed., *Otzar Midrashim*, vol. 2, New York, 1915, pp. 580–581. Cf. also Babylonian Talmud (edited circa 500 C.E.) *Avodah Zarah* 40b, a story in which rabbi expresses appreciation for foods that can cure. Although circumcision is not justified in the Jewish tradition in medical terms, it is instructive that the rabbis maintained that Jewish boys were not born circumcised specifically because God created the world such that it would need human fixing, a similar idea to the one articulated here on behalf of physicians' activity despite God's rule; see note 2 above.
17. Babylonian Talmud (edited circa 500 C.E.) *Shabbat* 10a, 119b. In the first of those passages, it is the judge who judges justly who is called God's partner; in the second, it is anyone who recites Genesis 2:1–3 (about God resting on the seventh day) on Friday night who thereby participates in God's ongoing act of creation. The Talmud in Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 38a specifically wanted the Sadducees *not* to be able to say that angels or any being other than humans participate with God in creation.
18. On this entire matter, the Rabbinical Assembly, the rabbinic body of the Conservative Movement within Judaism, has created a rabbinic letter designed for use with adults and with teenagers discussing the concepts, values, and laws of Judaism governing intimate relations, marriage, nonmarital sex, and homosexuality. See E.N. Dorff, *This Is My Beloved, This Is My Friend: A Rabbinic Letter on Intimate Relations*, Rabbinical Assembly, New York, 1996. See also E.N. Dorff, *Matters of Life and Death*, Jewish Publication Society, Philadelphia, PA, 1998, Ch. 3–6.
19. Genesis 2:18; cf. *Midrash Psalms* on Psalms 59:2. Exodus 21:10 prescribes that a woman has conjugal rights in marriage, just as a man does, and the rabbis then spell out exactly how often a man must offer to have sex with his wife and how long she can refuse his advances without losing part of her settlement in a divorce; see Mishnah (edited circa 200 C.E.); *Ketubbot* 5:6–7. Note that he may never force himself upon her.
20. For example, Deuteronomy 7:13–14; 28:4, 11; Psalms 128:6.
21. Genesis 15:5; 17:3–6, 15–21; 18:18; 28:14; 32:13.

22. Mishnah (edited circa 200 C.E.); *Yevamot* 6:6 (61b); Babylonian Talmud (edited circa 500 C.E.) *Yevamot* 65b–66a; Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Marriage* 15:2; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Even Ha'ezer* 1:1, 13.
23. Deuteronomy 6:5.
24. Babylonian Talmud (edited circa 500 C.E.) *Yevamot* 12b. For a thorough discussion of this, see D.M. Feldman, *Birth Control in Jewish Law*, New York University Press, New York, 1968; reprinted as *Marital Relations, Abortion, and Birth Control in Jewish Law*, Schocken, New York, 1973.
25. Cf. Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Forbidden Intercourse* 16:2, 6; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Even Ha'ezer* 5:2.
26. Babylonian Talmud (edited circa 500 C.E.) *Shabbat* 110b.
27. Cf. J.D. Bleich, *Judaism and Healing*, Ktav, New York, 1981, p. 65.

    D. Feldman and F. Rosner, eds., *Compendium on Medical Ethics*, New York, 1984, pp. 46–47.

    S. Freehof, *Reform Responsa*, Hebrew Union College Press, Cincinnati, OH, 1960, pp. 206–208.
28. Babylonian Talmud (edited circa 500 C.E.) *Yevamot* 69b. Rabbi I. Jakobovits [*Jewish Medical Ethics*, Bloch, New York, 1959, 1975, p. 275] notes that "40 days" in talmudic terms may mean just under two months in our modern way of calculating gestation due to improved methods of determining the date of conception.
29. Babylonian Talmud (edited circa 500 C.E.) *Niddah* 17a.
30. Cf. D.M. Feldman, *Birth Control in Jewish Law*, New York University Press, New York, 1968, pp. 265–266 and Ch. 15.
31. Mishnah (edited circa 200 C.E.) *Niddah* 3:5.
32. Mishnah (edited circa 200 C.E.) *Oholot* 7:6. There are variant versions of this. Like our Mishnah, Jerusalem Talmud (edited circa 400 C.E.) *Shabbat* 14:4 reads "its greater part;" Tosefta (edited circa 200 C.E.) *Yevamot* 9:9 and Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 72b have "its head;" and Jerusalem Talmud (edited circa 400 C.E.) *Sanhedrin* 8, end, has "its head or its greater part."
33. Cf. I. Jakobovits, *Jewish Medical Ethics*, Bloch, New York, 1975, pp. 186–187 and No. 173 on pp. 378–379.
34. Cf. D.M. Feldman, *Birth Control in Jewish Law*, New York University Press, New York, 1968, pp. 284–294.

    M.H. Spero, *Judaism and Psychology: Halakhic Perspectives*, Ktav, New York, 1980, Ch. 12.
35. I. Jakobovits, *Jewish Medical Ethics*, Bloch, New York, pp. 189–190.
36. J.D. Bleich, *Contemporary Halakhic Problems*, Ktav, New York, 1977, pp. 112–115.

    J.D. Bleich, in F. Rosner and J.D. Bleich, eds., *Jewish Bioethics*, Sanhedrin Press, New York, 1979, Ch. 9, esp. pp. 161 and 175, No. 97.
37. Cf. D.M. Feldman, *Birth Control in Jewish Law*, New York University Press, New York, 1968, pp. 284–294.
38. E. Waldenberg, *Responsa Tzitz Eliezer*, 9:51 (1967) and 13:102 (1978) [Hebrew]; S. Israeli, *Amud Hayemini*, No. 35 cited in *No'am*,16 (K.H.) 27 (note) [Hebrew]; L. Grossnass, *Responsa Lev Aryeh* 2:205 [Hebrew]; A.J. Goldman, *Judaism Confronts Contemporary Issues*, Ktav, New York, 1978, Ch. 3, esp. pp. 52–62.
39. I. Jakobovits, *Jewish Medical Ethics*, Bloch, New York, 1975, p. 264.

    J.D. Bleich, *Judaism and Healing: Halakhic Perspectives*, Ktav, New York, 1981, pp. 82–84.
40. J.D. Bleich [*Judaism and Healing*, Ktav, New York, 1981, pp. 80–84] cites all of the following authorities (all of whom wrote in Hebrew) as requiring physical contact of the genital organs for adultery to occur: R. Shalom Mordecai Schwadron, *Teshuvot Zekan Aharon*, II, No. 97; R. Yehoshua Baumol, *Teshuvot Emek Halakhah*, No. 68; R. Ben Zion Uziel, *Mishpetei Uziel, Even Ha-Ezer*, I, No. 19; R. Moshe Feinstein, *Iggrot Moshe, Even Ha-Ezer*, I, No. 10; and R. Eliyahu Meir Bloch, *Ha-Pardes*, Sivan 5713. Nevertheless, as he points out, these Orthodox rabbis would prohibit donor insemination on the grounds of potential incest in the next generation, as discussed in the next paragraph. Moreover some Orthodox rabbis Bleich cites (R. Yehudah Leib Zirelson, R. Ovadiah Hadaya, R. Eliezer Waldenberg) maintain that violating the prohibition against adultery does not require genital contact, and so they would object to donor insemination on that ground as well.
41. E.N. Dorff, Artificial insemination, egg donation, and adoption, *Conservative Judaism* **49**(1), 3–60 (1996). This rabbinic ruling, approved in 1994 by the Conservative Movement's Committee on Jewish Law and Standards, also discusses several ancillary concerns, such as the identity of the father for various purposes in Jewish law; the psychological issues raised by the asymmetry in the situation — namely that the child will be the biological product of the woman but not the man who will be raising him/her; and the psychological need of the child to know his/her genetic roots. This ruling was reprinted in E.N. Dorff, *Matters of Life and Death*, Jewish Publication Society, Philadelphia, PA, 1998, pp. 66–115.
42. R. Spirtas, S.C. Kaufman, and N.J. Alexander, *Fertility and Sterility* [J. Am. Fertil. Soc.] **59**(2), 291–292 (1993). I want to thank my friend, Dr. Michael Grodin, for sharing this article with me. The 1988 Congressional report also reported a number of other possible complications caused by commonly used drugs to stimulate the ovaries, including early pregnancy loss, multiple gestations (fetuses), ectopic pregnancies, headache, hair loss, pleuropulmonary fibrosis, increased blood viscosity, and hypertension, stroke, and myocardial infarction; see U.S. Congress, Office of Technology Assessment, *Infertility: Medical and Social Choices*, OTA-BA-358, U.S. Government Printing Office, (Washington, DC, 1988), pp. 128–129. Once again, the demonstrated risks are not so great as to make stimulation of the ovaries for egg donation prohibited as a violation of the Jewish command to guard our health, but they are sufficient to demand that caution be taken and that the number of times a woman donates eggs be limited.
43. Rabbi Aaron Mackler, *In vitro Fertilization*, Draft No. 3, November, 1995, adopted by the Committee on Jewish Law and Standards in December, 1995, (unpublished); see p. 12 of the typescript.
44. See, for example, J.D. Bleich, *Judaism and Healing: Halakhic Perspectives*, Ktav, New York, 1981, pp. 85–91.

    E.N. Dorff, *Conservative Judaism* **49**(1), 17–18, 47–48 (1996), with regard to donor insemination and egg donation, but the same considerations, although sometimes in different forms, apply to IVF, GIFT, and ZIFT. See also E.N. Dorff, *Matters of Life and Death*, Jewish Publication Society, Philadelphia, PA, 1998, Ch. 3 and 4.
45. E.N. Dorff, Artificial insemination, *Conservative Judaism* **49**(1), 47 (1996); *Matters of Life and Death*, Jewish Publication Society, Philadelphia, PA, 1998, pp. 56–57, 101–102, 129–130.

46. Genesis 16:2 uses a play on words in Hebrew when Sarai says to Abram, "Look, the Lord has kept me from bearing. Consort with my maid [Hagar]; perhaps I shall have a son [also, I shall be built up] through her." This indicates that Ishmael, the son that resulted from this union, was not only to be considered Abram's son, but Sarai's as well. Similarly Rachel tells Jacob, "Here is my maid Bilhah. Consort with her, that she may bear on my knees and that through her I too may have children" (Genesis 30:3). At that time, Rachel was infertile (she gave birth to Joseph and Benjamin only later), but Leah, who had already had four sons, also gave her handmaid, Zilpah, to Jacob, and when Zilpah bore two sons to Jacob, Leah says " 'What fortune!' meaning, 'Women will deem me fortunate' " (Genesis 30:13) — indicating that those two sons were ascribed to Leah as well. That Leah had stopped bearing and therefore resorts to the use of her handmaid Zilpah: Genesis 30:9. That she later bears three more children: Genesis 30:14–21.
47. See, for example, J.D. Bleich, *Judaism and Healing*, Ktav, New York, 1981, pp. 92–95. While Rabbi Bleich generally prohibits or limits the use of new medical procedures, here he specifically argues against Rabbi Jakobovits' claim that surrogacy is inherently immoral and spends most of his discussion on the question of the Jewish identity of the child. See also M. Gold, *And Hannah Wept: Infertility, Adoption, and the Jewish Couple*, Jewish Publication Society, Philadelphia, PA, 1988, pp. 120–127; and E. Spitz, Through her I too shall bear a child: Birth surrogates in Jewish law. *J. Religious Ethics* **24**(1), 65–97 (1996), a slightly different version of which was approved by the Conservative Movement's Committee on Jewish Law and Standards in September, 1997. At that same meeting the Committee also approved a rabbinic ruling, as yet unpublished, by Rabbi Aaron Mackler that ultimately permits surrogacy but with more disclaimers and restrictions.
48. For a good, popular article describing some of the issues involved, see L. Belkin, Getting the girl. *N.Y. Times Mag.*, July 25, 1999, pp. 26–31, 38, 54–55.

48a. *N.Y. Times*, April 28, pp. A1–A16 (2000).

49. E.g., Bleich, *Judaism and Healing*, Bloch, New York, 1981, p. 106.
50. Rabbi Avram Reisner reasons this way in a responsum entitled *Peri- and Neo-Natology*, approved by the Conservative Movement's Committee on Jewish Law and Standards in Fall, 1995 (unpublished). That a child's life is not confirmed until 30 days of age: Babylonian Talmud (edited circa 500 C.E.) *Shabbat* 135b.
51. For classical sources on this, see E.N. Dorff, *Matters of Life and Death*, Jewish Publication Society, Philadelphia, PA, 1998, Ch. 9.
52. Babylonian Talmud (edited circa 500 C.E.) *Yevamot* 69b. Rabbi Immanuel Jakobovits notes that "40 days" in talmudic terms may mean just under two months in our modern way of calculating gestation, since the Rabbis counted from the time of the first missed menstrual flow while we count from the time of conception, approximately two weeks earlier. See I. Jakobovits, *Jewish Medical Ethics*, Bloch, New York, 1959, 1975, p. 275.
53. *Hastings Center Report*, March–April, 1999, pp. 30–48.
54. See S.J. Gould, *The Mismeasure of Man*, Norton, New York, 1996, and G.J. Annas and M.A. Grodin, *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, New York, 1992.
55. For a thorough discussion of this blessing and concept in Jewish tradition, see C. Astor, "... *Who makes people different*": *Jewish Perspectives on the Disabled*, United Synagogue of America, New York, 1985.
56. Mishnah (edited circa 200 C.E.); *Semahot* 1:1–2; Mishnah (edited circa 200 C.E.); *Shabbat* 23:5 and Babylonian Talmud (edited circa 500 C.E.) *Shabbat* 151b; Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 78a; Maimonides' *Mishneh Torah* (1177 C.E.) *Laws of Murder* 2:7; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Yoreh De'ah* 339:2 and the comments of the Shakh and Rama there.
57. This includes even inanimate property that "belongs" to us, for God is the ultimate owner. Cf. Deuteronomy 20:19; Babylonian Talmud (edited circa 500 C.E.) *Bava Kamma* 8:6, 7; Babylonian Talmud (edited circa 500 C.E.) *Bava Kamma* 92a, 93a; Joseph Karo's *Shulhan Arukh* (1565 C.E.) with the glosses of Moses Isserles. *Hoshen Mishpat* 420:1, 31.
58. Cf. RaN, Babylonian Talmud (edited circa 500 C.E.) *Nedarim* 40a. The Talmud records such prayers: Babylonian Talmud (edited circa 500 C.E.) *Ketubbot* 104a, Babylonian Talmud (edited circa 500 C.E.) *Bava Mezia* 84a, and Babylonian Talmud (edited circa 500 C.E.) *Ta'anit* 23a. Note that this is not a form of passive euthanasia: in that, people refrain from acting, but here God is asked to act.
59. J.D. Bleich, *Judaism and Healing*, Ktav, New York, 1981, pp. 141–142.
60. E.g., I. Jakobovits, *Jewish Medical Ethics*, Bloch, New York, 1975, p. 124 and No. 46.

    A.I. Reisner, *Conservative Judaism* **43**(3), 52–89, esp. pp. 56–58 (1991).
61. E.N. Dorff, *Conservative Judaism* **43**(3), 3–51, esp. pp. 19–26 (1991); *United Synagogue Rev.* **44**(1, Fall), 21–22 (1991).
62. E.g., Rabbi Avram Israel Reisner, holds this view; see *Conservative Judaism* **43**(3), 62–64 (1991).
63. E.N. Dorff, *Conservative Judaism* **43**(3), 34–39 (1991).
64. Thus the Talmud specifically says, "We do not worry about mere hours of life" (Babylonian Talmud (edited circa 500 C.E.) *Avodah Zarah* 27b). The Talmud also says, however, that we may desecrate the Sabbath even if the chances are that it will only save mere hours of life (Babylonian Talmud (edited circa 500 C.E.) *Yoma* 85a). The latter source has led some Orthodox rabbis to insist in medical situations that every moment of life is holy and that therefore every medical therapy must be used to save even moments of life; e.g., see J.D. Bleich, *Judaism and Healing*, Ktav, New York, 1981, pp. 118–119, 134–145. The only exception is when a person is a *goses*, which Rabbi Bleich defines as within 72 hours of death, at which time passive, but not active, euthanasia may be practiced. He then uses the source in *Avodah Zarah* only to permit hazardous therapies that may hasten death if they do not succeed in lengthening life. Rabbi Bleich's position is *not*, however, necessitated by the sources. On the contrary, they specifically allow us (or, on some readings, command us) not to inhibit the process of dying when we can no longer cure, even long before 72 hours before death (however that is predicted).
65. Rabbi Reisner does not accept this "double-effect" argument, but he would agree that pain should be alleviated as much as possible up to, but not including, the dosage that would have the inevitable effect of hastening the person's death, even if not intended for that purpose. See A.I. Reisner, *Conservative Judaism* **43**(3), 66, 83–85, Notes 50–52 (1991); and see, in contrast, E.N. Dorff, *Conservative Judaism* **43**(3), 17–19, 43–45, Notes 24–27 (1991). See also Rabbi Reisner's summary of the differences between the Dorff and Reisner positions, *Conservative Judaism* **43**(3), 91 (1991).

66. Tosafot, Babylonian Talmud (edited circa 500 C.E.) *Avodah Zarah* 27b, s.v., *lehayyei sha'ah lo hyyshenan*. See E.N. Dorff, *Conservative Judaism* **43**(3), 15–17, 43, No. 22 (1991). For a contrasting interpretation of this source, see A.I. Reisner, *Conservative Judaism* **43**(3), 56–57, 72, No. 21.
67. Babylonian Talmud (edited circa 500 C.E.) *Yoma* 85a–b (with Rashi there); Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 74a–b; *Mekhilta on Exodus* 31:13; and for a general discussion of this topic, see I. Jakobovits, *Jewish Medical Ethics*, Bloch, New York, 1975, pp. 45–98.
68. Babylonian Talmud (edited circa 500 C.E.) *Sanhedrin* 73a.
69. Babylonian Talmud (edited circa 500 C.E.) *Bava Metzia* 62a.
70. I. Jakobovits, *Jewish Medical Ethics*, Bloch, New York, 1975, p. 291; cf. also pp. 96–98. That this is the generally held opinion regarding living donors is true not only for Orthodox rabbis, some of whom he cites but also for Conservative and Reform rabbis. For Orthodox opinions, see Moshe Feinstein, *Igrot Moshe*, Yoreh De'ah 229 and 230 [Hebrew]; E. Waldenberg, *Tzitz Eliezer*, vol. 9, No. 45; vol. 10, No. 25 [Hebrew]; Obadiah Yosef, *Dinei Yisrael*, vol. 7 [Hebrew]. For a Conservative position (the only one I know of to date on living donors), see E.N. Dorff, *Choose Life: A Jewish Perspective on Medical Ethics*, University of Judaism, Los Angeles, CA, 1985, p. 23. For Reform positions, see S.B. Freehof, *New Reform Responsa*, Hebrew Union College, Cincinnati, OH, 1980, p. 62ff.

    S.B. Freehof, *Current Reform Responsa*, Hebrew Union College Press, Cincinnati, OH, 1969, pp. 118–125.

    W. Jacob, *Contemporary American Reform Responsa*, Central Conference of American Rabbis, New York, 1987, pp. 128–133.
71. Babylonian Talmud (edited circa 500 C.E.) *Yoma* 85a; Pirkei de-Rabbi Eliezer, Ch. 52; *Yalkut Shim'oni*, "Lekh Lekha," No. 72.
72. Yoel Jakobovits, [Brain death and] heart transplants: The [Israeli] chief rabbinate's directives, *Tradition* 24:4 (Summer, 1989), pp. 1–14. For Conservative positions, see S. Siegel, Updating the criteria of death, *Conservative Judaism* **30**(2), (Winter), 23–30 (1976) D. Goldfarb, The definition of death, *Ibid.*, pp. 10–22; the Rabbinical Assembly resolution urging organ donation *Proc. Rabbinical Assembly* **52**, 279 (1990), and the 1996 responsum by Joseph Prousser making organ donation a positive obligation (unpublished). The Reform Movement officially adopted the Harvard criteria (presumably, as modified by the medical community) in 1980.

    See W. Jacob, ed., *American Reform Responsa*, Central Conference of American Rabbis, New York, 1983, pp. 273–274.
73. See "Animals, cruelty to," *Encyclopedia Judaica* 3:5–7.
74. E.g., see R. Gordis, *Judaic Ethics for a Lawless World*, Jewish Theological Seminary of America, New York, 1986, pp. 113–122; and E. Schwartz and B.D. Cytron, *Who Renews Creation*, United Synagogue of Conservative Judaism, New York, 1993.
75. The choice of animals: Leviticus 11; Deuteronomy 14. The dietary laws additionally require a specific mode of slaughter to minimize pain to the animal (based on Deuteronomy 12:21), that the blood be drained from the meat (Genesis 9:4; Deuteronomy 12:23–25), and that meat and dairy meals be separated. The prohibition of mixing seeds: Leviticus 19:19; Deuteronomy 22:9.
76. See K. Abelson and M. Rabinowitz, Definition of a *Davar Hadash. In Proceedings of the Committee on Jewish Law and Standards of the Conservative Movement, 1980–1985*, Rabbinical Assembly, New York, 1988, pp. 187–190.

ADDITIONAL READINGS

J.D. Bleich, *Contemporary Halakhic Problems*, Ktav, New York, 1977.

J.D. Bleich, *Judaism and Healing*, Ktav, New York, 1981.

E.N. Dorff, *Choose Life: A Jewish Perspective on Medical Ethics*, University of Judaism, Los Angeles, CA, 1985.

E.N. Dorff, A Jewish approach to end-stage medical care, *Conservative Judaism* **43**(3), 3–51 (1991).

E.N. Dorff, A time to live and a time to die, *United Synagogue Review* **44**(1), 21–22 (1991).

E.N. Dorff, Artificial insemination, egg donation, and adoption, *Conservative Judaism* **49**(1), 3–60 (1996).

E.N. Dorff, *This Is My Beloved, This Is My Friend: A Rabbinic Letter on Intimate Relations*, Rabbinical Assembly, New York, 1996.

E.N. Dorff, *Matters of Life and Death: A Jewish Approach to Modern Medical Ethics*, Jewish Publication Society, Philadelphia, PA, 1998.

E.N. Dorff and A. Rosett, *A Living Tree: The Roots and Growth of Jewish Law*, State University of New York Press, Albany, 1988.

D.M. Feldman, *Birth Control in Jewish Law*, New York University Press, New York, 1968; reprinted as *Marital Relations, Abortion, and Birth Control in Jewish Law*, Schocken, New York, 1973.

D.M. Feldman and F. Rosner, eds., *Compendium on Medical Ethics*, New York, 1984.

S. Freehof, *Reform Responsa*, Hebrew Union College Press, Cincinnati, OH, 1960.

S. Freehof, *Current Reform Responsa*, Hebrew Union College Press, Cincinnati, OH, 1969.

S. Freehof, *Reform Responsa for Our Time*, Hebrew Union College Press, Cincinnati, OH, 1977.

S. Freehof, *New Reform Responsa*, Hebrew Union College, Cincinnati, OH, 1980.

M. Gold, *And Hannah Wept: Infertility, Adoption, and the Jewish Couple*, Jewish Publication Society, Philadelphia, PA, 1988.

D. Goldfarb, The definition of death *Conservative Judaism* **30**(2), 10–22 (1976).

A.J. Goldman, *Judaism Confronts Contemporary Issues*, Ktav, New York, 1978.

R. Gordis, *Judaic Ethics for a Lawless World*, Jewish Theological Seminary of America, New York, 1986.

W. Jacob, ed., *American Reform Responsa*, Central Conference of American Rabbis, New York, 1983.

W. Jacob, ed., *Contemporary American Reform Responsa*, Central Conference of American Rabbis, New York, 1987.

I.Jakobovits, *Jewish Medical Ethics*, Bloch, New York, 1959, 1975.

Y. Jakobovits, [Brain Death and] Heart Transplants: The [Israeli] Chief Rabbinate's Directives, *Tradition* **24**(4), 1–14 (1989).

A. Mackler, *In Vitro Fertilization*, Draft No.3, November, 1995, adopted by the Conservative Movement's Committee on Jewish Law and Standards in December, 1995 (unpublished).

Proceedings of the Committee on Jewish Law and Standards of the Conservative Movement, 1980–1985, Rabbinical Assembly, New York, 1988.

J. Prousser, *The Obligation to Preserve Life and the Question of Post-Mortem Tissue Donation*, Rabbinic ruling adopted by the Conservative Movement's Committee on Jewish Law and Standards, 1996, (unpublished).

A.I. Reisner, A Halakhic ethic of care for the terminally ill *Conservative Judaism* **43**(3), 52–89 (1991).

A.I. Reisner, *Peri- and Neo-Natology*, Rabbinic ruling approved by the Conservative Movement's Committee on Jewish Law and Standards in Fall, 1995 (unpublished).

F. Rosner and J.D. Bleich, eds., *Jewish Bioethics*, Sanhedrin Press, New York, 1979.

E. Schwartz and B.D. Cytron, *Who Renews Creation*, United Synagogue of Conservative Judaism, New York, 1993.

S. Siegel, Updating the criteria of death, *Conservative Judaism* **30**(2), 23–30 (1976).

M.H. Spero, *Judaism and Psychology: Halakhic Perspectives*, Ktav, New York, 1980.

R. Spirtas, S.C. Kaufman, and N.J. Alexander, *Fertil. Steril.* **59**(2), 292–293 (1993).

E. Spitz, Through her I too shall bear a child: Birth surrogates in Jewish law, *J. Religious Ethics* **24**(1), 65–97 (1996).

U.S. Congress, Office of Technology Assessment, *Infertility: Medical and Social Choices*, OTA-BA-358, U.S. Government Printing Office, Washington, DC, 1988.

See other RELIGIOUS VIEWS ON BIOTECHNOLOGY entries.

RELIGIOUS VIEWS ON BIOTECHNOLOGY, PROTESTANT

COURTNEY S. CAMPBELL
Oregon State University
Corvallis, Oregon

OUTLINE

INTRODUCTION

Protestant thought has played a critical and ironic role in the development of the ethical justifications and practices of biotechnology. The influential theologians of the Protestant Reformation, such as Martin Luther (1483–1546) and John Calvin (1509–1564), made a sharp distinction between God and nature (including human nature). All value is derivative of the divine being; nature and human beings, having experienced the dramatic and enduring consequences of the human fall from divine grace, have been divested of ultimate value and stand in need of redemption and reconciliation. One result of these profound theological claims was, to borrow from Max Weber's memorable phrase, the "disenchantment of the world" (1), that is, nature loses its reverential hold on human attitudes and actions, and comes to be seen as a realm open to the many manipulations of science and technology. Ironically, these manipulations over time would culminate in skepticism over the existence of the God that Protestant theology wanted to bring to the forefront of human consciousness. For some contemporary Protestants, this unanticipated result of a basic theological claim leads to religious and moral criticism of biotechnology.

It is nonetheless the case that Protestant perspectives on biotechnology are multiple and complex, and connected with deeply rooted theological claims about God, human beings, and nature. This article will provide an overview of these perspectives, beginning with attention to some of the formative views that characterize a position as "Protestant."

PROTESTANT DISTINCTIVENESS

Protestantism is first of all a "religion," rather than an "ethic," in which the primary questions concern the nature of God, God's relationship to humanity, and human nature and salvation. Nonetheless, themes that are embedded in these fundamentally theological issues have implications for ethical actions in the world and for perspectives on biotechnology.

Authority of Scripture

Protestant reformers and theologians are united in affirming the primary authority of the Bible as revealing the word of God to human beings about God, our relationship to God, and our prospects for redemption and reconciliation. God is revealed as Creator, Sustainer, and Redeemer. Human beings are mirrors or "images" of God in the world but have, through participation in sin and evil, fallen short of divine glory. Human reconciliation with God is made possible through the suffering of Jesus Christ for the sins of the world.

The authority of Scripture in the realm of salvation carries over, with some qualifications, in the realm of ethics. In Protestant thought, salvation is a matter of grace and mercy, which means the ethical life expresses salvation rather than being an instrument to salvation, as it is formulated in those religious traditions that emphasize salvation primarily by works or actions. Therefore the question of why human beings should be moral is for Protestants placed within a context of gratitude and gracious response to the divine gift of salvation, and a recognition of humility and human dependency upon powers beyond our control. One issue

for Protestants in examining the ethical questions of biotechnology is whether scientific research and its application retains this motivational sense of gratitude and humility or effaces it by emphasizing human accomplishment.

There is debate within the various Protestant traditions as to the sufficiency of Scripture for moral conduct. While the Bible has moral authority within Protestantism, it may not be *the only* authority. Theologian James M. Gustafson has argued that four resources are required for a comprehensive and adequate Christian (including Protestant) ethics:

1. Scripture and its interpretation through the historical tradition of Christianity
2. Philosophical insights, methods, and principles
3. Scientific methods and information
4. Interpretations of human experience (2)

In this understanding, Scripture illuminates the context of ethical issues, but it is not possible to move directly from a biblical passage to a moral conclusion about a current controversy in an area such as biotechnology. While such a position has been very influential among prominent Protestant theologians, it is not necessarily shared by evangelical and fundamentalist Protestants and their communities (3).

One implication of the moral authority of Scripture for Protestants is a corresponding non-normativeness of nature or the natural order. Indeed, some Protestant thought affirms the ongoing "ordering" of creation, rather than the created "order," precisely to emphasize the fluidity and dynamic interactions of nature, including those changes introduced by human beings. Nature is created good but is now currently disordered and no less in need of redemption. According to Protestant theology, Scripture reveals Jesus as healer of the disorder in nature, and provides a pattern for human beings to emulate. This model of scriptural interpretation is invoked by one influential scholar, Ronald Cole-Turner, to support the genetic alteration of plants in order to enhance their disease resistance. By enhancing the usefulness of plants, and diminishing environmental damage, human beings participate in the divine workings of redemption (4).

Thus, even though Protestant perspectives on biotechnology are profoundly influenced by interpretations of nature and the non-normativeness of the natural world, this does not imply that nature is without normative significance. Under the governance of divine providence, nature is susceptible to human intervention; indeed, human labor and ingenuity can use nature in the service of preserving and enhancing human (and animal) life. In some views, moreover, human beings are called to restore fallen nature to an original, properly ordered condition, given constraints of human finitude and fallibility. Given that Scripture does not speak directly to concrete issues in biotechnology, and that nature does not possess normative status but does possess normative significance for human action, it is possible to establish a Protestant theological presumption in favor of biotechnology. Biotechnology should be directed and constrained by norms of love, freedom, and stewardship.

Christian Freedom

As illustrated by the preceding discussion, dissent or "protest" is itself a characteristic Protestant theological perspective. This internal disagreement is manifested in many areas in Protestant thought, including the nature of ecclesiastical authority and its relation to biblical authority and the authority of personal conscience, as well as the role of sacraments mediated by the Church and their relation to saving grace. Given these disputes on matters of profound ecclesiastical importance, it comes as little surprise that Protestant history has been marked by continual reformations and the founding of new and diverse churches. In addition to these ecclesiastical implications, important ethical issues are embedded within the principle of "Christian freedom."

Freedom and Choice. Protestantism expresses a pronounced commitment to the primary of personal freedom and choice. Christian freedom should not be conflated with the secular norm of autonomy, for freedom is directed and constrained by love and ultimate accountability before God. However, this emphasis does give Protestants significant personal discretion in moral action without necessary reliance on a structure of specific moral rules.

Moral Pluralism. These theological and ethical commitments inevitably give rise to moral pluralism within Protestantism, as different interpretations are offered of the requirements of love and freedom. On virtually any moral question of consequence, a range of Protestant perspectives can be identified, without any ecclesiastical teaching authority available to provide a definitive conclusion. This is no less true of approaches to biotechnology, where assessments span the spectrum from hostility to biotechnology as an arrogant intrusion upon God's created order to a celebration of biotechnology as a beneficial means of partnership with God in continuing creation. These models will be discussed more fully below.

The Image of God. The primary Protestant claim about human nature is derived from the biblical account of the earth's creation; in this narrative, human beings are created in "the image of God," a theological description with profound normative implications. By this designation, human beings are given a status that distinguishes them from both God and nature. First, human beings are not God; the creator is sovereign over the created being. In particular, human beings lack the abilities to predict the results of action, control actions once they are initiated, or to adequately evaluate outcomes. Human distinctness from God is manifest in limitations such as finitude and fallibility. Persons do not have the capacities of omniscience and omnipotence attributed to the divine. When human aspirations exceed human capabilities, the created being runs the risk of the sin of pride or hubris, and of "playing God." This theological anthropology means for Protestant thought that the general theological presumption supporting biotechnological interventions characteristically will be constrained by concern with unforeseen consequences, slippery slopes, and admonitions of caution.

Notwithstanding these limitations, human beings "image" God in the world, a status that distinguishes them from animals, plants, and other manifestations of creation. Human beings are given a mandate of dominion to (*1*) care for and (*2*) cultivate creation, a mandate that within Protestant thought has been subject to numerous interpretations, including models discussed below such as passivity, anthropocentric domination, stewardship, and co-creatorship. These various interpretations, however, all affirm a common theme that responsible dominion involves a covenant with God that the human person will exercise freedom and love toward others — these "others" include human beings, the earth, and its creatures — in a way that benefits the common good. Responsible dominion also entails a recognition that ultimate accountability must be rendered to God for one's actions.

Much of the ambivalence and caution in Protestant thought concerning biotechnology is generated from the intrinsic tensions of the dual mandates of dominion, those of care and of cultivation. The principle of "caring" implies practices of conservation and preservation of our current conditions; the principle of "cultivation" implies reliance on human creative potential in using natural resources to bring about improvements and progress. Theologies of human dominion will tend to give greater emphasis to one or the other of these basic principles, or hold them in some kind of balance or tension, which in turn shapes perspectives on biotechnology.

Since nature does not have normative status within Protestant thought, the use of living organisms in biotechnological methods for making or modifying products, such as pharmaceuticals, enhancing strains of plants and animals, or manipulating genes to provide therapies for humans does not seem intrinsically wrong on the grounds that, for example, such practices "violate nature" or constitute "playing God." This is significant since a study of public attitudes to biotechnology indicates opposition to biotechnology is frequently grounded in claims that it is "not natural," or is "against God's will" (5). Although such claims are frequently presented as religious arguments, it is not clear they are well-ground in Protestant theological discussion: While human beings are not God in Protestant views, humans are responsible to and accountable before God in bringing about the full fruition of creation and in redeeming it from its current disordered condition. This supports the normative theological presumption in favor of biotechnology, unless it can be shown that a particular application of the technology will violate norms of love, freedom, or responsibility. The following section develops a general typology of Protestant perspectives that seek to negate, limit, or enact this presumption.

TYPOLOGY OF PROTESTANT THEOLOGICAL PERSPECTIVES

Nonintervention

One strand of Protestant thought emphasizes a noninterventionist or passive posture regarding the use of biotechnology. This perspective gives primacy to the principle of "caring," while attributing diminished importance to the principle of "cultivating." The noninterventionist understanding does not claim that nature is intrinsically good, and should therefore be left alone. Like all creation, nature has fallen from its pristine paradisiacal state. Rather, the problem is that human beings, because of their finitude and fallibility, are through technological interventions likely to make matters worse rather than better. Thus one Protestant writer urges caution and warns of the potential dangers from genetic manipulations because of the limited understanding and knowledge of human beings regarding the genome (6).

Within this understanding, moral priority is given to a norm of nonharm (as derived from love) rather than seeking to provide benefits through technological progress. Nonetheless, it is open to criticism on several grounds. It must first address what is particularly distinctive about biotechnology relative to other forms of human interventions. In many cases the position of nonintervention does not portray anything intrinsically wrong or misguided about biotechnology but expresses instead concern, rooted in the anthropology of finitude, about human prospects for control of biotechnological applications. This minimized capacity for control increases the risks of harm from biotechnology, both in terms of probability of occurrence and severity of the harm, to substantial and even unacceptable levels. Since it is impossible to empirically establish a risk assessment without proceeding ahead with biotechnology, and since prevention of harms have moral primacy over promotion of possible benefits, the course of ethical wisdom on this view lies in forgoing biotechnology.

The noninterventionist or passive perspective may also be challenged on the grounds that by forgoing benefits, harm is inevitably caused. It is possible to engage in a thought experiment and readily determine what benefits humanity currently enjoys from medical and biotechnological interventions that would have been forgone had the noninterventionist perspective prevailed throughout history. Thus this strand of Protestant thinking appears to convey a "free-rider" approach, willing to receive and make use of the benefits bestowed from the scientific legacy of prior generations but unwilling to develop these capabilities still further to benefit future generations.

Finally, nonintervention seems theologically suspect because of its neglect of the mandate of "cultivation." Responsible dominion involves a judicious balancing of cultivation and care, for cultivation provides justification for human intervention, through biotechnology or other means, in the natural world, while care sets limits on the scope and extent of that intervention. That is, both are necessary principles of dominion, and neither is by itself sufficient. It is part of Christian freedom and responsibility to work out the practical implications of these principles when, as in some instances of biotechnology, they come into conflict.

Anthropocentric Domination

If nonintervention presents one pole of a continuum in which caring assumes primacy over cultivation, the

other end of the role is represented by perspective of anthropocentric domination, which emphasizes the mandate of cultivation to the neglect of caring. Moreover the distinctive status of "image of God" of humans means that human beings should be the primary beneficiaries of cultivation of nature. In this understanding, ontology implies moral superiority: Humans are not only the culmination of creation but also its measure and purpose. Thus human beings are held to receive divine permission to use the resources at their disposal, including the natural resources of the earth and their own intellectual and creative potential to improve human welfare.

The position of anthropocentric domination shares with that of nonintervention the view of a fallen world and nature, but it differs in two important respects. A fallen world invites improvement, which in some views may support efforts of biotechnology and medicine to restore conditions similar to those of an Edenic paradise; in other interpretations, while paradise may be ineradicably lost, there nonetheless is a mandate for cultivating more humane and beneficial conditions for living. Nor is human finitude and fallibility as paralyzing on the domination account as it is within noninterventionist accounts. Indeed, the record of human history, particularly within medicine over the past century, shows dramatic improvements in health and welfare through sustained investigation, understanding, and manipulation of nature. Thus there is confidence that human interventions will culminate in greater benefits than harms, and that the risks of subsequent interventions, such as through biotechnological methods, can be controlled and minimized.

This position has been very influential in Protestant thought, particularly since the dawn of the scientific and industrial revolutions, and it is no surprise that it has also been the recipient of the most sustained philosophical criticism. In a very significant essay, Lynn White, Jr. laid the blame for the current "ecological crisis" precisely at the door of anthropocentric domination. This version of Christianity, in White's view, "not only established a dualism of man and nature but also insisted that it is God's will that man exploit nature for his proper ends" (7). And, such attitudes can easily be reflected in and perpetuated by contemporary biotechnology.

While White's thesis has been very controversial, and has been challenged on several grounds, it is important to differentiate descriptive and normative implications. Descriptively it is the case that (*1*) anthropocentric domination has been present in some Protestant understandings of nature and technology and (*2*) it can be rendered as compatible with at least part of the mandate of human dominion, that of cultivation. Normatively, however, cultivation is not unlimited; it should be directed and constrained by the principle of care, with its emphasis on protection, preservation, and conservation. Moreover cultivation is also limited by responsibility and accountability before God. This is to say that an anthropocentric account of dominion — which is the primary focus of White's critique — may not possess the fidelity to Protestant Scripture and teaching as other interpretations (8).

Stewardship

A third perspective that seeks a balanced response to the mandates of care and cultivation has historically been designated as *stewardship* or *trusteeship*. An ethic of stewardship can itself be articulated in terms of the Protestant themes of authority, agency, and accountability. Human beings have been given divine authority over nature, as well as the moral freedom to make choices regarding the use of natural resources. However, the content of such choices should reflect a concern for the common good (which is not limited to what is good for human beings) and persons are held to assume accountability before God for their choices. In short, the "dominion" of human beings is much more inclusive of other creatures than implied by the anthropocentric interpretation, and as stewards of the earth, human beings are in the service of God to render service to others, with "others" defined holistically rather than anthropocentrically.

Neither care nor cultivation receives moral primacy in this ethic. Rather, the stewardship ethic tries to maintain a responsible balance that both justifies human interventions on the grounds of improving the world and human welfare and limiting those interventions when they overreach these goals. Thus there is recognition that both benefits and harms can occur through human technologies, and making decisions about their use under the human conditions of finitude and fallibility is complex and permeated by genuine ethical uncertainties and dilemmas. Thus, while biotechnology can be justified, good reasons must be offered in its support, and constraints must be acknowledged and adhered to.

There is a depth of kinship between humans and the earth and its creatures present in the stewardship ethic. Human beings are "earth creatures," created by God, to be sure, but of the dust of the earth. Indeed, "human" and "earth" share a common etymological root, "humus." This commonality brings awareness of a sense of interdependence, mutuality, and humility that precludes the attitude of anthropocentric conquest present in the domination perspective.

While certainly very influential in the history of Christian and Protestant thought, to the point that some interpreters have conflated anthropocentrism with stewardship, the stewardship perspective is also not immune from moral critique. Since it tries to hold two principles in some kind of equitable balance, it often is found limited with respect to practical issues and controversies, where some choice about whether to give priority to care or to cultivate is not a theological abstraction but a practical necessity. In addition it has been argued that stewardship is an abstract ideal that is not embedded in cultural practices, which historically have reflected a domination perspective. Thus, on both counts, the practical relevance of this perspective may be much less compelling than its theoretical appeal.

Partnership

While the stewardship perspective affirms that human beings are authorized agents of God in the world, the

purposes for which the earth should be cared for and cultivated are by and large given by divine design. The partnership perspective emphasizes by contrast a much more interactive and engaged role of human beings in shaping these ultimate purposes. It presumes that creation did not end at some prehistorical period, but that creation is a dynamic, ongoing process that human beings participate in as created co-creators with God, with human and bioecological destiny not predetermined but to be shaped contextually. Thus "human work, especially our technology, may be seen as a partnership with God in the continuing work of creation. "[O]ur genetic engineering has the potential for being an extension of the work of God" (4).

This perspective has emerged fairly recently within Protestant thought; it has been given thoughtful exposition in the context of genetics and ecology by such prominent writers as Ronald Cole-Turner, Philip Hefner (9), and Ted Peters (10), and will certainly be increasingly influential in attempts to facilitate dialogue between religious, scientific, and biotechnological interests. Partnership, or created co-creator, models rely more heavily on scientific understandings of cosmology that do their Protestant predecessors, which are largely formulated within biblical cosmologies. The analogical and substantive features of partnership, and by inference its differences from stewardship, are eloquently articulated by British theologian Arthur Peacocke:

> It is as if man has the possibility of acting as a participant in creation, as it were the leader of the orchestra in the performance which is God's continuing composition. . . . [M]an now has, at his present stage of intellectual, cultural, and social evolution, the opportunity of consciously becoming *co-creator* and *co-worker* with God in his work on Earth, and perhaps even a little beyond Earth (11).

Indeed, Peacocke places human beings within a cosmos whose designs and purposes are known neither by human beings or by God; thus persons are even "*co-explorers*" with God. Such an understanding is certainly very compatible with, and gives theological justification for, biotechnology as, to continue the orchestral metaphor, the lead violin in the creative and explorative composition; technology is a metaphor and symbol for divine creative activity.

Certainly the partnership perspective is not without its theological detractors and these objections seem twofold. First, the position presumes that "the future of creation is uncertain, because God has not guaranteed its outcome" (4). Yet, precisely because the future is indefinite, and to be partly shaped and directed by human creativity, the perspective does not seem capable of generating clear limits or constraints. The necessity for some purpose or goal is an issue of acute importance in ethical evaluations of biotechnology, lest the technology create its own justificatory role. Second, the position seems to overstate human capacities to the point of pretension to playing God. While Peters has argued that the theological anthropology embedded in partnership is to "play human" fully and authentically, the long-standing concerns of Protestant thought about sin, human finitude, and fallibility are not as prominent in expositions of partnership. And, as illustrated above, it is precisely those concerns that stand behind many of the limits or cautions some strands in Protestant ethics wish to impose on biotechnology.

LIMIT CIRCUMSTANCES

Certainly many of the innovations of biotechnology, particularly with respect to products that bring about improvements in human health (eg, genetically engineered human insulin) have been welcomed by Protestant thinkers and communities. In this respect, the efforts of biotechnology to alleviate or cure disease are commonly set within a context of divine creativity and redemption, working through the imaginative instrumentality of human beings. Thus, with very few exceptions, biotechnology per se seems to not be theologically suspect; especially on the partnership perspective, biotechnology can be theologically praiseworthy and even morally required. Nonetheless, given the background theological commitments delineated above, Protestant thinkers do raise questions about the purposes and the procedural controls of biotechnology. Some scholars have raised the possibility that "control" itself is fundamental to the enterprise of biotechnology, by seeking to diminish human vulnerabilities to the capriciousness of the natural world. Thus the theme of human control is an issue that cuts across both substantive and procedural questions. These two questions can be more carefully examined by considering selected innovations in biotechnology that have raised concern among a range of Protestant thinkers and traditions.

Gene Therapy

One of the most significant innovations in biotechnology is the possibility to alter the genetic makeup of a person. Protestant denominations, under the auspices of the National Council of Churches in Christ, joined in the late 1970s with Roman Catholic and Jewish ecclesiastical bodies to raise questions about the risks of genetic manipulations, as well as a perceived arrogance of human control and mastery presupposed by such manipulations. In 1982 theologian Roger L. Shinn articulated in congressional hearings five base points for Protestant reflection on genetic interventions:

1. A bias for the sacredness of human life requires minimization of risk to the patient or subject.
2. A sense of "human inviolability" both permits interventions and limits their scope.
3. Efforts to eliminate genetic-based diseases are justifiable.
4. Genetic enhancements cross the boundary of inviolability and are dangerous.
5. Equity and justice should guide the distribution of benefits and burdens in genetics research (12).

In the intervening years, Protestant thought on human gene therapy has tended to reflect the considerations delineated by Shinn, and skepticism has gradually given way to a cautious endorsement of gene therapy in some

circumstances. Some questions have been amenable to resolution through increasing knowledge of scientific and technical issues, as well as through procedural safeguards of public oversight and monitoring. And Protestant commentators have characteristically insisted upon a specific moral rationale for genetic manipulations, that is, that the designed intervention have therapeutic potential for alleviating disease.

These concerns have tended to direct Protestant thought into ethical positions that some times converge and some times diverge with other ethical traditions, religious or secular. As one point of convergence, the Protestant commitment to human equality and the dignity of individual persons has required respect for genetic diversity and correlative opposition to efforts at genetic enhancements and positive eugenics. However, an issue of controversy has occurred over the validity of the line drawn in many bioethics discussions between somatic cell therapy and germ-line therapy. Protestant thought has argued for a greater continuity between somatic cell and germ line therapy insofar as the rationale for either form of genetic intervention is disease-based. A policy document of the National Council of Churches developed in 1986, even before attempts at somatic cell therapy had been conducted, asserted that while germ-line therapy needed "stringent control," it could not be precluded because of prospects of substantial benefits in alleviating disease (13). This view has evolved into a "wait-and-see" position presented in a document of the United Methodist Church: "We oppose therapy that results in changes that can be passed to offspring (germ-line therapy) until its safety and the certainty of its effects can be demonstrated and until risks to human life can be demonstrated to be minimal" (14). Such conditional opposition, or what Shinn describes as "cautious openness" (15), suggests that there is nothing intrinsically theologically objectionable with germ-line therapy, but it does place the burden of proof on those who wish to proceed with such interventions, in the sense that safety and efficacy must be demonstrated, rather than those who wish to prohibit germ-line therapy because of possible risks.

However, there are Protestant dissenters to cautious openness to germ-line therapy, and they are instead characterized by Shinn as adopting a position of "emphatic rejection." Some conservative Protestants have argued that as germ-line therapy will likely involve manipulation of the cells of human embryos, this method constitutes unethical experimentation on the unborn, and ultimately erodes the sanctity of human life. Indeed, some Protestants are concerned about somatic cell therapy, not on its own merits but because the underlying rationale of cure of disease raises a new set of concerns about responsibility to the vulnerable who cannot consent when considered in the context of germ-line manipulations.

A distinguished group of Protestant scholars has, while generally supporting development of gene therapies, raised some additional questions that revolve around the risk of humans losing control of the technology. One concern focuses on the relativity and elasticity of the concept of "disease." It may be difficult to uphold a firm line distinguishing legitimate uses of genetic therapies on the basis of whether they are directed to healing diseases or not because the very concept of disease is so fluid. In addition successes with gene therapy may encourage society and medical researchers to develop innovative uses for genetic manipulation that depart from the disease-based rationale. "[P]roper therapy also directs control to the goods of life and health, to the goal of healing genetic disease. If people simply celebrate genetic control itself, ... we fear that they will lose the capacity to direct and limit this new power to therapeutic uses" (16). An attitude of celebrate without caution, or a practice of cultivation without caring, may open the door to designing human descendants after our own preferred image and characteristics. Thus the slippery slope of most concern to Protestants may not be that of somatic to germ-line therapy but from therapy to enhancement, for enhancement is considered a response to "cosmetic purposes or social advantage" (14). In each case of concern about genetic manipulations — germ-line therapy, the concept of disease, and genetic enhancements — some common ethical perspectives are reinforced in Protestant thought by convictions about human finitude, fallibility, and pretensions to arrogance and pride.

Transgenic Research

For some Protestants, manipulation of the genome of human beings is not the only theologically problematic dimension of biotechnology. While animals do not generally have the theological and moral status of human beings within Protestantism, as creatures, animals do fall under the domain of responsible human stewardship or partnership. Thus, on some accounts, the insertion of genes from one species into another species to produce a transgenic organism or animal raises questions and objections. One kind of argument stems in part from an imperative rooted in the biblical creation narratives to plant and animal species to "reproduce after their own kind" (6). Transgenic biotechnological research may thus be understood to use genetic information for purposes not intended in the origins of plant, animal, or human life. Even though scientific research possesses the power and capability to bring about such genetic alterations, use of that power violates a normative ideal of species integrity and perpetuation.

A second argument regarding transgenic research is that it may compromise human distinctiveness. The claim in this instance is that by eroding distinctions between species, including animal and human species, through the creation of transgenic organisms, it then seems arbitrary to draw a line that would allow for such research on animals but not on humans. Protestant opposition to transgenic animal research in the 1980s led one researcher to reply, "I don't know what they [Protestant opposition] mean when they talk about the integrity of species. ... Much of all genetic material is the same, from worms to humans." This appeal to the commonality of genetic information between species leads Protestant theologian Andrew Linzey to question whether, if researchers are really convinced of this point, what grounds they would then have for opposing transgenic and eugenic research with humans (17).

Theological arguments against transgenic research do not always invoke abstract possibilities; they commonly cite examples of actual research projects that have culminated in harms rather than benefits. The creation of transgenic pigs in the 1980s through insertion of the human growth hormone gene has been cited by Protestant writers as a case in point of the risks and moral mistakes implicit in transgenic research. The point of the research, which was sponsored by the U.S. Department of Research, was to develop pigs that had greater muscle mass and leaner fat content, thus making the animals more commercially desirable. However, the pigs turned out to be excessively hairy, arthritic, impotent, and lethargic.

Significantly, theological critiques of this project argued that transgenic research was proceeding without ethical sensitivity to the harms experienced by the animals, rather than, as would be expected within a framework of anthropocentric domination, to human beings. It is important to note, however, that theological evaluations that assess biotechnology relative to its benefits and harms primarily focus on scientific and technical considerations, rather than on intrinsic theological issues. Such arguments may not be so compelling when the harms are uncertain and unforeseen, as in more recent projects that have created a chicken-quail hybrid as a prelude to understanding brain disorders in human. And, as noted above, other Protestant scholars have supported transgenic plant research on the grounds of increased productivity and diminished harms.

Of particular concern for some evangelical Protestants is maintaining the integrity or purity of the human genome. This follows from the special status of human beings in contrast to animals. Thus, on one account, "genetic information from any other organism which does or did not exist in the human genome should not be placed within humans" (6). However, this position raises questions over which Protestants would differ. First, a prohibition of mixing genetic information from other species with humans seems to suggest that human distinctiveness is constituted by genetic differences, rather than, for example, all that is embedded in the concept of the "image of God." The preservation of human genetic integrity seems bought at a price of genetic essentialism and theological reductionism, or what one theologian criticizes as the "gene myth" (10). Second, such a position presumes that there is something distinct or unique about genetic information, even though there are many other ways by which human beings might absorb or consume animals or plants or their products. While it is clear that not all Protestants hold similar views on transgenic research and organisms, it is clear that this form of biotechnology raises some widely shared questions (18).

Issues of Ownership: Patenting and Distributive Justice

The tools of biotechnology have raised important questions of ownership, particularly with regard to human cells that have been retrieved, modified, and immortalized, or of genetically engineered animals and plants. One issue of debate has focused on the legitimacy of patenting genetically modified life forms, while another has concentrated on ownership of the economic gains from commercial development of biotechnological products. Protestant scholars have engaged both of these debates within their own faith communities and in public discourse.

Scholars and clergy from both conservative and mainstream wings of Protestantism have been in the forefront of challenging efforts to patent genetic material, derived either from animals or (of greater concern) from human beings. A report on biotechnology issued by the World Council of Churches in 1989 opposed patenting of genetically altered animal life forms, asserting that "the patenting of life encodes into law a reductionist conception of life which seeks to remove any distinction between living and non-living things" (19). More recently, in 1995, a group of 186 religious leaders, primarily from Protestant churches, issued a "Joint Appeal against Human and Animal Patenting": "We . . . oppose the patenting of human and animal life forms. We are disturbed by the U.S. Patent Office's recent decision to patent human body parts and several genetically engineered animals. We believe that humans and animals are creations of God, not humans, and as such should not be patented as human inventions" (20,21). The "Joint Appeal" provoked a storm of controversy, not only between religious and scientific communities but within the Protestant community itself, with some theologians arguing that it was misguided and reactionary. However, gene patenting may simply be the issue that crystallizes concerns of many religious communities about biotechnology and the new genetics, concerns that have a deeper and broader significance. Indeed, religious objections to patenting the results of biotechnology appear to stem from a diversity of rationales, including (*1*) symbolism about life, (*2*) scientific reductionism, (*3*) human artifice, and (*4*) anthropocentrism.

As a statement about biology, one would be mistaken to infer from approval of a patent for a gene that "life" is thereby under the realm of the patent office, as is suggested by the language of the WCC report and by comments of individual religious clergy who supported the 1995 "Joint Appeal." However, critics who dismiss this objection as misinformed about science or naive as to the patenting process and public policy may overlook the theological symbolism at stake. The claim is best understood as seeking to resonate at a symbolic rather than literal level of interpretation. In a very probing study, Dorothy Nelkin and Susan Lindee have illustrated how a gene sequence, or the double-helix structure, is invariably used in scientific, academic, and popular literature as a symbol for life as a whole; the gene has become invested with a spiritual or sacral significance historically attributed to the soul (22). Thus a gene is not simply an object for scientific study and manipulation, it has become embedded with a complex matrix of cultural, ethical, and religious meanings. In this respect the willingness to proceed with gene patenting signifies an effort to extricate science from these embedded social meanings, and thereby is interpreted by some opponents as diminishing the value of life.

The value of life is underscored by a second rationale for objections to genetic patenting, a concern about scientific reductionism, an ideology embedded in contemporary

biomedicine that is viewed as blurring the boundaries between life and nonlife. If patenting is an instance of such a reductionist ideology, then some theological questions assume a greater legitimacy. In particular, scientific research and technological applications of parts of the body, such as tissues, cells, or genes, can conflict with religious values about bodily integrity that are central to the Protestant understanding of the "image of God" present in human beings. E. Richard Gold draws on these themes to contrast scientific instrumentalism with theological intrinsicalism: "The body, from a scientific viewpoint, is a source of knowledge of physical development, aging, and disease. From a religious perspective, the body is understood as a sacred object, being created in the image of God. . . . The scientist values the body instrumentally, as a means to acquire knowledge; the believer values the body intrinsically, for being an image of God" (23).

Put another way, Protestant (and Western) religious thought begins with a strong presumption that the status of the body as a whole is greater than the sum of its parts. The interest of medical science in the body stems from the prospect of gaining information about human character traits and behaviors, including susceptibilities to illness and bodily responses to disease, through study and understanding of the basic components of life, such as genes. The scientific value of the body as a totality is instrumental to the goal of deciphering the codes, messages, and functions of the fundamental components of the body that contain valuable genetic information. In this respect, gene patenting may reflect scientific reductionism in that (*1*) genes are viewed as scientifically more significant than the organic body totality and (*2*) the value of a cell or gene resides primarily in the information it provides researchers rather than as a symbol of life's dynamism and processes.

The third Protestant objection raised to gene patenting seeks to maintain a distinction between the realm of divine creation and the realm of human invention. The argument suggests that "creation" is the work of God, and within the Protestant Christian context, God creates *ex nihilo*, or "out of nothing." This is then differentiated from "invention" as a human affair, in which human beings re-organize, or more to the point with genetics, re-combine, already existing material elements to produce a new life form, such as the Harvard OncoMouse.

Given this distinction, however, it is difficult to see just how a patent application is any more an encroachment on divine creativity than the original research that inserted the human gene into the mouse. The latter might be objectionable with regard to the issues delineated concerning transgenic organisms, but it would then seem that the theological line needs to be drawn on that issue, not over the legal rights and restrictions granted by a patent. As with the objection that equates patenting DNA with patenting life, the creativity–inventiveness objection also seems to function as a symbol for objections that are important but not fully articulated.

Perhaps the strongest argument presented against biotechnology and patenting has been put forward by theologian Andrew Linzey who, in a broad-ranging theological critique of human use of animals more generally, sees in patenting the culmination of a departure from Christian stewardship and an embrace of anthropocentric domination. Linzey contends that "biotechnology in animal farming represents the apotheosis of human domination" (17). This development is the technological, if not logical, end point of the anthropocentric perspective that animals (despite also being created by God) belong to and exist for the benefit of human beings, and have no intrinsic value. While Linzey is aware that the human species has always made use of animal species, he nonetheless claims something distinctive is present in the application of biotechnology and genetic engineering to animals, namely it employs "the technological means of absolutely subjugating the nature of animals so that they become totally and completely human property" (17). In this interpretation, genetic engineering and patenting of animals is the moral equivalent of human slavery.

The Christian theology of responsible stewardship entails for Linzey maintenance and promotion of the good that already exists. The "artificial creation" of animals with disease-bearing characteristics, such as the OncoMouse, simply violates the integrity and design of creation. Moreover acceptance of patenting of genetically engineered animals, by which legal recognition is given to human property claims over animals, symbolizes biotechnological enslavement. The granting of patents over animals will "reduce their status to no more than human inventions, and signifies the effective abdication of that special God-given responsibility that all humans have towards the well-being and autonomy of sentient species" (17). Thus, on Linzey's account, animal patenting is a form of "idolatry" because it supplants God with human beings as owners of creation and "represents an attempt to perpetuate, to institutionalize, and to commercialize, suffering to animals" (17).

In explicating these Protestant reservations about genetic and animal patenting, it is important to acknowledge that other Protestant scholars have not found these reservations at all compelling. Ronald Cole-Turner, one of the most influential Protestant scholars at the intersection of theology and genetics, has maintained that because "there is no metaphysical difference between DNA and other complex chemicals, . . . there is no distinctly religious ground for objecting to patenting DNA" (21). Cole-Turner does not contend that there cannot be legitimate objections to patenting, but only that there are no specifically religious grounds for those objections. This claim enables Cole-Turner to encourage the initiation of dialogue between religious communities and scientific researchers, for "religion gives science its purpose, and science gives religion its eyes and its hands." It also provides Cole-Turner a basis for interpreting genetic engineering not in terms of anthropocentric domination, but rather as participation in divine creative activity.

One feature of the Protestant debate over gene patenting raises a more general concern in Protestant discourse about biotechnology, the issue of "commodification." This question may take two different but related forms. First, objections may be raised against biotechnology on the grounds that it transforms what is found in the world, such as genetic material, into commodities for commercial

development in accord with market values. This implies an understanding and valuation of genetic material that may be considered theologically objectionable (24). Thus, in its generally cautious but favorable appraisal of biotechnology, the World Council of Church admonishes, "the integrity of creation is damaged if biotechnology is utilized by commercial pressures to manufacture new life forms that are valued only as economic commodities" (19). Part of the concern embedded in this claim reiterates the Protestant concern with authoritarian control (i.e., commercial pressures) over biotechnology.

A second kind of argument may accept that certain things are legitimately classified as commodities, but protests against biotechnology on the ground that certain peoples and nations are better-positioned to participation in the biomarket and gain access to the benefits of biotechnology, while other peoples and nations will be excluded for economic reasons. In particular, the fruits of the biotechnological revolution are likely to be harvested by first-world peoples, while third-world countries may find themselves on the margins of the technologies. This runs contrary to the norm of distributive justice and the example of the ministry of Jesus, wherein Christians are encouraged to pay special attention to lifting the burdens and meeting the needs of the poor, the vulnerable, and the outcast. Some Protestant scholars believe the community of the vulnerable, and the moral primacy of responding to their needs, must be broadened from the "near" neighbor to the "stranger." Even though Protestant discussion about patenting of genetically engineered life forms is still in its embryonic stages, it seems clear that it will engage Protestants in dialogue and criticism with each other, and with the scientific and biotechnological communities.

Human Cloning

Recent scientific reports on successful mammalian cloning through the process of somatic cell nuclear transfer, which have in turn raised the prospects of human cloning in the near future, have revealed the pluralism of Protestant ethics perhaps more than any other question in biotechnology. Protestant theologians were invited to testify before the National Bioethics Advisory Commission established by President Clinton to recommend public policy on human cloning, and Protestant scholars have begun to contribute to the emerging ethics literature on this question (25). Yet this biotechnological development of the late 1990s is not a new question within Protestantism. Protestant theologians Joseph Fletcher and Paul Ramsey participated in influential academic and scientific forums in the 1960s and 1970s when cloning was first proposed as a scientific solution to many of the ills of the world. Fletcher and Ramsey staked out diametrically opposed positions and envisioned a world of human cloning that is remarkably prescient given the state of current discussion.

Fletcher advocated expansion of human freedom (self-determination) and control over human reproduction. He portrayed human cloning as one among a variety of present and prospective reproductive options that could be ethically justified under circumstances of overriding societal benefit. Indeed, for Fletcher, human cloning was a preferable method of reproduction relative to the "genetic roulette" of sexual reproduction: laboratory reproduction was "radically human" because it was deliberate, designed, chosen, and willed (26).

By contrast, Paul Ramsey portrayed cloning as a "borderline" or moral boundary for medicine and society that could be crossed only at risk of compromise to humanity and to procreation. He identified three "horizontal" (person–person) and two "vertical" (person–God) border-crossings of cloning. (*1*) Clonal reproduction would require dictated or managed breeding to serve the scientific ends of a controlled gene pool. (*2*) Cloning would involve non-therapeutic experimentation on the unborn. (*3*) Cloning would assault the meaning of parenthood by transforming "procreation" into "reproduction" and by severing the unitive and the procreative ends of human sexual expression. Theologically, cloning represented (*4*) the sins of pride or hubris, and (*5*) of self-creation in which human beings aspire to become a man-God (27). The legacy of the themes identified by Fletcher and Ramsey concerning human cloning have been perpetuated in both recent religious and secular reflection on cloning. Within the Protestant religious communities, these debates have revolved around several contested themes.

Sanctity of Life. Protestant evangelicals have appealed to the sanctity of human life to argue against human cloning. The process of somatic cell nuclear transfer for the purpose of making a new human being, at least as illustrated in the current animal studies, would inevitably result in loss of human embryonic life. In addition evangelical positions claim that contemporary societal disregard for the sanctity of human life could possibly lead to a re-definition of humanity, such that the clone may be treated as a repository for spare organs and tissues (28).

Parenthood. Conservative and evangelical Protestants also object to human cloning on the basis of an intrinsic connection between the unitive and procreative purposes of sexuality as embedded in the *Genesis* creation story. Sexuality is understood to be a divine gift with the twin purposes of uniting the partners through a physical expression of their love and for bringing offspring into the world. In this context human cloning runs contrary to critical biological, emotional, and symbolic connections between spouses and between parent and child. In particular, the idea of a child as a "gift" is effaced as the child becomes both a project and a projection of the self (29). This argument interprets human cloning to diminish humanity to "raw material" out of which an artifice can be designed and constructed in our image, rather than the "image of God," thus leading to power over other humans rather than enhanced choices.

The Image of God. Conservative and evangelical Protestant scholars maintain that as bearers of God's image, human beings gain insight into self-understanding and human uniqueness and receive a distinctive status relative to the rest of creation. Cloning risks devaluing this image of the person by suggesting genetics is the essence of personhood, or by valuing the clone because of its replication of valued characteristics of another person.

Some mainstream Protestant theologians have argued, by contrast, that human cloning can express the creative dimensions of the *imago Dei* insofar as the new genetics promotes human dignity and welfare (30). Moreover the Christian vocation of freedom warrants the pursuit of scientific knowledge, when coupled with the obligation of accountability delineated above. Even though the reality of sin will manifest itself in an ongoing disparity between a designed future and its reality, this position holds that Christians are given permission to "sin bravely" in the pursuit of progress. Thus, if further research on human cloning can establish a reasonable expectation of benefits, and ensure human dignity, then both research and eventually human cloning seem warranted.

CONCLUSION

Protestant thought can celebrate biotechnology because of the prospects of revealing more about God's creation and applying that knowledge for human betterment, and the betterment of life on this planet. Simultaneously Protestant thought characteristically urges caution about the biotechnological revolution, lest use be transformed into abuse. The powers that human beings can wield through biotechnology must be acknowledged as limited and beyond our capacity to fully control, but Protestant thought has historically been concerned not simply with the external action but what such action reflects or expresses about a person's moral character. In this regard Protestant theological ethics forces the question of what kind of persons we need to be in order to wield such powers for good rather than ill.

BIBLIOGRAPHY

1. M. Weber, *The Protestant Ethic and the Spirit of Capitalism*, Scribner's, New York, 1958.
2. J.M. Gustafson, *Protestant and Roman Catholic Ethics: Prospects for Rapprochement*, University of Chicago Press, Chicago, IL, 1978.
3. A.D. Verhey, in J.F. Childress and J. Macquarrie, eds., *The Westminster Dictionary of Christian Ethics*, Westminster Press, Philadelphia, PA, 1986.
4. R. Cole-Turner, *The New Genesis: Theology and the Genetic Revolution*, Westminster/John Knox Press, Louisville, KY, 1993.
5. T.J. Hoban and P.A. Kendall, *Consumer Attitudes about the Use of Biotechnology in Agriculture and Food Production*, North Carolina State University Press, Raleigh, 1992.
6. K.P. Wise, in R.D. Land and L.A. Moore, eds., *Life At Risk: The Crises in Medical Ethics*, Broadman & Holman, Nashville, TN, 1995.
7. L. White, *Science* **155**, 1203–1207 (1967).
8. J.B. Cobb, Jr., *Second Opinion* **18**, 11–21 (1992).
9. P. Hefner, in T. Peters, ed., *The Cosmos as Creation: Theology and Science in Consonance*, Abingdon Press, Nashville, TN, 1989.
10. T. Peters, *Playing God? Genetic Discrimination and Human Freedom*, Routledge, New York, 1997.
11. A.R. Peacocke, *Creation and the World of Science*, Clarendon Press, Oxford, UK, 1979.
12. J.R. Nelson, *Hum. Gene Ther.* **1**, 43–48 (1990).
13. National Council of Churches, *Genetic Science for Human Benefit*, National Council of Churches of Christ in the U.S.A., New York, 1986.
14. United Methodist Church, *Genetic Science*, General Board of Church and Society, Washington, DC, 1992.
15. R.L. Shinn, *The New Genetics*, Moyer Bell, Wakefield, RI, 1996.
16. H. Bouma et al., *Christian Faith, Health, and Medical Practice*, William B. Eerdmans Publishing, Grand Rapids, MI, 1989.
17. A. Linzey, *Animal Theology*, University of Illinois Press, Chicago, IL, 1994.
18. J.R. Nelson, *On The New Frontiers of Genetics and Religion*, William B. Eerdmans Publishing, Grand Rapids, MI, 1994.
19. World Council of Churches, *Biotechnology: The Challenge to the Churches and to the World*, World Council of Churches, Geneva, 1989.
20. Call for a Moratorium on Gene Patenting, *The Christian Century* **112**(20), 633 (1995).
21. R. Cole-Turner, *Science* **270**, 52 (1995).
22. D. Nelkin and M.S. Lindee, *The DNA Mystique: The Gene as a Cultural Icon*, Freeman, New York, 1995.
23. E.R. Gold, *Body Parts: Property Rights and the Ownership of Human Biological Materials*, Georgetown University Press, Washington, DC, 1996.
24. M.J. Hanson, *Hastings Cent. Rep.* **27**(6), 1–22 (1997).
25. R. Cole-Turner, *Human Cloning: Religious Responses*, Westminster/John Knox Press, Philadelphia, PA, 1997.
26. J. Fletcher, *The Ethics of Genetic Control*, Anchor Press, Garden City, NY, 1974.
27. P. Ramsey, *Fabricated Man: The Ethics of Genetic Control*, Yale University Press, New Haven, CT, 1970.
28. J.K. Anderson, *Genetic Engineering*, Zondeman Publishing House, Grand Rapids, MI, 1982.
29. G. Meilander, *Bio Law* **II**(S), 114–118 (1997).
30. C.S. Campbell, in *Cloning Human Beings*, vol. 2, National Bioethics Advisory Commission, Rockville, MD, 1997, pp. D1–D64.

See other RELIGIOUS VIEWS ON BIOTECHNOLOGY entries.

REPRODUCTION, ETHICS, MORAL STATUS OF THE FETUS

BONNIE STEINBOCK
SUNY Albany
Albany, New York

OUTLINE

INTRODUCTION

The moral status of the fetus is not only central to the abortion debate but is also relevant to reflection on such issues as assisted reproductive technology (ART), the moral status of extracorporeal and frozen embryos, prenatal genetic testing, fetal and embryo research, and fetal therapy. This article discusses the meaning of moral status, and its relation to moral value and moral rights. It goes on to discuss the moral status of the fetus in the context of the abortion debate, outlining the main positions that have been taken. These include genetic humanity, brain birth, viability, personhood, potential personhood, the possession of a "future like ours," sentience, and a multi-criterial approach.

WHAT IS MORAL STATUS?

Mary Anne Warren succinctly explains the concept of moral status:

> To have moral status is to be morally considerable, or to have moral standing. It is to be an entity towards which moral agents have, or can have, moral obligations. If an entity has moral status, then we may not treat it in just any way we please; we are morally obliged to give weight in our deliberations to its needs, interests, or well-being. Furthermore, we are morally obliged to do this not merely because protecting it may benefit ourselves or other persons, but because its needs have moral importance in their own right (1).

Some entities clearly have moral status: for example, people. It is difficult even to imagine a moral view that did not require us to consider the interests and well-being of people. Indeed, this requirement may define "the moral point of view" (2). Just as clearly, some entities do not have moral status: for example, ordinary rocks. They are just things, of no particular value or importance. The hard questions fall in between people and more things. What about nonhuman animals, permanently unconscious humans, plants, species, and the environment? A coherent, nonarbitrary answer to the question of whether these beings have moral status requires a general theory of moral status.

Moral Status and Moral Value

While some "mere things" have little or no value, others are valuable in various ways. Some things have commercial value, some aesthetic value, some scientific value, and so on. All of these ways of being valuable can be distinguished from being morally valuable or having moral value. What, then, is it for something to have moral value? A plausible suggestion is that something has moral value if there are moral reasons for valuing it. Consider, for example, the flag of the United States. Viewed solely as a piece of material, the flag has relatively little value. But it is more than that. It is the emblem of our country. School children pledge allegiance to the flag, it is flown over government buildings, lowered to half-mast when important people die, and so forth. Many people have strong feelings about the flag: They respect it, revere it, even love it. Of course it is not the flag itself that inspires these feelings, but what it stands for. This explains why it is held to be wrong to fly the flag in tattered condition or to allow it to touch the ground. It also explains the outrage felt by many patriots when the flag is burned in a political protest. The point is not that flag burning is necessarily wrong. In fact burning is the recommended way of disposing of a worn-out flag. In addition flag-burning could be a legitimate form of political protest. The point is rather that it is a not a matter of moral indifference what one does to a flag. There are moral reasons to treat flags in certain ways. This suggests that the flag, because of its symbolic significance, has moral value.

Bonnie Steinbock has suggested that moral status be distinguished from moral value (3). The difference lies in the reasons why certain treatment is regarded as morally wrong. In the case of the flag, the reason why it is held to be wrong to burn a flag in protest is that this shows disrespect for the flag and the country for which it stands. The reason is not a "golden rule" type reason, along the lines of "how would you like it if you were a flag and someone burned you?" In making this distinction, we recognize that it does not matter to a flag what is done with it, and this differentiates flags from, say, people or animals. To accord something moral status is not merely to consider it important or valuable, and worthy of protection; it is to take its perspective into account when making moral decisions. In this view, flags might have moral value, but they lack moral status. This is because flags do not have a point of view. They cannot have a point of view because they lack awareness of any kind. Mere things can be destroyed, but whatever is done to them cannot matter to them. They do not have a stake in their preservation or well-being (if mere things can be said to *have* a well-being). They have no interests because only beings who can care about how they are treated or what is done to them can have interests of their own. Lacking interests, mere things lack moral status, but it does not follow that it is morally permissible to treat them any way you like.

Many environmentalists object to the view that only sentiment beings can have moral status: the "sentience only" view. For "deep ecologists," like Aldo Leopold, moral status is not limited to sentient beings. Natural plant and animal species, populations, and habitats can all have moral status, just as much as sentient beings. As Mary Anne Warren expresses the point, "To many environmentalists, a theory which allows us to have moral obligations *regarding* the nonsentient elements of the

natural world but never *to* them, seems just as inadequate as the Kantian theory, which allows us to have duties regarding animals, but never to them" (1, p. 72). While this debate is important for environmental ethics, it is not directly relevant to the moral status of the fetus, as those who maintain that the fetus has moral status do not usually base its status on its being an element of the natural world but rather on such features as its being genetically human, potentially a person, or the kind of future it will have.

Moral Status and Moral Rights

All beings that have moral rights have moral status. Indeed, if someone has a right to something, this imposes obligations on others to behave in certain ways. Your moral rights limit my freedom with respect to how I am permitted morally to treat you. This implies that from a moral point of view, someone who has a right counts or matters, which is the same as ascribing to it moral status. However, the reverse does not hold. It does not follow that every being that has moral status has moral rights. It is even possible that there are no moral rights, as some utilitarians maintain. There can be an account of morality that does not include moral rights, but there cannot be an account of morality that does not include a theory of moral status.

MORAL STATUS AND ABORTION

We start with the problem of abortion, where the moral status of the fetus has been a central and contentious issue. First, a word about terminology. During the first week of its existence, the fertilized egg is known as a conceptus. The term "embryo" refers to the entity between the second and eighth weeks. From the eighth week until birth, it is a fetus. However, the term "fetus" is often used to refer generally to the unborn throughout pregnancy. This article follows that usage, except where the different phases of gestation have a bearing on moral status and need to be distinguished. Those who are "pro-life" (often referred to as "conservatives" on abortion) argue that fetuses, throughout gestation, have the same moral status as born human beings, and therefore killing fetuses is seriously wrong, as wrong as killing born human beings. Those who are "pro-choice" (often referred to as "liberals" on abortion) usually argue that fetuses differ in morally important ways from born human beings, and for this reason, lack full moral status and in particular a right to life. Abortion, while not desirable, is not seriously wrong, in this view, and is certainly not equivalent to killing a born human being.

A number of writers have pointed out that the abortion issue does not turn solely on the moral status of the fetus. For example, Thomas Murray argues that "Fetal personhood is only one strand in the web holding our moral judgments about abortion" (4, p. 146). Sociologist Kristin Luker argues that it is a mistake to think that the views of pro-choice and pro-life activists about abortion are determined by their philosophical or religious views on the moral status of the fetus (5). Rather, their views on the morality of abortion stem from their differing views on the meaning and value of sexuality, motherhood, and the proper role of women. How they view the fetus is determined by how they regard abortion, not the reverse. Legal philosopher Ronald Dworkin argues that the abortion debate is not really about the moral status of the fetus, despite the rhetoric on both sides. According to Dworkin, even those vehemently opposed to abortion do not actually believe that a fetus is from the moment of its conception a full moral person with rights and interests equal in importance to those of any other member of the moral community. Nor do those committed to protecting a woman's "right to choose" think of the developing fetus as just a part of the pregnant woman's body. "The disagreement that actually divides people is a markedly less polar disagreement about how best to respect a fundamental idea we almost all share in some form: That individual human life is sacred" (6, p. 13). Judith Thomson maintains that it is a mistake to think that the abortion debate is over if the premise that the fetus is a person, with a right to life, is accepted. For the abortion debate also raises the question of whether the fetus has a right to the use of the pregnant woman's body, and how much of a sacrifice to sustain its life she is required to make. Thomson thinks that even if fetuses are persons, at least some abortions could be justified (7). Finally, some feminists regard the inquiry into the status of the fetus as irrelevant to the problem of abortion (8). They view abstract inquiries into the moral status of the fetus as distracting from the real issues, which have to do with creating the social conditions that permit women to make genuine reproductive choices.

All of these voices have deepened the abortion debate, but it is doubtful that any has proved the irrelevance of moral status. Even if Luker is right about the origin of people's views about the moral status of the fetus, that does not address the correctness or plausibility of their views. Feminists who focus solely on the interests of women can be fairly charged with simply avoiding the question of fetal moral status. Even if sexism, racism, poverty, and other bad social conditions were eliminated, there would undoubtedly still be unwanted pregnancies, and women would still want abortions. The question is whether abortion is a morally permissible choice. It is hard to see how this question can be answered without considering the arguments of those who claim that killing fetuses is seriously wrong. Finally, while Dworkin's view is ingenious, it is doubtful that most pro-lifers at least would accept his reconceptualization. For them, it is not a matter of how best to respect the sanctity of life, but rather a matter of preventing the murder of innocents. Their opposition to abortion cannot be taken seriously if this claim is not addressed.

THEORIES OF MORAL STATUS

Genetic Humanity

According to Roman Catholic teaching, abortion is murder. It is murder because it is "the deliberate and direct killing ... of a human being in the initial phase of his or her

existence . . .” (9,10). In other words, there can be no doubt that a human fetus has the moral status of any other human being, because “from the time that the ovum is fertilized, a life is begun which is neither that of the father nor the mother; it is rather the life of a new human being with his own growth. *It would never be made human if it were not human already*” (9, p. 22, emphasis added). Despite the fact that a fertilized egg does not look like the human beings we know, and has none of the characteristic attributes of born human beings, its possession of its own genetic code makes it a separate, unique, and individual human person. It is worth noting that this argument, though religious in origin, is not religious in nature. It makes no appeal to religious notions, like God or the soul. Rather, the humanity of the fetus is based on its being biologically or genetically human. John Noonan makes the same point when he says, “The positive argument for conception as the decisive moment of humanization is that at conception the new being receives the genetic code. . . . A being with a human genetic code is man” (11, p. 264). All and only genetic human beings have full moral status, and this full or human moral status begins at conception.

Birth

Conception as the point at which a human being acquires moral status is often defended by showing the arbitrariness of any other stage. Consider, for example, birth. Live birth is often taken to be a significant landmark both in Anglo-American law and in religious traditions, such as Judaism. At birth, the fetus becomes an infant, separate from its mother and no longer physiologically dependent on her. Moreover, birth is a precise moment, noted on birth certificates. However, conservatives argue that the difference between a newborn moments after birth and a fetus moments before birth is insignificant. How can location alone determine moral status?

Viability

Moving back into gestation, some regard viability — the stage at which the fetus can survive outside the womb, albeit with artificial aid — as having moral significance. In *Roe* v. *Wade*, the Supreme Court chose viability as the point at which there was no longer a constitutional right to an abortion:

> With respect to the State’s important and legitimate interest in potential life, the “compelling” point is at viability . . . State regulation protective of fetal life after viability thus has both logical and biological justifications. If the State is interested in protecting fetal life after viability, it may go so far as to proscribe abortion during that period except when it is necessary to preserve the life or health of the mother . . . (12).

Of course, the Court’s decision was not intended to mark the point at which the fetus acquires *moral* status, since that is not a legal issue. Nor did the Court stipulate that a fetus undergoes a change in *legal* status at viability. *Roe* v. *Wade* did not hold that a fetus becomes a legal person, with a right to life, at viability. Rather, viability marks the point at which the *state’s* interest in protecting fetal life becomes “compelling,” namely the point at which the state can outweigh the woman’s privacy right to make the abortion decision. However, unless viability has moral significance, it would be a completely arbitrary point at which to allow states to ban abortion. It is clear that the Court did not regard it as an arbitrary point, but rather considered viability as a dividing line with “logical and biological” justification. But what exactly is this justification, and what does it have to do with fetal moral status?

The Court noted that at viability the fetus has the capacity for meaningful life outside the mother’s womb. However, as several commentators have noted, this is not a justification for extending protection to it, but rather an explanation of what viability means. The question is why the capacity for meaningful, independent existence bears on moral status. Nancy Rhoden has suggested that the Court focused on viability because, especially in 1973, the capacity for independent existence was connected with late gestation and fetal development (13). In other words, by the time a fetus becomes viable, it shares many characteristics with infants, who are entitled to the law’s protection. A late-gestation fetus looks like a baby, it can probably feel and hear, it sucks its thumb. It is sufficiently like a newborn that all of our protective feelings for babies “kick in” and incline us to extend the same protection to viable fetuses.

None of this is persuasive to the conservative, who asks, “Why should the dependence of the nonviable fetus on its mother deprive it of human moral status?” Moreover, a viable fetus has a chance at survival if removed from the uterus. The fact that a nonviable fetus cannot survive outside the womb is all the more reason not to eject it. Thus the “logical and biological justification” for permitting abortion of previable fetuses is puzzling. In any event, the conservative does not regard any part of fetal development as having special moral significance. If allowed to grow and develop, the fetus will acquire a nervous system, organs, a brain. It will begin to move and look like a born human being. Eventually it will have sensations and hear sounds. None of these stages makes the fetus a human being. It is already human, and each landmark is nothing more than a stage in its development.

Brain Birth

Most conservatives on abortion base moral status on being human, but not all believe that human life begins at conception. For example, Baruch Brody argues that a functioning brain is essential for being human (14). His position is based on a parallel with the end of human life. If a human being dies and goes out of existence when the brain irrevocably and completely stops functioning, then human life can be said to start when the fetal brain begins to function, or when brain waves can be detected, somewhere around six weeks after fertilization. It is not simply the parallel with the current criteria for death that motivates Brody’s focus on brain birth. In addition brain function is the biological basis for consciousness, thought, and feeling. The importance of this for Brody is indicated when he says, “One of the characteristics essential to a human being is the capacity for conscious experience, at least at a primitive level. Before the sixth week, as far as we know, the fetus does not have

this capacity. Thereafter, as the electroencephalographic evidence indicates, it does. Consequently, that is the time at which the fetus becomes a human being" (14, p. 83). However, from the fact that brain waves can be detected in a six-week old fetus, it does not follow that it can feel. While there is some disagreement on this point (15), and it is difficult to pinpoint the exact time at which the fetus becomes aware of its surroundings, most scientists believe that more than brain waves are necessary before the fetus has the capacity for conscious experience. It seems likely that sentience does not occur until well into the second trimester (probably between 22 and 24 weeks) as prior to that time the neural pathways are not sufficiently developed to be able to transmit pain (or any experiential) messages to the fetal cortex (16,17). So, if it is conscious experience that marks the beginning of human life, then this very likely does not occur until the second trimester of pregnancy. The occurrence of brain waves is certainly a necessary part of this development, but it is not clear why the beginning of brain function should have particular moral significance. Certainly most pro-lifers do not view "brain birth" as having any more significance than any other developmental stage in the life of the fetus.

Personhood

A different approach to the moral status of the fetus focuses on the fact that fetuses, while certainly genetically human, are not persons. Persons are essentially defined or recognized, not by their biology, but by certain psychological features, such as consciousness or an awareness of their surroundings, sentience, self-consciousness, thought, and the use of language. Fetuses, especially at the beginning of pregnancy, when 90 percent of abortions occur, have none of these characteristics. How, proponents of a personhood view ask, can a clump of cells be compared to an individual who feels, thinks, worries, enjoys, and so forth? It is persons who have a special moral status and a right to life. The mistake conservatives make is to equate being human with being a person.

This is a natural enough mistake, since all the persons we know are human beings. In contrast, personhood proponents argue that the two concepts are distinct in theory, however much they coincide in our experience. To see this, think about encountering an intelligent alien like *E.T.* in the movie of that name. Certainly E.T. is not human in a genetic or biological sense. He is not a member of the species *Homo sapiens*. Nevertheless, if we should meet someone like E.T., we would have no hesitation in according him the moral status of a person because, whatever his species membership, he resembles us in morally significant ways. That is, he is conscious, sentient, rational (indeed, far more advanced than human beings), self-conscious, uses language, and is a moral agent. Mary Anne Warren (18) claims that it is these features that have moral significance, not membership in a particular species. Why should a merely biological category make a moral difference? Or, as Don Marquis (an opponent of abortion who nevertheless criticizes the species criterion of moral status) puts the point, "Why ... is it any more reasonable to base a moral conclusion on the number of chromosomes in one's cells than on the color of one's skin?" (19, pp. 26–27). On the other hand, if "human being" is taken to be a *moral* category, then the fetus's claim to be a human being cannot be a *premise* in the antiabortion argument, for that is precisely what is to be proved. The conservative argument appears either to be based on an arbitary, morally irrelevant category (species membership) or to beg the question. To say this is not to say that the genetic humanity criterion has been shown to be wrong, only that it has not been adequately defended. Perhaps membership in the human species can be shown to have moral significance so that the conservative argument can be rendered noncircular.

What are the implications of a personhood view of moral status for the fetus? Warren does not claim that possession of *all* the person-making characteristics is necessary to be a person. Her point is rather that a being who possessed *none* of the relevant characteristics would not be a person. Indeed, she thinks that anyone who claimed that a being who possessed none of the person-making traits was a person all the same "would thereby demonstrate that he had no notion at all of what a person is — perhaps because he had confused the concept of a person with that of genetic humanity" (18, p. 68). On Warren's view, early-gestation fetuses are certainly not persons, as they lack all of the characteristics of persons. But what are we to say about more developed fetuses, who have some of the characteristics of persons, such as sentience and a rudimentary form of consciousness? Warren notes that even late-gestation fetuses are not fully conscious, in the way that an infant of a few months is, and that it cannot reason or use language, and so forth. "Thus, in the *relevant respects*, a fetus, even a fully developed one, is considerably less personlike than is the average mature mammal, indeed the average fish" (18, p. 69). Warren concludes that a fetus at any stage cannot have more of a right to life than a newborn guppy "and that a right of that magnitude could never override a woman's right to obtain an abortion, at any stage of her pregnancy" (18, p. 69).

There are two central problems with a personhood criterion of moral status. The first is that it applies not only to fetuses but to other human beings, notably, newborns. If abortion is not seriously morally wrong, is infanticide equally morally neutral, something best left to parents to decide? Warren's response to this is to offer consequentialist reasons for keeping infanticide illegal and to acknowledge that infanticide need not be seriously wrong in a society that cannot care for all the infants who are born. Whatever the merits of this argument, the personhood criterion seems to leave out many individuals most of us thought were full-fledged members of the moral community: elderly senile people, for example, and people with severe mental disabilities. Are we to say that their deficits in rationality, language usage, and so forth, deprive them of "human moral status" and that therefore killing them is not seriously morally wrong? One way to avoid this unpleasant conclusion is to set the requirements for personhood relatively low. Perhaps consciousness and sentience will do. However, that will mean that a great many animals will qualify as persons whom it is presumably wrong to kill. This is a possible moral view, but one that requires radical revision

of ordinary moral thought and practice. On the other hand, if the requirements for personhood are set relatively high, and rationality and language use are necessary conditions, that will keep out the animals, but at the cost of denying full moral status to many human beings, and not just fetuses.

The other problem with Warren's personhood view is that it suffers from the same defect she criticized in the conservative argument (20). According to Warren, the conservative confuses genetic humanity with moral humanity (or moral personhood). This confusion leads the conservative to think, wrongly, that just because the fetus is genetically human, it must be morally human, that is, have the moral status that human persons have, including a right to life. What is missing in the conservative's argument is an explanation of why species membership should endow a being with a particular moral status. However, Warren can be criticized for making the same mistake, for she does not explain what it is about personhood that endows a being with moral status. As Marquis explains the criticism:

> The principle "Only persons have the right to life" also suffers from an ambiguity. The term 'person' is typically defined in terms of psychological characteristics, although there will certainly be disagreement concerning which characteristics are most important. Supposing that this matter can be settled, the pro-choicer is left with the problem of explaining why *psychological* characteristics should make a *moral* difference (19, p. 27).

Warren says that it is "self-evident" that descriptive persons are moral persons, but that equation is no more self-evident than the conservative's claim that genetic humans are moral humans (i.e., entitled to be treated in certain ways). This is not to say that Warren's claim cannot be defended, but rather to say that it is not self-evident, and an argument connecting the psychological properties of personhood and moral status is needed. Even if such an argument can be given, it does not follow that moral personhood is limited to descriptive persons. The possibility remains open that it is justifiable to confer *normative personhood* (including a right to life) on human beings, such as babies and those with severe mental deficits, who lack the psychological properties of descriptive persons.

Being a Person and Having a Right to Life

Like Warren, Michael Tooley thinks that personhood is essential to moral status, but unlike Warren, he provides an argument to show why only descriptive persons can have a right to life (21). Tooley starts with Joel Feinberg's analysis of right-bearers as beings who can have interests (22). Interests are necessary for rights because the function of rights is to protect interests; if a being had no interests, there would be nothing to protect and the ascription of a right meaningless. Tooley then takes Feinberg's view one step further, arguing that particular rights are connected with specific sorts of interests. According to the "particular-interests principle," an individual cannot have a right to R unless it is capable of having an interest in R. The question then is what is required for a being to have an interest in R. This depends, Tooley maintains, on the sort of thing that R is. Consider the interest in not being subjected to painful stimuli. To have an interest in that requires only that the being be capable of experiencing pain as a disagreeable sensation. Thus we could intelligibly ascribe a right not to be subjected to painful stimuli to any sentient being. Other rights require more conceptual abilities. It would be absurd to ascribe a right of freedom of expression to a cow, because cows have no interests that can be furthered by such a right. Now what about life? It might be argued that life is in the interest of all living things, but this is just what Tooley wants to deny. He maintains that the right to life protects the interest in one's own continued existence, and that therefore only beings who can have a concept of their own continued existence can have a meaningful right to life. Tooley's view is much more stringent than Warren's. It limits a right to life to those who, first, are able to think of themselves as continuing to exist into the future and, second, have desires about that future, in particular, that it exist. This rules out not only human fetuses and most animals, but also babies and young children. It is not clear when children obtain a conception of themselves as existing in the future. It is probably not before they become language users, somewhere around the age of two.

The practical implications of Tooley's view (that it makes infanticide morally permissible) are not the only objection to it. In addition it can be objected that a right to life can be meaningfully ascribed to any being whose life is a good to it now, even if it lacks the capacity to envisage, and have desires about, a future existence. Animals, babies, and severely retarded adult human beings can enjoy their lives; why then cannot we preserve their lives *for their own sake?* If the reason for not killing them is that their lives are a good *to them*, this suggests that they have an interest in living, which can be the basis for ascribing to them a right to life.

Potential Personhood

Some pro-lifers do not regard genetic humanity as intrinsically significant, but rather as an indication that the being will become a descriptive person. An embryo does not now have any of the properties of a person. It is not even sentient or conscious, much less capable of communicating or relating to other persons. However, even an embryo is potentially just like us. If left alone (i.e., not aborted), it will grow and develop into a human person. Therefore we ought not to thwart its natural development. "On its strongest interpretation," Stephen Buckle explains, "the argument is thought to establish that we should treat a potential human subject as if it were already an actual human subject" (23, p. 227).

Is it in fact true that embryos, if not deliberately aborted, will develop into persons? A great deal has been learned about the rate of miscarriage in the last 20 years. When John Noonan was writing in 1970, he claimed that only 20 percent of pregnancies ended in spontaneous abortion. This figure still holds for pregnancies that are physiologically recognized. However, it is now thought that up to 75 percent of all human conceptions are aborted

spontaneously (24). Many of these spontaneous abortions occur before the woman realizes she is pregnant. With such a high rate of pregnancy loss, can it be maintained that every embryo is a potential person? The vast majority do not develop into persons, and would not develop into persons, even if allowed to develop naturally. This is one reason why implantation might be chosen by potentiality theorists as the moment at which the conceptus attains full moral status. After implantation, the prospects for live birth improve considerably.

There are other problems with arguments based on potentiality. The first is known as "the logical problem with potentiality." It is directed at the strongest version of the potentiality argument, which claims that the potential of the fetus to become a person gives it now the rights of a person. But, it may be objected, why should mere potential convey actual rights? As Stanley Benn has put it, "A potential president of the United States is not on that account Commander-in-Chief" (25, p. 143).

The logical problem can be avoided if the claim is weakened. It is not that the embryo or fetus *now* has the right to life. Rather, the claim is that because we think that the lives of persons are valuable and deserving of protection, so too we ought to recognize the value of entities that have the potential to become persons. While this might not accord to fetuses a full-fledged right to life, it would at least require that the reasons for killing a fetus be substantial ones.

A serious problem for potentiality arguments is that they seem vulnerable to a *reductio ad absurdum*. If the objection to abortion is that it kills potential person, why cannot the same complaint be made of contraceptive techniques that kill sperm, or prevent sperm and egg from joining? Why isn't an unfertilized ovum also a potential person? John Harris makes the point this way (although he refers to "human being" rather than "person"):

> To say that a fertilized egg is potentially a human being is just to say that if certain things happen to it (like implantation) and certain other things do not (like spontaneous abortion), it will eventually become a human being. But the same is true of the unfertilized egg and the sperm. If certain things happen to the egg (like meeting a sperm) and certain things happen to the sperm (like meeting an egg) and thereafter certain other things do not (like meeting a contraceptive), then they will eventually become a new human being (26, pp. 11–12).

If abortion is seriously wrong because it kills a potential person, then using Delfen foam (a spermicide) is mass murder! Indeed, even abstinence would have to be justified, since failing to have intercourse, at least during a woman's fertile period, would prevent the development of a new human being. Since virtually all potentiality theorists wish to differentiate between contraception (which they regard as morally neutral) and abortion (which they regard as seriously wrong), they must explain why an embryo is a potential person in a way that a gamete is not.

Some theorists look to probability to defend the claim that an embryo is a potential person, but a gamete is not. Even if a fertilized egg has only a 25 percent chance of becoming a person, this does not compare with a sperm's chance: about one in 200 million. A given ovum has a better chance to become a person than a given sperm, but still much less chance than that of a fertilized egg. However, it is not clear that potential should be understood in terms of probability. As Steinbock points out, "Is not every entrant in a lottery a potential winner, even if the odds of winning are extremely low?" (3, p. 63)

Others regard potentiality not in terms of the odds of success but rather in terms of natural development. Many fertilized eggs do grow and develop into embryos, fetuses, and babies. A gamete, on the other hand, is not growing or developing into anything (27,28). However, basing potential personhood on what fertilized eggs become in the natural course of events has notable consequences for extracorporeal embryos. Since they cannot develop into persons without human intervention (i.e., being placed in a uterus), extracorporeal embryos are not potential persons, and presumably, it is not seriously wrong to kill them. This conflicts with the opinion of many pro-lifers and the Catholic Church that extracorporeal embryos have the same moral status as embryos in a uterus. Indeed, some have regarded frozen embryos as "pre-born children" (29).

Some defenders of a potentiality principle try to distinguish between a gamete and a zygote by saying that prior to fertilization, no particular individual exists. Once the complete human genome is present, there is a new human being, the same individual that will be born, grow up, and die. However, as Mary Anne Warren notes, this claim can be disputed on empirical grounds:

> It is not clear that the zygote is the same organism or proto-organism as the embryo that will later develop from it. During the first few days of its existence, the conceptus subdivides into a set of virtually identical cells, each of which is "totipotent"—capable of giving rise to an embryo. Spontaneous division of the conceptus during this period can lead to the birth of genetically identical twins or triplets. Moreover, it is thought that two originally distinct zygotes sometimes merge, giving rise to a single and otherwise normal embryo. These facts lead some bioethicists to conclude that there is no individuated human organism prior to about fourteen days after fertilization, when the 'primitive streak' that will become the spinal cord of the embryo begins to form (1, pp. 203–204).

In this view, an implanted embryo is a potential person, while neither a gamete nor a newly fertilized ovum is. But some theorists argue that while the implanted embryo may be identified with some *particular* person, in a way that neither the zygote nor the constituent gametes are, nevertheless gametes and zygotes are potential persons. Potentiality is one thing; uniqueness or identity another (30).

A Future Like Ours

A variation on the potentiality principle is offered by Don Marquis in "Why Abortion Is Immoral" (19). According to Marquis, the reason why killing people is generally wrong is that killing deprives the individual of a valuable future, a future like ours (FLO). Killing a fetus by having an abortion deprives the fetus of its valuable future, and therefore abortion is (usually) seriously morally wrong. Marquis notes two ways in which his view differs from

traditional accounts. First, it is not "species-ist." Moral status does not depend on being genetically human but rather on having a valuable future like ours. If there are members of other species elsewhere in the universe who have valuable futures like ours, then it would be seriously wrong to kill them. For that matter, it is possible that some nonhuman animals have FLO, and that killing them is seriously morally wrong. Marquis leaves indeterminate precisely what FLO consists in, but presumably he has in mind the kinds of capacities that make us persons: rationality, self-consciousness, the ability to have relationships with others, and so forth. Therefore, although his account does not refer explicitly to the wrongness of killing persons or potential persons, this is implicit in his account, since the beings it is seriously wrong to kill (i.e., those who have FLO) are persons or potential persons. If a being is neither a person, nor capable of developing into a descriptive person, then presumably it does not have a valuable future like ours. Second, Marquis's view differs from a sanctity of life approach that holds that killing people is always wrong. According to Marquis, killing someone who does not have a valuable future is not necessarily wrong. Thus Marquis's account is compatible with voluntary euthanasia and even nonvoluntary euthanasia of infants and fetuses whose lives will be filled with suffering or empty of the things that make life worth living.

An objection to Marquis's view is that it is vulnerable to the same *reductio* as other potentiality views. If it is wrong to kill a fetus because of its FLO, why is it not equally wrong to kill gametes? Why do not gametes have a valuable future? Of course, it is true that most gametes do not have much of a future at all. They pass out of or are reabsorbed into the body. However, use of a contraceptive prevents those few gametes that might develop into persons from doing so. It would seem that contraception should be somewhat morally problematic, in Marquis's view. He, however, denies this. We can ascribe a valuable future to a fetus because the fetus is identified with the person it becomes. It has the *same* future as the born individual. However, the born individual does not share a future with the egg and sperm that conjoined to form it. Neither the egg nor the sperm is the person; therefore neither has his or her future. As Marquis expresses it:

> If I were the same individual as a sperm and also the same individual as an ovum, then a particular sperm and a particular ovum were the same individual. This is obviously false. It follows that any argument that I was once a sperm (or an unfertilized ovum) is unsatisfactory. Therefore, the contraception objection fails (30).

However, why should the fact that I am not identical with either the sperm or the egg entail that they do not have valuable futures, of which they would be denied if they were killed or otherwise prevented from conjoining? Marquis thinks that neither gamete can have a future because it is only when the two conjoin that there is anything that has a valuable future. The valuable future is that of the new being, and it is not identical with either of its two components. However, it is not clear why only the new being can be said to have a valuable future. To be sure, neither gamete can have a valuable future all by itself, but why does that matter? It might be objected that Marquis has confused identity with having a valuable future.

One of the interesting implications of Marquis's account is the possibility that pre-implantation embryos lack FLO, and therefore lack moral status. Since a pre-implantation embryo might turn out to be two or three people, it cannot be identified with any particular individual. If having FLO depends on identification with a particular future person, then pre-implantation embryos do not have FLO. This means that killing a pre-embryo using a very early abortifacient, such as the morning-after pill, is not seriously wrong. In addition Marquis presumably would not oppose creating embryos either for research or possible implantation, and then discarding them. Only implanted embryos, which have the primitive streak and can no longer become twins, can be identified with the subsequent person, have FLO and are seriously wrong to kill. The fact that most conservatives would be unwilling to accept these implications is not an argument against the FLO theory. However, it remains an open question whether it, like potentiality theories generally, are vulnerable to the contraception objection.

Sentience-Based Views

Sentience is the ability to experience pain or pleasure. It is relevant to moral status because we normally assume that it is wrong to inflict pain without a good reason. Many pro-lifers seem to base their objection to abortion on the premise that abortion causes the fetus to suffer. One pro-lifer was quoted as saying that abortion is "mean." Since most methods of abortion require the embryo or fetus to be ripped from the uterine wall, and often cause it to be torn apart, the idea that abortion hurts the fetus is not surprising. However, as we have seen, it is extremely unlikely that a fetus in the first trimester can experience pain or anything else.

Rejection of unnecessary infliction of pain suggests that sentience is a sufficient condition of moral status. However, the infliction of pain is not the only kind of action with moral importance. Killing is also something that often requires justification, even when done painlessly. We need a deeper reason for thinking that sentience is a necessary, as well as sufficient, condition of moral status. The deeper reason comes from the connection between moral status and interests, and the connection between interests and sentience.

To say that a being has moral status is to say that it has moral claims against us. This in turn suggests that we should do, or forbear from doing, certain things for its sake. We are required to consider its welfare in deciding what to do. The next question is what kinds of beings can have a sake or a welfare. For some philosophers, anything that can be protected or preserved can have a welfare. Others limit a welfare to beings to whom it can matter how they are treated. In this view, it may be morally

wrong to destroy the environment, but it is not a wrong *to* the environment because the environment does not, indeed, cannot, care what is done to it. The environment has no stake in what happens to it. Because nothing can matter to it, it has no interests. If it has no interests, then its interests cannot be considered, and therefore it lacks moral status.

The conceptual connection between moral status and the possession of interests seems self-evident. To accord something moral status is to require that its interests be considered. If a being has no interests, no sake, and no welfare, it is hard to see how it could have moral claims on us. More controversial is the claim that sentience is necessary for a being to have interests. What is it about the ability to experience pain and pleasure that enables a being to have interests? The intuitive idea is that sentient beings care about what happens to them, while nonsentient beings do not. However, perhaps it is not sentience that enables beings to have interests but rather simply consciousness: the ability to have experiences. Mary Anne Warren suggests this when she writes:

> One can imagine a being that has conscious experiences of many sorts, but that never experiences pleasure or pain, or any other positive or negative feeling, mood, or emotion. Such a being would be conscious, but it would not be sentient. Data, the brilliant and personable android of the television series *Star Trek: The Next Generation*, is described by himself and other characters as such a being. Although he is conscious, rational, morally responsible, and highly self-aware, his programming includes no capacity to experience pain, pleasure, or emotion . . . such a being would have strong moral status by virtue of its moral agency; but it could not have any moral status that is contingent upon sentience (3, p. 56).

The ability to experience physical pain or pleasure does not seem essential to having interests. There are human beings who lack the ability to feel pain; they still have all kinds of interests, including an interest in continuing to live. Sentience, understood in purely physical terms, does not seem to be necessary for moral status. However, Data is alleged to have no feelings or emotions at all. This being the case, does it matter to him what happens to him? Does he care if he is destroyed or protected? And if he is indifferent to what happens to him, can preserving him be said to be in *his* interest? It might be argued that the connection between sentience (broadly understood to include emotions and feelings) and interests is not so easily broken.

Tom Regan argues that the belief that sentience is necessary for the possession of interests stems from a failure to distinguish between two senses of "interest." One sense of "interest" refers to what individuals *take* an interest in; what they are interested in or care about, what matters to them. It seems clear that nonconscious, nonsentient beings cannot have interests in this sense. However, there is another sense of "interest" that refers to what is *in* a being's interest. That these are not the same is easily shown. It can be *in* someone's interest to give up smoking, exercise regularly, and eat moderately, and yet the person in question might have no interest in doing so. Conversely, people often have an interest in things that are not, on balance, in their interest. The question is whether it makes sense to talk about what is *in* the interest of beings that do not *take* an interest in anything. Certainly we do sometimes talk this way, recommending that certain actions be taken "in the interest of preserving the environment." But it does not follow that the interests in question are those of the environment, nor that an inanimate object can have a good or welfare of its own in the same sense as a being that cares, if only in the most rudimentary sense, about what is done to it.

Some sentience-based theorists regard sentience alone as having moral significance. For example, Peter Singer is a "sentience-only" proponent, whose theory of moral status is inspired by that of Bentham and Sidgwick. According to Singer, the comparable interests of all sentient beings should be given equal weight in our moral deliberations (31). This does not entail treating all sentient beings alike, since their needs and interests will differ. It does not even mean valuing the lives of all sentient beings equally. Singer acknowledges that the life of a human — that is, a rational, self-conscious, morally autonomous agent — may properly be considered more important than the life of a mouse. However, he thinks that there is no justification for valuing the pain experienced by a person over the (comparable) pain experienced by a mouse. Pain is pain, no matter who feels it.

Not all sentience-based theorists accept the principle of equal consideration of interests. For example, L.W. Sumner, who thinks that sentience is a necessary and sufficient condition of moral status, argues that both sentience and moral status come in degrees (32). The moral status of a being is proportional to its degree of sentience. This is supposed to explain the intuitive view that the interests of people count for more than the interests of mice. However, it is not clear why rationality, self-consciousness, and moral agency should be conceived of as degrees of sentience. Nor does it seem impossible that a being lacking in those morally relevant features could nevertheless be intensely sensitive to pain and pleasure. Bonnie Steinbock does not accept the degrees of sentience view advocated by Sumner; she thinks that there can be features aside from sentience that are relevant to moral status. Thus she argues that while the possession of interests is a necessary condition of moral status, and potential personhood by itself does not endow beings who lack interests with moral status, the potential of a sentient being to become a person must be regarded as enhancing its moral status. Both human infants and nonhuman animals have minimal moral status as sentient beings, but the degree of moral status possessed by human infants is greater than that of nonhuman animals because the infants are potential persons. Another factor is relationships with others. Some humans, due to brain defects, will not develop into descriptive persons, capable of language, reasoning, and moral responsibility. Nevertheless, they remain normative persons. In part, this is because of the place they occupy in a network of affections. A retarded child does not cease to be someone's son or daughter, or loved the less because

of that. Their moral status does not change because they will not fulfill normal human potential. And even if the parents reject the child, society should take on this caring role, out of kindness and compassion for the helpless child who needs our care.

The Multi-Criterial View

Steinbock's suggestion that there might be features besides sentience relevant to the degree of moral status is taken up explicitly by Mary Anne Warren, who argues for a "multi-criterial" view (1). She rejects uni-criterial views, whether based on life, sentience, or personhood, as simplistic and inconsistent with elements of commonsense morality that we cannot reasonably be expected to jettison. Her multi-criterial approach consists of seven principles that she regards as implicit elements of commonsense morality. The principles that might be thought relevant to fetal moral status are respect for life and the transitivity of respect. (The anticruelty principle applies only to beings who can be treated cruelly, that is, sentient beings. It does not apply to early-gestation fetuses.) The respect for life principle accords some moral status to fetuses, but not enough for full moral status. Moreover, the reasons why women seek abortions are sufficiently compelling to justify the destruction of a living thing that is not yet sentient and not yet a member of a human social community. The transitivity of respect principle requires us, within the limits of the other principles and to the extent that it is morally feasible, to respect other people's attributions of moral status. Since some people do regard fetuses as having full moral status, this justifies regarding fetuses as having some moral status. However, the transitivity of respect principle is limited by the basic moral rights of moral agents. Warren concludes that "although the fetus gains in moral status as it becomes increasingly likely to be capable of sentience, until it has been born it cannot be accorded a fully equal moral and legal status without endangering women's basic rights to life and liberty" (1, p. 222).

BIBLIOGRAPHY

1. M.A. Warren, *Moral Status: Obligations to Persons and Other Living Things*, Clarendon Press, Oxford, UK, 1997.
2. K. Baier, *The Moral Point of View*, Random House, New York, 1965.
3. B. Steinbock, *Life Before Birth: The Moral and Legal Status of Embryos and Fetuses*, Oxford University Press, New York, 1992.
4. T. Murray, *The Worth of a Child*, University of California Press, Berkeley, 1996.
5. K. Luker, *Abortion and the Politics of Motherhood*, University of California Press, Berkeley, 1984.
6. R. Dworkin, *Life's Dominion: An Argument about Abortion, Euthanasia, and Individual Freedom*, Knopf, New York, 1993.
7. J.J. Thomson, *Philos. Public Affairs* **1**(1), 47–66 (1971).
8. For example, S. Harding, in S. Callahan and D. Callahan, eds., *Abortion: Understanding Differences*, Plenum, New York, 1984.
9. John Paul II, *Evangelium Vitae*, Encyclical Letter, August 16, 1993. Copyright Libreria Editrice Vaticana, Rome, 1993, reprinted in Dwyer and Feinberg (10, pp. 21–23). Page references are to this edition.
10. S. Dwyer and J. Feinberg, eds., *The Problem of Abortion*, 3rd ed., Wadsworth, Belmont, CA, 1997.
11. J.T. Noonan, Jr., in J.T. Noonan, Jr., ed., *The Morality of Abortion: Legal and Historical Perspectives*, Harvard University Press, Cambridge, MA, 1970, reprinted in J.D. Arras and N.K. Rhoden, eds., *Ethical Issues in Modern Medicine*, 3rd ed., Mayfield Publishing, Mountain View, CA, 1989, pp. 261–265. Page references are to this edition.
12. *Roe v. Wade*, 410 U.S. 113, 163 (1973).
13. N.K. Rhoden, *Yale Law J.* **95**(4), 639–697 (1986).
14. B. Brody, *Abortion and the Sanctity of Life*, MIT Press, Cambridge, MA, 1975.
15. P. McCullagh, *The Foetus as Transplant Donor: Scientific, Social and Ethical Perspectives*, Wiley, Chichester, 1987.
16. K.J.S. Anand and P.R. Hickey, *N. Engl. J. Med.* **317**(21) (1987).
17. M.J. Flower, *J. Med. Philos.* **10**(7), 237–251 (1985).
18. M.A. Warren, *The Monist* **57**, 1973, reprinted in Dwyer and Feinberg (10, pp. 59–74). Page references are to this edition.
19. D. Marquis, *J. Philos.* **86**(4), 183–202 (1989), reprinted in Dwyer and Feinberg (10, pp. 24–39). Page references are to this edition.
20. J. Feinberg, in T. Regan, ed., *Matters of Life and Death: New Introductory Essays in Moral Philosophy*, 2nd ed., Random House, New York, 1986, pp. 256–293.
21. M. Tooley, *Abortion and Infanticide*, Clarendon Press, Oxford, UK, 1983.
22. J. Feinberg, in W.T. Blackstone, *Philosophy & Environmental Crisis*, University of Georgia Press, Athens, 1974.
23. S. Buckle, *Bioethics* **2**(3), 227–253 (1988).
24. J.W. Knight and J.C. Callahan, *Preventing Birth: Contemporary Methods and Related Moral Controversies*, University of Utah Press, Salt Lake City, 1989.
25. S.I. Benn, in J. Feinberg, *The Problem of Abortion*, 2nd ed., Wadsworth, Belmont, CA, 1984, pp. 135–144.
26. J. Harris, *The Value of Life: An Introduction to Medical Ethics*, Routledge & Kegan Paul, London, 1985.
27. R. Hursthouse, *Beginning Lives*, Basil Blackwell, Oxford, UK, 1987.
28. R. Warner, *Soc. Theory Pract.* **3**(4) (1974), revised and reprinted in R.A. Wasserstrom, ed., *Today's Moral Problems*, 2nd ed., Macmillan, New York, 1979.
29. P. Singer and K. Dawson, *Philos. Public Affairs* **17**(2), 87–104 (1988).
30. D. Marquis, in R. Edwards and E.E. Bittar, eds., *Bioethics for Medical Education*, JAI Press, Greenwich, CT, 1999.
31. P. Singer, *Practical Ethics*, Cambridge University Press, Cambridge, UK, 1979.
32. L.W. Sumner, *Abortion and Moral Theory*, Princeton University Press, Princeton, NJ, 1981.
33. P. Singer, *Animal Liberation*, Avon Books, New York, 1975.
34. T. Regan, *The South. J. Philos.* **14**, 485–498 (1976).

See other entries GENE THERAPY, ETHICS, GENE THERAPY FOR FETUSES AND EMBRYOS; see also REPRODUCTION, ETHICS entries; REPRODUCTION, LAW, REGULATION OF REPRODUCTIVE TECHNOLOGIES.

REPRODUCTION, ETHICS, PRENATAL TESTING, AND THE DISABILITY RIGHTS CRITIQUE

ERIK PARENS
The Hastings Center
Garrison, New York

ADRIENNE ASCH
Wellesley College
Wellesley, Massachusetts

CYNTHIA POWELL
University of North Carolina
Chapel Hill, North Carolina

OUTLINE

TECHNOLOGICAL CONTEXT

History of Prenatal Testing

The oldest reference of prenatal diagnosis is for anencephaly (absent skull and brain) diagnosed by X rays in 1917 by James T. Case. The first amniocentesis is attributed to Schatz in 1883 for the purpose of treating hydramnios (excessive amniotic fluid) (1). Amniocentesis to detect erythroblastosis fetalis (complication of Rh incompatibility) began in the 1950s. Prenatal testing with amniocentesis for genetic disorders began in 1955 when it was discovered that the sex of a human fetus could be predicted by analysis of fetal cells in amniotic fluid (2–5). Initially the testing was used to determine fetal sex when a woman was at risk for having a child with an X-linked condition such as hemophilia (6). In this situation she would have a fifty percent risk of having an affected male infant. If the baby were female, she would have a fifty percent risk of carrying the gene, but would be clinically unaffected. The sex chromatin body (Barr body) could be identified in nondividing amniotic fluid cells. A male fetus (absent Barr body) could be identified and the pregnancy terminated despite a 50 percent probability that it would be unaffected. The first report of the procedure being done for this purpose was from Denmark (7).

It was reported in 1959 that Down syndrome is due to an extra chromosome 21 and use of amniocentesis to detect fetal chromosome abnormalities began in the 1960s (8–11). Widespread use of amniocentesis for increased maternal age in the United States is partly attributed to lawsuit settlements in the late 1970s in cases where the patients had not been referred for testing and gave birth to children with disabilities. In 1983 the American Academy of Pediatrics and the American College of Obstetricians and Gynecologists recommended that all women over the age of 35 be offered amniocentesis; thus amniocentesis became a routine part of obstetric care (12). Other contributing factors to the rapid increase in demand for prenatal diagnostic services include the liberalization of abortion statutes, changes in cultural attitudes toward family size, and extensive media coverage and publicity given to prenatal diagnosis (13).

Chorionic villus sampling (CVS), aspiration of tissue that will become the placenta, was first done in Copenhagen in the late 1960s (14,15). Because of various technical problems it did not come into common use until the 1980s (16). This test has the advantage of being performed in the first trimester of pregnancy (9–13 weeks) as compared to amniocentesis, which is done in the second trimester (15–18 weeks). However, the spontaneous abortion rate following the procedure is approximately 1 percent, which is higher than for amniocentesis (1 in 300–400). There is also a higher rate of mosaicism (two or more cell lines with different chromosomal constitution) with CVS as compared to amniocentesis (0.5–1 percent for cultured CVS cells vs. 0.2–0.3 percent for amniocytes) (17). In most cases the abnormal cell line is confined to the placenta, so a follow-up amniocentesis is often needed to look for the possibility of true mosaicism in the fetus. CVS performed earlier than nine weeks may be associated with an increased risk of limb and facial malformations. And different from amniocentesis, it is not possible to measure alpha-fetoprotein levels with CVS; as a result CVS cannot be used to detect neural tube impairments. Moreover, because CVS is a technically more difficult procedure than is amniocentesis, fewer obstetricians offer CVS in their office. Not all tertiary care facilities or major medical centers offer it.

Screening for neural tube impairments was first demonstrated in 1972 by determining amniotic fluid concentrations of alpha-fetoprotein (AFP). Elevated levels were associated with spina bifida and anencephaly in the fetus (18). Elevated levels of AFP were also found in maternal blood serum when the fetus had an open neural tube (19). By the 1980s most women were being offered serum AFP screening during pregnancy. Maternal serum AFP (MSAFP) screening, ultrasound and amniocentesis are capable of detecting 100 percent of cases of anencephaly and 80 to 90 percent of cases of spina bifida (20).

In 1984 it was reported that AFP values in maternal serum were lower than expected when the fetus had Down syndrome or trisomy 18 (21). AFP measurement along with age of the mother were used to detect pregnancies with an increased likelihood of a chromosome abnormality in women under the age of 35. This technique could detect 20 percent of cases of Down syndrome in this population

who otherwise would not be identified as having an increased risk. Other biochemical markers were identified which varied from normal when a woman was carrying a fetus with Down syndrome or trisomy 18. These include unconjugated estriol and human chorionic gonadotropin. Use of these three markers to screen pregnancies is commonly referred to as the "triple screen" and is generally done at the sixteenth to eighteenth week of pregnancy. Abnormal values are followed by ultrasound to confirm dating. If dating is correct, amniocentesis is offered for definitive testing for chromosome abnormalities. Detection rates with use of the triple screen have been reported as 67 to 75 percent, with a false positive rate of 4 to 5 percent (22).

Ultrasound was first developed in World War I as sonar, to look for underwater submarines. Using it to look at fetuses began in the 1960s (23). Ultrasound is used for determining gestational age, locating structures prior to invasive testing procedures, and identifying structural abnormalities in the fetus. Conditions such as anencephaly, spina bifida, congenital heart defects, hydrocephalus, and kidney abnormalities are detectable with ultrasound. Most women in this country have at least one ultrasound during their pregnancy. Studies have not shown a risk to the fetus as a result of ultrasound, but pregnancy outcomes have not been shown to improve through routine ultrasound use.

Fetoscopy, the direct visualization of the fetus by using an optical instrument, was first used in 1954 but is now rarely done except for fetal skin biopsy. Fetoscopy was used in the past to obtain fetal blood samples, but percutaneous umbilical blood sampling (PUBS) — also known as cordocentesis is currently the method of choice. In this procedure ultrasound guides placement of a needle inserted through the maternal abdomen into the umbilical vein. PUBS is used for diagnosis of chromosome abnormalities, fetal infections, coagulation defects, hemoglobin and red cell disorders, metabolic and immunologic diseases.

The first reported prenatal diagnosis established by molecular genetic techniques was for alpha-thalassemia by means of linkage analysis in 1976, and for sickle cell anemia by analysis of gene mutation in 1978.

Today's Prenatal Tests

Prenatal diagnosis is available for hundreds of genetic conditions including chromosome abnormalities, inborn errors of metabolism, neural tube impairments, and single gene disorders. Ultrasound detects many different structural defects including hydrocephalus, congenital heart impairments, limb anomalies, skeletal dysplasias, and diaphragmatic hernias. In the most recent edition of Catalog of Prenatally Diagnosed Conditions there are 940 conditions listed that have been diagnosed prenatally. These include chromosome abnormalities, congenital malformations, dermatologic disorders, fetal infections, hematologic disorders, inborn errors of metabolism, tumors and cysts, and multiple congenital anomalies of unknown etiology.

Methods of prenatal diagnosis include maternal serum screening for biochemical markers to look for neural tube impairments as well as chromosomal aneuploidies (i.e., abnormal numbers of chromosomes). The triple screen detects up to 75 percent of fetuses with Down syndrome. However, these are screening tests and must be followed by additional procedures such as ultrasound in the case of neural tube impairments and amniocentesis or chorionic villus sampling for chromosome abnormalities.

Data from the Council of Regional Networks for Genetic Services (CORN) for 1989 estimated that 50 percent of pregnancies were being screened for MSAFP, this number is most likely now increased. The 1989 data from this group showed that increased maternal age was the most common indication for prenatal tests (62 percent), with abnormal MSAFP accounting for 14 percent, positive family history in 7 percent, previous spontaneous abortion or stillbirth 1 percent, abnormal ultrasound 1 percent, parental concern or anxiety 1 percent. "Other" and "unknown or unrecorded" accounted for an additional 11 percent (24). In a survey of a university-based cytogenetics laboratory over a five month period in 1997, of 476 amniotic fluid samples, 52 percent were obtained for increased maternal age, 34 percent were for abnormal triple screen, 7 percent were for abnormal ultrasound, 3 percent were for a family history of a genetic disorder, 1 percent were for DNA testing, and the remaining 3 percent were for elevated MSAFP, multiple miscarriages, and maternal anxiety. The DNA diagnostic testing included testing for achondroplasia, Rh, Kell antibody, sickle cell, and X-linked hydrocephalus.

Who Is Offering Tests

Most prenatal genetic testing is obtained through obstetricians in private practices or public health clinics. The remainder is obtained through tertiary referral centers such as university medical centers or private genetic centers. In a U.S. study that looked at discussions between obstetricians or nurse-midwives and their patients during their first prenatal visit, it was found that time devoted to discussion of genetic testing averaged 3.7 minutes $\pm$3.9 minutes (range 0–25.3 minutes). A comprehensive family history was not taken in any of the visits. Discussion of topics such as abortion or continuation of pregnancy if an anomaly were detected, or a description of the disorders for which testing was offered occurred in a minority of visits (24). Guidelines for Perinatal Care (1997) from the American Academy of Pediatrics and American College of Obstetrics and Gynecology does not recommend referral for genetic counseling for women of advanced maternal age because it states that the primary care physicians can explain the risks (25).

When an abnormality is detected through prenatal testing, information and counseling usually comes from an obstetrician, a genetic counselor, or a nurse. The information given varies depending upon the knowledge and experience the professional has of the specific condition. Many health professionals who offer prenatal genetic counseling have not had direct contact with children and adults with developmental disabilities and genetic disorders. Obstetricians often have had little contact with such patients since their medical school training, and even then this contact may have been

minimal. Few genetic counselors work in both prenatal and pediatric genetics. Few genetic counseling training programs give students an opportunity to work with developmentally disabled children and adults. Therefore, most women who receive information about a specific chromosome abnormality or genetic impairment in their fetus receive this information from health care providers without personal knowledge of the natural history and outcomes of the condition.

SOCIAL CONTEXT OF THE DISABILITY RIGHTS CRITIQUE

Outright Discrimination Against and Unexamined Attitudes About People With Disabilities

The history of discrimination against people with disabilities, including episodes of infanticide and compulsory sterilization, is long, ugly, and well documented (26–28). Even with such important steps as the passage of the Americans with Disabilities Act (ADA) and the Individuals with Disabilities Education Act (IDEA), discrimination is far from over. People with disabilities are still often treated as inferior to nondisabled people. As disability studies scholar Lennard Davis has observed, even the most educated of Americans, professors who make a living by writing about the nature of discriminatory practices and who decry discrimination against women, people of color, and other minorities, leave their attitudes toward people with disabilities largely unexamined. According to Davis, in the writings of these literary theorists, while "others" whose bodies are normal become vivid, others whose bodies are abnormal remain invisible (29).

It is not just practitioners of fashionable literary theory who sometimes harbor unexamined and discriminatory attitudes toward people with disabilities. The bioethics and medical literatures of the last decade too reveal misinformation and stereotypic thinking about what disability means for individuals, families, and society. Many clinicians and bioethicists take it for granted that health status is mostly responsible for the reduced life chances of people with a disability, largely ignoring the role of societal factors such as educational and employment discrimination. Furthermore, these clinicians and bioethicists often discount data indicating that people with disabilities and their families do not view their lives in solely or even predominantly negative terms (30); instead, they may insist that such data reflect a denial of reality or an exceptional ability to cope with problems (31–32).

People who make policy concerning the dissemination of genetic information have reached a consensus that the purpose of prenatal testing is to enhance reproductive choice for women and families — not to decrease the number of children with disabilities who are born. Some have acknowledged, however, that there is a tension between the goals of enhancing reproductive choice and preventing the births of children who would have disabilities. Writing about screening programs for cystic fibrosis in the pages of the *American Journal of Human Genetics*, medical geneticist A.L. Beaudet observed: "Although some would argue that the success of the program should be judged solely by the effectiveness of the educational programs (that is, whether screenees understood the information), it is clear that prevention of CF is also, at some level, a measure of a screening program, since few would advocate expanding the substantial resources involved if very few families wish to avoid the disease (33, p. 603). Beaudet acknowledges that, in tension with the genetic professional's stated goal of educating individuals (without any investment in the particular decision those individuals might reach), those who pay for such education do so in part with a view to reducing the number of — and costs associated with — children born with cystic fibrosis.

The profession of genetic counseling is based on a deep commitment to helping clients discover what course of action, upon reflection, is best for them. Some evidence suggests, however, that when disabilities are involved, both trained genetic counselors and others who deliver genetic information do not always live up to that commitment. A recent study designed to understand the experience of mothers who received a prenatal diagnosis of Down syndrome and chose to continue the pregnancy found problematic attitudes toward people with disabilities, evidenced in the way that medical professionals spoke to those prospective mothers. According to David T. Helm, one of the mothers who received a diagnosis of Down syndrome reported the following exchange:

> Obstetrician: *You have to move quickly. There is a doctor at [Hospital X] who does late-term abortions*. Mother: *No, I told you I'm not going to have an abortion*. Obstetrician: *Talk to your husband. You might want to think about it* (34, p. 57).

Because Helm only provides this portion of a longer exchange, the reader cannot confidently interpret the exchange he reports. Advising a patient to discuss a major life decision with her spouse is not prima facie problematic, much less discriminatory. According to Helm's interpretation (and the interpretation of the Disability Rights Community), however, these words reveal the physician's unwillingness or inability to respect this woman's already stated decision to continue the pregnancy with the fetus carrying a disabling trait. The reported exchange provides no evidence that this obstetrician understands the ways in which many families welcome and nourish — and are nourished by — children with Down syndrome.

Research has shown that obstetricians may be more likely than genetic counselors to urge particular actions upon their patients (35,36). Helm's study also reports, however, that some genetic counselors reacted negatively to women who intended to bear and raise children with Down syndrome. A woman who was told that the fetus she was carrying would have Down syndrome reported the following: "[The genetic counselor] treated me as though I couldn't accept this news, although I told her I could. She asked, 'What are you going to say to people when they ask you how you could bring a child like this into the world?'" (34, p. 57) Those words suggest that this counselor has not thought deeply about what disabilities mean for individuals who live with them and for their families. At least from what we learn of her from Helm, she does not seem to appreciate that

welcoming a child with Down syndrome into a family is not a decision that needs to be defended; she does not seem to appreciate that parental attitudes differ, that traits that matter a great deal to one couple may seem inconsequential to another. Such exchanges are probably not rare exceptions; similar examples can be found in other discussions of genetic counseling practices in the prenatal testing situation (37,38).

Nonetheless, many genetic counselors and physicians work extremely hard to live up to the central genetic counseling values of informed consent and nondirectiveness, and many of them are not only aware of but share the concerns voiced by the disability rights community. For example, at the New England Medical Center, women whose fetuses are diagnosed with Down syndrome are routinely scheduled to meet with a pediatric medical geneticist and a nurse clinician who specializes in the care of pediatric genetic patients. These women are scheduled to meet with pediatricians who specialize in genetics rather than obstetricians because pediatric geneticists understand better how Down syndrome influences the lives of children and their families. According to Dr. Diana Bianchi, who practices at the New England Medical Center, every attempt is made to introduce the pregnant woman and her partner to families who are raising infants, children, and/or young adults with Down syndrome. She reports that in her practice, only 62 percent of women who discover they are carrying a fetus with Down syndrome decide to have abortions. That rate of abortion upon a positive finding is believed to be relatively low. Disability critics point to such facts to suggest that when prospective parents obtain more accurate information about what life with disability is like, many realize that parenting a child who has a disability can be as gratifying as parenting a child who does not.

The disability critique proceeds from the view that discrimination results when people in one group fail to imagine that people in some "other" group lead lives as rich and complex as their own. The disability rights critics believe that everyone from literary theorists to bioethicists to obstetricians and genetic counselors are susceptible to such failures of imagination. Moreover they think that the desire of prospective parents to avoid raising children with disabilities may depend on that same failure.

Plurality of Disabling Traits and Plurality of Attitudes Toward Prenatal Diagnosis

In thinking about the meaning of using prenatal diagnosis to detect disabling traits, it is important to notice that the class of "disabling traits" is exceedingly heterogeneous. Prenatal diagnosis can now detect conditions as different as Lesch-Nyhan syndrome and ectrodactyly (a trait involving a partial fusion of the bones of the fingers and toes). Further, not only are the traits heterogeneous, but so are perceptions of their significance and/or seriousness. Nancy Press's research reveals that some generalizations can be made about what people take to be "serious": for example, mothers considering prenatal testing are most fearful of conditions like Lesch-Nyhan, which results in early and painful death (39). But as the infamous Bree Walker Lampley case indicates, there is debate about the seriousness of ectrodactyly. In 1991 Bree Walker Lampley, a television news woman in Los Angeles who had ectrodactyly, discovered that the fetus she was carrying had the trait and, when asked, made it known that she had no interest in terminating for such a trait; some suggested that it was "irresponsible" to bring a child into the world with such a serious trait (40). Indeed, the research of Dorothy Wertz and colleagues suggests that even genetics professionals have very different ideas of what is and what is not "serious" (41). In one of Wertz's surveys, cleft lip/palate, neurofibromatosis, hereditary deafness, insulin-dependent diabetes, Huntington's disease, cystic fibrosis, sickle cell anemia, Down syndrome, and manic depression were deemed serious by some professionals and not serious by others (42).

A similar plurality of views exists within the disability community. Many groups representing people with disabilities, such as the National Down Syndrome Congress and Little People of America, have position statements affirming the value of life with disability for individuals and families (43,44). However, there is nuance and disagreement among groups, and in fact within some groups. This complexity is suggested by attitudes within the membership of Little People of America. Many of those who live with achondroplasia are concerned that prenatal testing, which can identify heterozygotes (i.e., fetuses that will develop into long-lived people with achondroplasia) will be used to obliterate the Little People of America community. In fact some members of that community might use the technology to select for the trait. Nevertheless, many couples who are heterozygous for achondroplasia would like to use prenatal testing to identify fetuses that are homozygous for the allele associated with achondroplasia. Homozygous achondroplasia is a uniformly fatal condition, and they would like to spare themselves the experience of bearing a child who will soon die. Adding to the complexity, some people with disabilities would use prenatal testing to selectively abort a fetus with the trait they themselves carry — and some people who would not abort a fetus carrying their own disability might abort a fetus if it carried a trait incompatible with their own understanding of a life they want for themselves and their child.

A similar diversity of views toward prenatal testing and abortion can be found among parents raising a child with a disability. Many such parents do not use prenatal diagnosis to determine whether their present fetus is affected (45). The reasons for this are many; to some, the trait has come to be unimportant or irrelevant. Some may refuse it on the ground that using the technology would say something hurtful to or about their existing child. Other parents of children with disabilities decide to use these technologies.

The point about the plurality of traits and attitudes toward testing is not to suggest that the terrain is too complex to be amenable to policy response. The point is simply that people committed to ending discrimination and improving life for people who have disabilities are not monolithic on the prenatal testing issue, any more than all feminists are monolithic on a host of "women's issues" or than members of racial minorities are monolithic in

their stance toward affirmative action or other practices that affect them. Comprehensive evaluations of prenatal genetic testing will have to take such complexities into account.

Reproductive Liberty Premise

The proliferation of prenatal genetic testing occurs against the background of the controversy about abortion. Prenatal testing for genetic disability elicits unexpected responses from both sides of the abortion debate: Many of those who are uneasy with abortion based on a prenatal finding of a disabling trait are pro-choice. And many who, in general, are against the right to abortion nonetheless approve of abortions performed on a fetus carrying a disabling trait.

Virtually all the major work in the disability critique of prenatal testing emerges from those who are also committed to a prochoice, feminist agenda: Adrienne Asch, Marsha Saxton, Anne Finger, and Deborah Kaplan, for example (46–49). Other pro-choice feminists, including Ruth Hubbard, Abby Lippman, Carole Browner, and Nancy Press, draw on the disability critique to question the impact of prenatal testing (50–53). The shared premise of these scholars is that women (and men) have the right to determine when and how many children they will have; within the first two trimesters of pregnancy, abortion is a legally and morally defensible means of exercising that right.

What is new about prenatal testing is that it enables prospective parents to some extent to determine not only when and how many but also what kind of children they will have. With the exception of revealing the sex of the fetus, current prenatal testing is used to detect traits considered medically disabling — characteristics deemed undesirable or departures from species-typical functioning. In the future it may be increasingly possible to select for traits that we do value. That, however, is not the possibility that has motivated the disability critique; the motivation for the disability critique is the reality of using prenatal testing and selective abortion to avoid bringing to term fetuses that carry disabling traits.

ETHICAL ARGUMENTS OF THE DISABILITY RIGHTS CRITIQUE

As mentioned above, the number and variety of conditions for which prenatal genetic tests are available grows almost daily (54). Today we test for one trait at a time. In the future, however, with advances in diagnostic technology, it will be possible to test simultaneously for as many traits as one would like. In principle, it will be able to test for any trait one wishes that has been associated with any given allele. Not only will the cost of such testing likely decrease as the diagnostic technology advances, but advances in the technology will make it possible to do the testing earlier in the pregnancy. As mentioned earlier, one such technology will isolate the very small number of fetal cells that circulate in the maternal blood. Insofar as these earlier tests will be performed on fetal cells obtained from the mother's blood (rather than from the amniotic sac or chorionic villi) they will be minimally invasive. Thus it will be possible to do many more tests, at once, and with less cost to the pregnant woman in time, inconvenience, risk, or dollars, than is now the case (55).

As the ease of testing increases, so does the perception within both the medical and broader communities that prenatal testing is a logical extension of good prenatal care: The idea is that prenatal testing helps prospective parents have healthy babies. On the one hand, this perception is quite reasonable. Although no researcher has yet even attempted to correct a genetic impairment with in utero gene therapy, increasingly there are nongenetic approaches to such impairments. At the time of this writing, more than 50 fetuses have undergone in utero surgery to repair neural tube impairments (myleomeningoceles) (56). Moreover, negative (or reassuring) prenatal test results will reduce the anxiety felt by many prospective parents, and this in itself can be construed as part of good prenatal care. On the other hand, as long as in utero interventions remain relatively rare, and as long as the number of people seeking prenatal genetic information to prepare for the birth of a child with a disability remains small, prospective parents will use positive prenatal test results primarily as the basis of a decision to abort fetuses that carry mutations associated with disease and/or disability. Thus there is a sense in which prenatal testing is not simply a logical extension of the idea of good prenatal care.

Logical extension or no, using prenatal tests to prevent the birth of babies with disabilities seems to be self-evidently good to many people. Even if the testing will not help bring a healthy baby to term this time, it gives prospective parents a chance to try again to conceive. To others, however, prenatal testing looks rather different. A moment's reflection about the history of our society's treatment of people with disabilities makes it easy to appreciate why people identified with the disability rights movement might regard such testing as dangerous. Critics contend that prenatal diagnosis reinforces the medical model that disability itself, not societal discrimination against people with disabilities, is the problem to be solved. The charge that such testing is dangerous is supported by two, broad lines of argument. The first is that prenatal testing followed by selective abortion is *morally problematic*. The second line of argument is that the desire to undertake prenatal testing is based on *misinformation* about what disability is like for people with disabilities and for their families.

Prenatal Testing Is Morally Problematic

The disability critique holds that selective abortion after prenatal diagnosis is morally problematic, and for two reasons. First, selective abortion expresses negative or discriminatory attitudes not merely about a disabling trait, but about those who carry it. Second, it signals an intolerance of diversity not merely in the society but in the family, and ultimately it could harm parental attitudes toward children.

The Expressivist Argument. The argument that selective abortion expresses discriminatory attitudes has been called the *expressivist* argument (57). Its central claim

is that prenatal tests to select against disabling traits express a hurtful attitude about and send a hurtful message to people who live with those same traits. In the late 1980s Adrienne Asch put the concern this way: "Do not disparage the lives of existing and future disabled people by trying to screen for and prevent the birth of babies with their characteristics" (58, p. 81). More recently, she has clarified what the hurtful or disparaging message is:

> As with discrimination more generally, with prenatal diagnosis, a single trait stands in for the whole, the trait obliterates the whole. With both discrimination and prenatal diagnosis, nobody finds out about the rest. The tests send the message that there's no need to find out about the rest (59).

Indeed, many people with disabilities, who daily experience being seen past because of some single trait they bear, worry that prenatal testing repeats and reinforces that same tendency toward letting the part stand in for the whole. Prenatal testing seems to be more of the discriminatory same: a single trait stands in for the whole (potential) person. Knowledge of the single trait is enough to warrant the abortion of an otherwise wanted fetus. On Asch's more recent formulation, the test sends the hurtful message that people are reducible to a single, perceived-to-be-undesirable trait.

This observation about letting the part stand in for the whole is surely enormously important. In everyday life, traits do often stand in for the whole, people do get looked past because of them. Indeed, one form of the expressivist argument has been regarded rather highly in another context. Many people who are concerned to support women's rights, have argued that prenatal sex selection is morally problematic because it embodies and reinforces discriminatory attitudes toward women (60). The sex trait is allowed to obliterate the whole, as if the parents were saying, "We don't want to find out about 'the rest' of this fetus; we don't want a girl."

Marsha Saxton has put the expressivist argument this way:

> The message at the heart of widespread selective abortion on the basis of prenatal diagnosis is the greatest insult: some of us are "too flawed" in our very DNA to exists; we are unworthy of being born. . . .[F]ighting for this issue, our right and worthiness to be born, is the fundamental challenge to disability oppression; it underpins our most basic claim to justice and equality—we are indeed worthy of being born, worth the help and expense, and we know it (61).

And as Nancy Press has argued, by developing and offering tests to detect some characteristics and not others, the professional community is expressing the view that some characteristics, but not all, warrant the attention of prospective parents (62).

For several reasons, however, there is disagreement about the merit of the expressivist argument as a basis for any public policy regarding prenatal diagnosis of disability. Individual women and families have a host of motives and reasons for seeking out genetic information, and as James Lindemann Nelson and Eva Feder Kittay argue, it is impossible to conclude just what "message" is being sent by any one decision to obtain prenatal testing (63,64). Acts (and the messages they convey) rarely have either a single motivation or meaning.

Some prospective parents no doubt have wholly negative attitudes toward what they imagine a life with a disability would be like for them and their child; others may believe that life could be rich for the child, but suspect that their own lives would be compromised. Others who have disabilities perhaps see passing on their disabling trait as passing on a part of life that for them has been negative. Parents of one child with a disability may believe that they don't have the emotional or financial resources for another. The point is that the meaning of prenatal testing for would-be parents is not clear or singular. In any case, those sympathetic to at least some forms of prenatal testing point out that prospective parents do not decide about testing to hurt existing disabled people but to implement their own familial goals. In that sense, there is no "message" being sent at all.

To many in the disability rights movement, however, regardless of the parental motive to avoid the birth of a child who will have a disability, the parent may still be letting a part stand in for the whole. That prospective parents do not intend to send a hurtful message does not speak to the fact that many people with disabilities receive such a message and are pained by it.

A second criticism of the expressivist argument is that it calls into question the morality of virtually all abortions. The argument presumes that we can distinguish between aborting "any" fetus and a "particular" fetus that has a disability—what Adrienne Asch has called the *any-particular distinction*. According to Asch, most abortions reflect a decision not to bring any fetus to term at this time; selective abortions involve a decision not to bring this particular fetus to term because of its traits. Pro-choice individuals within and outside the disability community agree that it is morally defensible for a woman to decide, for example, that she doesn't want any child at a given time because she thinks she's too young to mother well, or because it would thwart her life plan, or because she has all the children she wants to raise. The question is whether that decision is morally different from a decision to abort an otherwise-wanted fetus.

But it is not clear that the distinction is adequate. Sometimes the decision to abort "any" fetus can be recast as a decision to abort a "particular" fetus. James Lindemann Nelson, for example, argues that if parents of three children chose to end a pregnancy that would have produced a fourth child, such parents would not be making a statement about the worthwhileness of other families with four children, or about the worth of fourth-born children as human beings (64). Rather, they would be deciding what would be right for their particular situation. If, as Asch and others have argued, prenatal testing is morally suspect because it lets a trait stand in for the whole potential person, precisely the same argument would apply to aborting a fetus because it was the fourth child. The trait of being fourth-born makes the prospective parents ignore every other respect in which that fetus could become a child that would be a blessing to its family and community. Nelson's example of the potential fourth-born child suggests one reason to doubt the merit of the

any-particular distinction; he thinks that the disability critics have failed to explain why traits like being fourth-born could be a legitimate basis for an abortion while disabling traits could not.

A third criticism of the expressivist argument is that it presumes that selective abortion based on prenatal testing is morally problematic in a way that other means of preventing disability are not. Such other means include, for example, taking folic acid to reduce the likelihood of spina bifida, or eschewing medication that is known to stunt the growth or harm the organs or limbs of a developing fetus. Such acts (or refraining from such acts) on the part of the pregnant woman are designed to protect the health of the developing fetus.

Disability critics hold, however, that abortion does not protect the developing fetus from anything. It prevents disability by simply killing the fetus. Proponents of this disability critique hold a strong prochoice position. Their objection is only to a certain way of using abortion.

But those from the mainstream prochoice community think of selective abortion in different terms. They do not see an important moral difference between selective abortion and other modes of preventing disability in large part because they do see an important moral distinction between a born child with a disabling trait and an embryo or fetus with a disabling trait. They argue that parents of all born children have an obligation to love and care for those children — regardless of their traits. They also argue, however, that the pregnant woman (and her partner) are not "parents" before the child is born. Just as a woman or couple may decide during the first two trimesters of any pregnancy that becoming a parent to a first child, or to any child, is not in accord with their life plans, so may they make the same decision on the grounds that the fetus has disabling traits. The woman may terminate the pregnancy and try again to become pregnant with a fetus that has not been identified as carrying a disabling trait. On this view, if it is reasonable to prevent disability in a developing child by adhering to a particular lifestyle, taking specified medications or refraining from taking others, it is equally acceptable to opt for abortion to prevent the birth of a child with a significant disability (65).

Even if expressivist arguments will not dissuade all people from using tests in making reproductive decisions for their own lives, there is widespread agreement that policies that would in any way penalize those who continue pregnancies despite knowing that their child will live with a disabling trait must be avoided. That is, there is widespread agreement that prospective parents who either forgo prenatal testing or decide that they want to continue a pregnancy despite the detection of a disabling trait should not have to contend with losing medical services or benefits for their child, nor feel obliged to justify their decisions.

The Parental Attitude Argument. The second argument that prenatal testing is morally problematic may be called the *parental attitude* argument. According to it, using prenatal tests to select against some traits indicates a problematic conception of and attitude toward parenthood. Part of the argument is that prenatal testing is rooted in a "fantasy and fallacy" that "parents can guarantee or create perfection" for their children (58). If parents were to understand what they really should seek in parenting, then they would see how relatively unimportant are the particular traits of their children.

The parental attitude argument also involves the thought that in the context of prenatal testing, a part, a disability, stands in for the whole, a person. The prospective parent who wants to avoid raising a child with a diagnosable disability forgets that along with the disabling trait come other traits, many of which are likely to be as enjoyable, pride-giving, positive (and as problematic, annoying, and complicated), as any other child's traits. If prospective parents imagine that disability precludes everything else that could be wonderful about the child, they are likely acting on misinformation and stereotype.

According to the parental attitude argument, prospective parents should keep in mind that the disabling trait is only one of a fetus' characteristics. The activity of appreciating and nurturing the particular child one has is what the critics of selection view as the essence of good parenting. Loving and nurturing a child entails appreciating, enjoying, and developing as best one can the characteristics of the child one has, not turning the child into someone she is not or lamenting what she is not. If we were to notice that it is a fantasy and fallacy to think that parents can guarantee or create perfection for their child, if we were to recognize what is really important about the experience of parenting, we would see that we should be concerned with certain attitudes toward parenting, not with "disabling" traits in our children. Good parents will care about raising whatever child they receive and about the relationship they will develop, not about the traits the child bears. In short, what bothers those wary of prenatal diagnosis is what might be called "the selective mentality." The attention to particular traits indicates a morally troubling conception of parenthood, a preoccupation with what is trivial and an ignorance of what is profound.

Those who connect acceptance of disability to what is desirable in any parent–child relationship worry that our attitudes toward parenthood and ultimately toward each other are changing as a result of technologies like prenatal diagnosis (66,67). Do these technologies lead us, one might ask, toward the commodification of children, toward thinking about them and treating them as products rather than as "gifts" or "ends in themselves"? Is it making us as a society less resilient in the face of the inevitable risks that our children face, and less willing to acknowledge the essential fragility of our species? When members of our society are confronted with, for example, sex selection or with the possibility of selecting for non-health-related traits like sexual orientation, they often raise concerns about the selective mentality. Indeed, those who want to reject the parental attitude argument in the context of disabling traits should recognize that they are criticizing an argument that they themselves may sometimes use in the context of non-health-related traits. Certainly many worry about the cumulative effect of individual choices, about the technologization of reproduction, and about a decreasing cultural ability or willingness to accept the

reality of uncontrollable events. These concerns trouble even those who profess to be comfortable with genetic testing and selective abortion.

Nonetheless, many find significant problems with the parental attitude argument. One of the most important is that it makes what William Ruddick has called the "maternalist assumption," namely that "a woman who wants a child should want any child she gets" (68). Ruddick acknowledges that many women do hold "maternalist" conceptions of pregnancy and motherhood, out of which that assumption grows. But he argues that there are other legitimate conceptions of pregnancy and motherhood that do not depend on or give rise to the same assumption. He suggests that some prospective parents may legitimately adopt a "projectivist" or "familial" conception of parenthood, and that either of these views is compatible with trying to ensure that any child they raise has characteristics that accord with these parental goals. In the projectivist parent's understanding of child rearing, the child is a part of her parental projects, and within limits, parents may legitimately undertake to ensure that a child starts out with the requisites for fulfilling these parental hopes and aims. Ruddick is not claiming that projectivist parents could ignore a child's manifested commitments to things beyond the parents' life plans, but he is saying, for example, that, the parent passionate about music may legitimately select against a future child whose deafness would make a love of some forms of music impossible. If a hearing child turns out to be tone deaf and enthusiastic about rock collecting and bird watching but not music, and if the parent views these activities as inimical to her parental values or projects, she need not support them, or (within limits) allow other people to do so.

According to Ruddick, the "familial" conception of parenthood highlights a parent's vision of her child as herself a parent, sibling — a participant in a nuclear and extended family that gives central meaning to life. For example, parents whose dreams of child rearing include envisioning their own child as a parent would be acting consistently with their conception of parenthood if they decided not to raise a boy with cystic fibrosis, whose sterility and shortened life span might preclude either biological or adoptive parenthood. A child of such a parent might, of course, reject family life in favor of solitude or communal adult companionship, but in using available technology to avoid raising a child who would never be able to fulfill a deeply cherished parental dream, the parent is acting in accordance with a legitimate conception of parenthood.

Though many share the disability community's concern that prenatal testing may threaten our attitudes toward children, parenthood, and ultimately ourselves, arguments such as Ruddick's and the others mentioned above make it unlikely that such concerns can undergird specific policies regarding prenatal testing for disabling traits.

Prenatal Testing Is Based on Misinformation

The second major claim of the disability critique is that prenatal testing depends on a misunderstanding of what life with disability is like for children with disabilities and their families. Connected with this claim is the question whether disability is one more form of "neutral" human variation, or whether it is different from variations usually thought of as nondisabling traits, such as eye color, skin color, or musicality.

There are many widely accepted beliefs about what life with disability is like for children and their families. Most of these beliefs are not based on data. They include assumptions that people with disabilities lead lives of relentless agony and frustration and that most marriages break up under the strain of having a child with a disability. Recent studies suggest, for example, that many members of the health professions view childhood disability as predominantly negative for children and their families, in contrast to what research on the life satisfaction of people with disabilities and their families has actually shown (69–70). For example, disability researchers Philip Ferguson, Alan Gartner, and Dorothy Lipsky have reviewed empirical data on the impact of children with disabilities on families (71), and have concluded that the adaptational profiles of families that have a child with a disability basically resemble those of families that do not.

According to Ferguson, Gartner, and Lipsky's reading of the data, families that include disabled children fare on average no better or worse than families in general. Some families founder, others flourish. Ferguson, Gartner, and Lipsky do not deny that families are often distressed upon first learning that their child has a disability. And they acknowledge that families with children who evince significantly challenging behavior experience more disruption than do other families. But recent research on raising a child with a disability offers happier news for families than many in our society have been led to expect. The Ferguson, Gartner, and Lipsky review of scores of studies about family life where children have significant cognitive, physical, and sensory disabilities, behavioral and health problems, suggests that on average, families with and without disabled children fare about the same on such measures as parental stress, marital satisfaction, and family functioning.

Although families of children with a variety of conditions have been studied, families of children with Down syndrome have received the most extensive examination. Nonetheless, all of those studies concerned children whose conditions are incontrovertibly disabilities of some consequence. The findings indicate that challenging behavior of a child is much more likely to disrupt families and cause negative consequences than significant intellectual disability or health problems. While studies differ in methodology, population studied, questions pursued, and types of conclusions, according to Ferguson, Gartner, and Lipsky's interpretation, what most reviewers find is that the mild-severe continuum is not the important one in terms of family outcomes. Behavior of the child is a much stronger predictor of negative consequences than is intellectual or physical impairment.

Studies of family adaptation, too, have begun to recognize the prevalence of positive outcomes in many families (72,73). Indeed, one recent study found that parents

of disabled adolescents reported more positive perceptions of their children than do parents of nondisabled adolescents (74).

In a 1995 study intended to learn how a child's disability affected the work lives of dual career families, the authors found that the needs and concerns of families with and without children with disabilities were "strikingly similar." They did, however, observe:

> What seems to distinguish families of children with disabilities from other working families is the intensity and complexity of the arrangements required to balance work and home responsibilities successfully. For example, parents of children with disabilities, particularly those with serious medical or behavioral problems, find it more difficult to locate appropriate, affordable child care. . . .Similarly, these families are more dependent upon health insurance policies with comprehensive coverage (75, p. 511).

This same study also suggests, however, that a child's disability may sometimes alter the customary parent–child life cycle, in which parents gradually relinquish daily guidance and caretaking and — if they are fortunate — see their children take on adult productive and caretaking roles. Depending on the impairment and on the social arrangements that parents help a growing child construct, some people with disabilities may require their parents' help through adulthood in securing shelter, social support, and safety. Increasingly, adults with disabilities such as muscular dystrophy, spina bifida, cystic fibrosis, Down syndrome, and other conditions do not stay "eternal children," as they were once thought to do. Nonetheless, some, albeit small, portion of the population of disabled people will be more vulnerable for longer than others, and more in need of what Kittay (borrowing from Sara Ruddick) describes as "attentive, protective love" (76).

While it is important to demolish the myth that disability entails relentless agony for the child and family, there is still considerable disagreement about what conclusions to draw from the literature on the family impact of a child with disability. In the view of the disability community, this literature suggests that prenatal testing to select against disabling traits is misguided in the sense that it is based on misinformation. That is, if prospective parents could see that families with children who have disabilities fare much better than the myth would have it, then parents would be less enthusiastic about the technology.

However, recognizing that there are erroneous beliefs that need to be dispelled may not show that the desire for prenatal testing stems from misinformation alone. The first problem with the argument from misinformation has to do with the difference between retrospective and prospective judgments. It is one thing to look back on a stressful but ultimately rewarding experience and say, I'm glad I did that. It is another to look forward to the possibility of a stressful and perhaps ultimately rewarding experience and say, I'm glad to give it a try. To appreciate that many families respond well to stress does not commit one to thinking that it would be a mistake for families to try to avoid it. It may be true that, as one of the studies of working families points out, the concerns of working parents with disabled children very much resemble the concerns of any working parent — ensuring that children are safe, happy, stimulated, and well cared for at home, at school, and in after-school activities. But that study also acknowledges that working parents of children with special medical or behavioral needs find that meeting those needs takes more time, ingenuity, and energy than they think would have to be spent on the needs of nondisabled children. To appreciate that many families emerge stronger, wiser, and even better as a result of such an experience may not suggest that it is unreasonable or morally problematic to try to avert it.

Disability in Society. One of the most difficult issues that emerges in the argument from misinformation concerns what having a disability is "really" like for people themselves and for their families. Just how much of the problem of disability is socially constructed? Is it reasonable to say that in a differently constructed social environment, what are now disabling traits would become "neutral" characteristics?

Undoubtedly, more of the problem of disability is socially constructed than many people generally believe. But does that imply that having a characteristic like cystic fibrosis or spina bifida is of no more consequence than being left-handed or being a man who is five feet, three inches tall? According to the disability rights critique of prenatal testing, if people with disabilities were fully integrated into society, then there would be no need for the testing. In the world they seek to create, if a given health status turned out to be a handicap, that would be because of societal, not personal, characteristics; the appropriate response would be to change society so that the person could live a full life with a range of talents, capacities, and difficulties that exist for everyone. In a society that welcomed the disabled as well as the nondisabled, there would be no reason to prevent the births of people with traits now called disabling.

Those sympathetic to at least some forms of prenatal testing are struck by the fact that, for reasons that seem to be complex, members of the disability community speak at different times in different modes about the nature of disability. Sometimes, members of that community are clear about the fact that disabling traits have a "biological reality" or are not neutral. Adrienne Asch writes, "The inability to move without mechanical aid, to see, to hear, or to learn is not inherently neutral. Disability itself limits some options" (58, p. 73). At other times, however, and this is the mode usually emphasized in critiques of prenatal testing, those in the disability rights movement speak as if those traits indeed are inherently neutral. Thus Deborah Kent writes: "I premised my life on the conviction that blindness was a neutral characteristic (77). In this other mode, the disability community argument often is that, different from what prospective parents imagine, these so-called disabling traits are not, to coin a term, "disvaluable" in themselves; they are disvaluable because of the way they are socially constructed.

Nora Groce's work illustrates the point about how social arrangements shape whether a characteristic is disabling (78). In Martha's Vineyard in the nineteenth

century, Groce argues, being unable to hear was not disabling because everyone spoke sign language. Groce's work establishes that much of what is difficult about having a disability stems from manifold facets of society, from architecture to education to aesthetic preferences. In choosing how to construct our societies, we do, as Allen Buchanan puts it, "choose who will be disabled" (79). We could choose differently than we have, and if we were to choose differently, what's disabling about what we now call disabilities would be largely eliminated. Plainly, then, the social constructionist argument is powerful. The objection concerns, rather, what appears to be a correlative claim of the disability position: that so-called disabling traits are neither disabling nor "disvaluable," but neutral.

Again, adherents of the disability critique acknowledge that some characteristics now labeled disabilities are easier to incorporate into today's society, or into a reconstructed society, than are others. Thus no one would deny that disabling traits — departures from species-typical functioning — foreclose some options, or that some disabilities foreclose more options than others. A child with Down syndrome may never climb Mount Rainier because his strength, agility, and stamina may preclude it; he may also never read philosophy because he does not have the skills to decipher abstract material. Granting that people who can climb mountains and read abstract papers derive enjoyment and meaning from such activities, then being foreclosed from them, not by one's own choice, is regrettable. The lack of possibility is widely seen as disvaluable. In addition these lacks of capacity stem from the characteristics of the individual who is not strong enough or agile enough to climb, or who is unable by any teaching now known to us to grasp complex abstract discourse. In that sense, disability community critics acknowledge that these facets of some disabilities are "real," inherent in the characteristic itself and not an artifact of any interaction with the environment. Even if all traits are to some extent "socially constructed," that is irrelevant to the fact that the existence of these traits forecloses for those who have them the opportunity to engage in some highly desirable and valuable activities; not being able to engage in those activities is disvaluable. To the extent that spina bifida, Down syndrome, blindness, or cystic fibrosis currently preclude people from undertaking some parts of life that people who do not have those traits might experience, the disability critique acknowledges that disability puts some limits on the "open future" (80,81) people seek for themselves and their children.

As Bonnie Steinbock argues, if we really thought disability "neutral," we would not work as we do to maintain, restore, and promote health in ourselves and others. We use medicine in the hope that it will cure or ameliorate illness and disability. We urge pregnant women to refrain from activities that risk harming the fetus. If we thought that disabilities were "neutral," then we could tell women who smoke or drink during pregnancy to rest easy, for developmental delay, low birth weight, and fetal alcohol syndrome would all be just "neutral variations," of no consequence to the future child (82).

While disability community critics acknowledge that some disabilities foreclose some opportunities, they also hold that calling attention to the foreclosure obscures two important points. The first is that rather than dwell on the extent to which opportunities to engage in some activities are truncated, we should concentrate on finding ways for people with disabilities to enjoy alternative modes of those same activities. Philip Ferguson puts it this way:

> The point is not so much whether ... a blind person cannot enjoy a Rembrandt ... but whether social arrangements can be imagined that allow blind people to have intense aesthetic experiences. ...People in wheelchairs may not be able to climb mountains, but how hard is it to create a society where the barriers are removed to their experiences of physical exhilaration? ...Someone with Down syndrome may not be able to experience the exquisite joy of reading bioethics papers and debating ethical theory, but ... that person can experience the joy of thinking hard about something and reflecting on what he or she really believes. ...The challenge is to create the society that will allow as many different paths as possible to the qualities of life that make us all part of the human community (83).

The second fundamental point is that rather than concentrate on the truncation or loss of some opportunities, our society generally — and prospective parents in particular — should concentrate on the nearly infinite range of remaining opportunities. Every life course necessarily closes off some opportunities in the pursuit of others. Thus, while the disability critics of prenatal diagnosis acknowledge that disability is likely to entail some amount of physical, psychological, social, and economic hardship, they hold that when viewed alongside any other life, on balance, life is no worse for people who have disabilities than it is for people who do not. No parent should assume that disability assures a worse life for a child, one with more suffering and less quality, than will be had by those children with whom she or he will grow up.

The claim then is that overall, there is no more stress in raising a child with a disability than in raising any other child, even if at some times there is more stress, or different stress. In that sense the disability community claims that disability is on balance neutral. Even here, however, many find that the terms "neutral" and "normal" are either inaccurate characterizations of disability or are being used in confusing ways. Specifically some worry that these terms are used sometimes only to describe or evaluate traits and at other times to describe or evaluate persons.

Evaluations of Traits Versus Evaluations of Persons. As already mentioned, the disability community itself sometimes speaks about the descriptive and evaluative senses in which disabling traits are not neutral, not normal. Legislation like the ADA could not exist without a recognition that in some sense disabling traits are neither neutral nor normal. Indeed, the societal provision of special resources and services to people with disabilities depends on noticing the descriptive and evaluative senses in which disabling traits are not neutral, and how the needs of the people who live with them are, descriptively speaking, not normal. Yet the recognition of the obligation to provide those special resources is rooted in a commitment to the fundamental idea that the people living with those traits are,

morally speaking, "normal"; the people bearing the traits are evaluatively normal in the sense of deserving the normal respect due equally to all persons. Unequal or special funding expresses a commitment to moral equality. Recognizing the nonneutrality of the trait and the "abnormality" of the person's needs is necessary for expressing the commitment to moral equality and equal opportunity. There is nothing paradoxical about appreciating the descriptive sense in which people with disabling traits are abnormal while also appreciating the evaluative or moral sense in which they are normal.

Some who are sympathetic to prenatal testing worry that people in the disability community (as well as others) often conflate descriptive claims about traits and evaluative or moral claims about persons. For example, Deborah Kent, who is blind, writes:

> When I was growing up people called my parents "wonderful." They were praised for raising me "like a normal child." As far as I could tell, they were like most of the other parents in my neighborhood, sometimes wonderful and sometimes very annoying. And from my point of view I wasn't like a normal child — I was normal (77).

What does Kent mean when she says that she "was normal"? As a descriptive claim, it is not reasonable to say that the trait of blindness is normal. Statistically speaking, it is not. Also, as an evaluative claim, insofar as the trait can make it impossible to enjoy some wonderful opportunities, it does not seem reasonable to say that the trait is neutral. The trait may indeed seem neutral and insignificant when viewed in the context of the whole person, but that is a claim about the person, not the trait. On the view of those sympathetic to testing, the descriptive and evaluative claims about the trait do not bear a necessary logical relation to evaluative claims about the person who bears it. As an evaluative or moral claim about the person, it makes perfect sense to say that a person who is blind is normal; she is normal in the sense that she deserves the normal, usual, equal respect that all human beings deserve.

But if it is easy to notice the difference between the descriptive and evaluative claims about traits and the evaluative claims about persons, why do people in the disability community (and others) keep slipping between the two? Erik Parens has suggested that there may be an important reason for this seemingly imprecise slipping (84). Discrimination against people with disabilities often involves a tendency to allow the part to stand in for the whole; perhaps members of the disability community sometimes succumb to a similar, equally problematic error. It could be that as the majority community sometimes uses the trait to deny the moral significance of the person, the disability community sometimes uses the moral significance of the person to deny the significance of the trait. The majority community slips from an observation about a trait to a claim about a person; the disability community slips from an observation about a person to a claim about a trait. At important moments, both groups fail to distinguish evaluations of traits from evaluations of persons. While such slippage may be easily committed in both communities, and particularly understandable on the part of the disability community, it may be equally counterproductive in both.

Regardless of whether one is or is not persuaded by the disability community arguments regarding prenatal testing, it is important to remember that the disability community arguments are not intended to justify wholesale restrictions on prenatal testing for genetic disability. Rather, they are intended to make prospective parents pause and think about what they are doing, and to challenge professionals to help parents better examine their decisions. They are intended to help make the decisions of prospective parents thoughtful and informed, as opposed to thoughtless and automatic. As the prenatal testing technology marches forward, the need for thoughtful private and public conversations about its uses will become increasingly great. The disability rights arguments will be an invaluable resource in the promotion of those conversations.

BIBLIOGRAPHY

1. D.D. Weaver, *Catalog of Prenatally Diagnosed Conditions*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1999.
2. D.M. Serr, L. Sachs, and M. Danon, *Bull. Res. Council Israel* **5B**, 137 (1955).
3. F. Fuchs and P. Riss, *Nature* **177**, 330 (1956).
4. E.L. Makowski, K.A. Prem, and I.H. Kaiser, *Science* **123**, 542 (1956).
5. L.B. Shettles, *Am. J. Obstet. Gynecol.* **71**, 834 (1956).
6. R. Schwartz Cowan, *Fetal Diagnostic Ther.* **8**(suppl. 1), 10–17 (1993).
7. P. Riis and F. Fuchs, *Lancet* **2**, 180 (1960).
8. J. LeJeune, M. Gauthier, and R. Turpin, *C. R. Acad. Sci.* **248**, 1721–1722 (1959).
9. M.W. Steele and W.R. Breg, *Lancet* **1**, 383 (1966).
10. H.L. Nadler, *Pediatrics* **42**, 912 (1968).
11. C.B. Jacobson and R.H. Barter, "Intrauterine diagnosis and management of genetic defects," *Am. J. Obstet. Gynecol.* **99**, 796–807 (1967).
12. American Academy of Pediatrics, American College of Obstetricians and Gynecologists, *Guidelines for Perinatal Care*, Washington, DC, 1983.
13. M.M. Kaback, *N. Engl. J. Med.* **289**, 1090–1091 (1973).
14. J. Mohr, *Acta Pathol. Microbiol. Scand.* **73**, 73–77 (1968).
15. N. Hahnemann, *Clin. Genet.* **6**, 294–306 (1974).
16. Z. Kazy, I.S. Rosovsky, and V.A. Bakharev, *Prenat. Diag.* **2**, 39–45 (1982).
17. Canadian collaborative CVS-Amniocentesis Trial Group, *Lancet* **1**, 1–7 (1989).
18. D.J. Brock and R.G. Sutcliffe, *Lancet* **2**, 197 (1972).
19. D.J. Brock, A.E. Bolton, and J.M. Monaghan, *Lancet* **2**, 923–924 (1973).
20. A. Milunsky, ed., *Genetic Disorders and the Fetus Diagnosis, Prevention, and Treatment*, 4th ed., Johns Hopkins University Press, Baltimore, MD, 1998.
21. R. Merkatz et al., *Am. J. Obstet. Gynecol.* **184**, 896 (1984).
22. A. Kolker and B.M. Burke, *Prenatal Testing, A Sociological Perspective*, Bergin and Garvey, Westport, CT, 1994.
23. F.J. Meaney, S.M. Riggle, and G.C. Cunningham, *Fetal Diagnosis Ther.* **8**(suppl.), 18–27 (1993).

24. B.A. Bernhardt et al., *Obstet. Gynecol.* **91**, 649–655 (1998).
25. American Academy of Pediatrics and American College of Obstetrics and Gynecology, *Guidelines for Perinatal Care*, 1997.
26. A. Gartner and J. Tome, eds., *Images of the Disabled: Disabling Images*, Praeger, New York, 1987.
27. J. Shapiro, *No Pity*, Times Books, New York, 1992.
28. J. West, ed., *The Americans with Disabilities Act: From Policy to Practice*, Milbank Memorial Fund, New York, 1991.
29. L.J. Davis, *Enforcing Normalcy: Disability, Deafness, and the Body*, Verso, London, 1995.
30. P. Ferguson, A. Gartner, and D. Lipsky, in E. Parens and A. Asch, eds., *Prenatal Testing and Disability Rights*, Georgetown University Press, Washington, DC, 2000.
31. J.E. Tyson and R.S. Broyles, *J. Am. Med. Assoc.* **276**, 492–493 (1996).
32. National Organization on Disability, *Survey of Americans with Disabilities*, 1998.
33. A.L. Beaudet, *Am. J. Hum. Genet.* **47**(4), 603–605 (1990).
34. D.T. Helm, S. Miranda, and N.A. Chedd, *Mental Retard.* **36**(1), 55–61 (1998).
35. B.A. Bernhardt et al., *Obstet. Gynecol.* **91**, 648–655 (1998).
36. T.M. Marteau, J. Kidd, and M. Plenicar, *J. Reprod. Infant Psychol.* **11**, 3–10 (1993).
37. C. Dunne and C. Warren, *Issues Law Med.* **14**(2), 165–202 (1998).
38. T. Marteau, H. Drake, and M. Bobrow, *J. Med. Genet.* **31**, 864–867 (1994).
39. N. Press et al., in S. Franklin and H. Ragone, eds., *Reproducing Reproduction: Kinship, Power, and Technological Innovation*, University of Pennsylvania Press, Philadelphia, PA, 1998.
40. A. Kolker and B.M. Burke, *Prenatal Testing: A Sociological Persepective*, Bergin and Garvey, Westport, CT, 1994.
41. B.M. Knoppers et al., *Am. J. Hum. Genet.* **57**(4, suppl.), A296, 1723 (1995).
42. D. Wertz, Abstract, National Society of Genetic Counselors meetings, October 1998.
43. National Down Syndrome Congress, *Position Statement on Prenatal Testing and Eugeneics: Families' Rights and Needs*, Available at: *http://members.carol.net/ndsc/eugenics.html*
44. Little People of America, *Position Statement on Genetic Discoveries in Dwarfism*, Available at: *http://www2.shore.net/~dkennedy/dwarfism-genetics.html*
45. D.C. Wertz, in H. Holmes, ed., *Issues in Reproductive Technologies I*, Garland Publishers, New York, 1992.
46. M. Fine and A. Asch, *Reprod. Rights Newslett.* **4**(3), 19–20 (1982).
47. M. Saxton, in E.H. Barucch, A.E. D'Amado, and J. Seager, eds., *Embryos, Ethics and Women's Rights*, Haworth Press, New York, 1988.
48. A. Finger, *Past Due: Disability, Pregnancy, and Birth*, Seal Press, Seattle, WA, 1987.
49. D. Kaplan, in K.H. Rothenberg and E.J. Thomson, eds., *Women and Prenatal Testing: Facing the Challenges of Genetic Testing*, Ohio State University Press, Columbus, OH, 1994.
50. R. Hubbard, *The Politics of Women's Biology*, Rutgers University Press, New Brunswick, 1990.
51. A. Lippman, *Am. J. Law Med.* **17**(1–2), 15–50 (1991).
52. C. Browner and N. Press, *Med. Anthropol. Q.* **10**(2), 141–156 (1996).
53. M.A. Field, *Harvard Women's Law J.* **16**, 79–138 (1993).
54. C.M. Powell, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
55. T.H. Murray, *The Worth of a Child*, University of California Press, Berkeley, CA, 1996.
56. D.W. Bianchi, T.M. Crombleholme, and M. D'Alton, *Fetology: Diagnosis and Management of the Fetal Patient*, McGraw-Hill, New York, forthcoming.
57. A.E. Buchanan, *Soc. Philos. Policy* **13**, 18–46 (1996).
58. A. Asch, "Reproductive Technology and Disability," p. 81.
59. A. Asch, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
60. D.C. Wertz and J.C. Fletcher, in H.B. Holms and L.M. Purdy, eds., *Feminist Perspectives in Medical Ethics*, Indiana University Press, Bloomington, IN, 1992.
61. M. Saxton, in R. Solinger, ed., *Abortion Wars: A Half Century of Struggle, 1950–2000*, University of California Press, Berkeley, CA, 1997.
62. N. Press, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
63. E.F. Kittay and L. Kittay, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
64. J.L. Nelson, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
65. B. Steinbock, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
66. T.H. Murray, *Worth of a Child*, University of California Press, Berkeley, CA, 1996.
67. A. Asch and G. Geller, in S.M. Wolf, ed., *Feminism and Bioethics: Beyond Reproduction*, Oxford University Press, New York, 1996.
68. W. Ruddick, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
69. J.A. Blier Blaymore et al., *Clin. Pediat.* **35**(3), 113–117 (1996).
70. M.L. Wollraich, G.N. Siperstein, and P. O'Keefe, *Pediatrics* **80**(5), 643–649 (1987).
71. P. Ferguson, A. Gartner, and D. Lipsky, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
72. D.A. Abbott and W.H. Meredith, *Family Rel.* **35**, 371–375 (1986).
73. A.P. Turnbull et al., eds., *Cognitive Coping: Families and Disability*, Paul H. Brookes, Baltimore, MD, 1993.
74. J.P. Lehman and K. Roberto, *Mental Retard.* **34**, 27–38 (1996).
75. R.I. Freedman, L. Litchfield, and M.E. Warfield, *Families in Society: J. Contemp. Hum. Services* (October), 507–514 (1995).
76. S. Ruddick, *Maternal Thinking: Toward a Politics of Peace*, Beacon Press, Boston, MA, 1989.
77. D. Kent, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.

78. N.E. Groce, *Everyone Here Spoke Sign Language: Hereditary Deafness on Martha's Vineyard*, Harvard University Press, Cambridge, MA, 1985.
79. A. Buchanan, *Soc. Philos. Policy* **13**(2), 18–46 (1996).
80. D.S. Davis, *Hastings Center Rep.* **24**(6), 15–21 (1997).
81. R. Green, *J. Law, Med. Ethics* **25**(1), 5–16 (1997).
82. B. Steinbock, in E. Parens and A. Asch, eds., *Prenatal Testing and the Disability Rights*, Georgetown University Press, Washington, DC, 2000.
83. P. Ferguson, Personal communication.
84. E. Parens and A. Asch, *Hastings Center Rep.* **29**(5, special suppl.), S1–S22 (1999).

See other entries DISABILITY AND BIOTECHNOLOGY; GENETIC COUNSELING; see also REPRODUCTION, ETHICS entries; REPRODUCTION, LAW, IS INFERTILITY A DISABILITY?; REPRODUCTION, LAW, REGULATION OF REPRODUCTIVE TECHNOLOGIES; REPRODUCTION, LAW, WRONGFUL BIRTH, AND WRONGFUL LIFE ACTIONS.

REPRODUCTION, ETHICS, SEX SELECTION

BONNIE S. LEROY
University of Minnesota, Institute of Human Genetics
Minneapolis, Minnesota

DIANNE M. BARTELS
University of Minnesota, Center for Bioethics
Minneapolis, Minnesota

OUTLINE

INTRODUCTION

There are basically two ways to approach the issue of sex selection. The first includes an assortment of methods with variable results that can be used to enhance the odds of achieving a conception of one sex over the other. None of these methods ensure that a pregnancy will result in the desired sex, rather they are thought to increase the chances. This approach is often referred to as *sex preselection* as the methods concerned are employed prior to conception. The second approach involves techniques used to determine the sex of an embryo or fetus by directly analyzing the genetic material. Both approaches elicit considerable ethical controversy, but the second is more controversial since a conception has already been achieved. The following discussion will first review the history of sex selection and then examine present and possible future technologies. We will summarize a variety of attitudes regarding sex selection, present a case study, and finally, consider a number of ethical arguments that have been advanced toward this subject.

BACKGROUND

Myriad arguments have been extended in support of and in opposition to sex selection. Prior to discussing the issues surrounding sex selection, it is important to look at the history and understand present and possible future technological interventions. Attempts to influence the conception of a desired sex extend far back in our history. Gledhill, in one paper, and Reubinoff and Schenker, in another publication, present useful portrayals of the early history of this practice (1,2). In early Greece one theory held that males developed on the right side of the uterus while females developed on the left. This led to the belief that the sex of an offspring could be controlled by the position of the woman during intercourse. Another belief was that the right testis produced male sperm while the left produced female sperm. This theory sustained for quite some time and in the eighteenth century led to the procedure of removing the left testicle so that one supposedly would be guaranteed a boy. Even Aristotle also had an opinion on this subject. He argued that the partner who was most active during intercourse would determine the sex of a child.

The Talmud suggests that one can influence the sex of offspring by one partner having an orgasm before the other. If a woman has an orgasm first, the child was said to be a male and if a man was first to have an orgasm, the child would be female. Diet was also, at times, thought to play a role. In the Middle Ages, wine and lion's blood and later in the early twentieth century, bitter and sour foods and a diet rich in red meat consumed by the mother were thought to enhance conception of a male. In 1917 it was reported that the right ovary contained male eggs and the left contained female eggs. Also it was thought that ovulation occurred in an alternate fashion, releasing a male egg from the right ovary one month and a female from the left the next. Therefore it was believed to be possible to predict the sex of a conception by counting the months since the last child.

In the 1950s reports began to emerge that supported a belief that one could influence the conception of a child of the desired sex with the timing of intercourse. This theory maintains that the male determining sperm swim faster but have a shorter survival time than the female determining sperm. Therefore intercourse close to the time of ovulation would result in a male, while intercourse days prior to ovulation would more likely result in a female. Many researchers later followed with reports on the effects that timing of conception had on controlling the determination of sex. Some research also included enhancement techniques such as variations in vaginal penetration and douching with an alkaline solution prior to intercourse. Most of these methods have been alternatively

supported and disputed with many attempts to reproduce the results.

Although today much skepticism surrounds these methods, it is clear from this history that considerable effort has been expended over centuries in the attempt to gain the ability to control the sex of offspring. This history provides some insight into the substantial role that sex of our offspring plays in reproductive issues and the personal identity of parents today.

Current technology involves three distinct approaches: the separation of a sperm sample prior to insemination, evaluation of the genetic material of an early embryo prior to implantation in the uterus, and evaluation of the sex of a fetus in an established pregnancy. Each approach will be discussed in the following text. Before examining the methods, it necessary to know that sex in humans is determined by the genetic constitution of the sperm cell that fertilizes the egg. Females normally have two chromosomes designated as Xs and males normally have one X chromosome and one Y. The combination of a normal X chromosome and a normal Y chromosome in a fertilized egg will result in the development of a male fetus and the presence of two normal X chromosomes will result in the development of a female. Females contribute an X chromosome in each oocyte while males produce sperm with either an X chromosome or a Y. A sample of sperm will normally, on average, be composed of equal numbers of X and Y bearing sperm. Given no abnormalities in the genetic material, the odds of conceiving a male versus a female are equal.

SPERM SEPARATION

Efforts to separate X from Y bearing sperm prior to insemination have revolved around various methods that utilize differences such as size, shape, density, charge, swimming characteristics, and DNA content (1–3). Most of these methods fail to produce consistent results and the research conclusions reported conflict with one another (4–6). Some researchers claim success rates of over 80 percent, while others are unable to replicate these results (7–10). However, some recent research appears to hold promise. Investigators using flow cytometry, a method utilizing the 2.9 percent difference in DNA content between an X bearing sperm and a Y bearing sperm, assert to be able to produce a sperm separation of 80 to 90 percent for X enriched samples and 65 to 70 percent for Y enriched samples (2,11,12). Because this method greatly reduces the number of sperm in the sample, it is necessary to couple it with in vitro fertilization (IVF) technology to achieve pregnancy. This makes the procedure complicated, time-consuming and expensive. As investigators become more proficient with this and other methods in the near future, it will likely be possible to effectively separate X from Y bearing sperm.

PREIMPLANTATION DIAGNOSIS

A second technique is preimplantation diagnosis, which is the evaluation of genetic material in an early embryo prior to implantation in the uterus (2,13). Eggs are harvested from the mother and IVF is performed. One or two cells are then taken from the early developing embryo for genetic analysis. This analysis can include chromosome studies to evaluate the number and structure of all of the chromosomes including those determining sex. In addition direct gene analysis can be performed for a limited number of genetic diseases. This technique is also complicated and expensive. Current use focuses mostly on the determination of the sex of an embryo to prevent a pregnancy with a sex-linked genetic disease, or to detect a known single gene disorder (14).

FETAL SEX DETERMINATION

The final group of techniques that can be used for sex selection involves the determination of fetal sex in an established pregnancy (15). Chorionic villus sampling (CVS) entails harvesting a small amount of early placental tissue containing fetal cells at about 10 weeks into the pregnancy. Chromosome analysis, as well as other biochemical and genetic tests, can be performed on these cells, and the sex of the fetus can be determined. Genetic amniocentesis is a procedure that has been used for quite some time to diagnose a pregnancy affected with a chromosomal or inherited disorder. Amniotic fluid is collected at about 15 to 16 weeks into a pregnancy, and this fluid contains fetal cells that can be analyzed. As in CVS, the sex of a fetus can be learned through chromosomal analysis from fetal cells present in the amniotic fluid.

Since CVS and amniocentesis directly analyze the genetic material that, among other things, determines the sex of the fetus, the accuracy of the testing is extremely high. In addition current prenatal ultrasonographic equipment provides the resolution to ascertain fetal sex with a high degree of precision. Ultrasound uses sound waves to provide a picture of the developing fetus. CVS, amniocentesis, and ultrasound are the most widely used procedures to determine the sex of a fetus, since they currently provide the greatest testing accuracy at the least cost.

ATTITUDES TOWARD SEX SELECTION

A number of studies describe the attitudes of specific groups toward sex selection by examining views of a number of diverse cultures, religions, and professional providers. Most studies separated the use of sex selection for prevention of the birth of a child with a sex-linked genetic disease, from sex selection for the sole purpose of choosing the desired sex. A majority of groups surveyed were supportive of sex preselection and sex determination followed by abortion as options for the prevention of a sex-linked genetic disease in offspring. The following discussion will describe attitudes toward sex selection for nonmedical reasons, that is, for the sole purpose of choosing the desired sex in offspring.

In exploring preferences for a child of one sex over another, reports describe a decided cultural difference between Western views and those of Asia, India, and

some other less industrialized societies. In the United States, studies showed that couples preferred a son as their firstborn and a daughter for their second child. The inclination was clearly toward a balanced sex composition in the family, although the majority of women in the United States do not approve of controlling the sex of their offspring (16). In the early 1960s unmarried college students were asked about sex preference if they were to have only one child. At that time, 91 percent of men and 66 percent of women stated that they would prefer a boy (16). In 1972 and 1987 similar surveys showed that the preference for boys had dropped. In 1972, 55 percent of all students would prefer their only child to be a male, and in 1987 that number was down to 52 percent in favor of males (17). In Great Britain one study asked over two thousand pregnant women if they preferred one sex to the other for their child. The majority, 58 percent, stated no preference while only 18 percent leaned toward a boy and 25 percent toward a girl. When asked if they wanted to know the sex of their baby before birth, 62 percent said no and 20 percent said they were unsure. In a small number of Afro-Caribbean and Asian respondents, there was a slight bias toward wanting boys (18). Another study in Great Britain looked at academic and nonacademic men and women between the ages of 18 and 20 years and found that over 75 percent of all respondents did not support the idea of choosing a baby's sex. No differences were found in the responses based either on sex or between the academic and nonacademic populations (19).

In 1993 the London Gender Clinic opened in London, England. Data gathered on couples who attended the clinic during the first 18 months of service provides some insight into those who would use this technology (20). The ethnic distribution among clients was 57.8 percent East Indian, 32 percent European, 3.6 percent Chinese, and 6.8 percent designated as other. Of all the couples participating in the clinic, 80.6 percent stated that they would have had another baby even if sex preselection were not an option. As expected, Asian and East Indian couples overwhelmingly wanted a boy, while European couples stated a slight preference (62.9 percent) for girls. The main reason given for wanting a girl was the desire of the mother to have a daughter. Couples from the Indian community repeatedly stated the need for a boy to carry on the family name for religious and social reasons. Most interestingly, a major reason for seeking these services involved wanting to avoid having a large family in order to get a son. This was important since 54 percent of the Indian couples in the clinic population already had 3 or more girls in their family and 94 percent had not yet had a son. The authors concluded from their experience that those interested in sex selection are mostly couples with two to three children of the same sex who want one last child of the opposite sex to complete their families. A study based in New York City considered which populations were using sex preselection technology and why. Out of 178 couples studied, 58 were from other countries. All non-American couples in the study wanted a boy, while the American couples wanted boys and girls with equal frequency depending on the sex of the children they already had. They expressed a desire to balance their family. In the non-American couples, reasons for wanting a boy included the custom in their country for a son to support their parents at old age, the belief that it was essential for a male to run a family business, for inheritance purposes, the belief that males are more intelligent, and cultural pressures where males are preferred (21).

Canada assembled the Royal Commission on New Reproductive Technologies to make recommendations to the Canadian federal government by the end of 1993 (22). Before making their recommendations, they held public hearings and conducted random surveys involving more than 40,000 people. Decisions were made in light of research findings, public input, and a set of guiding principles including autonomy, equality, and noncommercialization of reproduction. In addressing the issue of sex selection for nonmedical reasons, they concluded, "The commission viewed sex selection for preference as contrary to its guiding principles, and to generally held Canadian values. Policies are needed to ensure that the values of citizens are respected." In Japan a similar national commission, the Japanese Medical Ethics Advisory Board, set guidelines limiting the availability of sex selection due to the concern that couples would overwhelmingly choose boys if reliable methods of sex preselection were available. However, it was noted that physicians are not legally bound by these guidelines (16). In the Netherlands in 1995, attempts to open the first private clinic to offer sex selection for nonmedical reasons failed due to opposition from doctors and politicians (23). The Dutch health secretary stated that he would take legal measures to ban it because, "the clinic's claims are ethically unjust." Physicians were mostly said to be opposed because, "it crosses the border of what is ethically acceptable in the Netherlands." The Royal Dutch Medical Association said that doctors should not cooperate with the clinic.

In India, son preference is so strong that the use of sex selection technologies followed by abortion is widely accepted. Between the years of 1978 and 1983, about 78,000 female fetuses were aborted following the use of sex selection technologies in India (24). Khanna published a study about the practice of sex selection in Shahargaon, a small village in north India (24). It dramatically demonstrates the effects of wide usage of sex selection procedures on a society. The sex composition of the children in this community for ages 0 to 5 years was found to be 691 females to 1000 males. Khanna, the author of this study, reports, "In this society, the birth of a son is considered an economic and political asset associated with the honor of a family, whereas a daughter is born as an expense and as a moral burden." Lobbying groups brought their concerns surrounding this practice to the Indian government and in response the government has attempted to regulate the use of prenatal diagnostic technologies. The results of these regulations has been an increase in cost for services by the clinics to offset the risks of practice, and an increase in the number of illegal "unregistered" clinics. One of many concerns in India is that a sex-selection industry is rapidly developing because of the great potential for profit.

Attitudes toward sex selection vary among different religions as well. Grazi, Wolowelsky, and Jewelewicz

compared the position of traditional Jewish law with that of Roman Catholicism on the subject of assisted reproduction that included sex selection as one of the issues (25). In the Roman Catholic document *The Instruction on Respect for Human Life and the Dignity of Procreation*, reproductive technologies and IVF specifically are considered to be morally and absolutely illicit practices. This position includes the attempt to conceive a baby of a desired sex for any reason. The Catholic church's position articulates a belief held by many that a fetus has a right to life from the moment of conception. Termination for sex selection would be a most egregious violation of that right.

In Jewish law, gametes and unimplanted embryos have no standing. This would suggest that IVF, if used in treatment of infertility followed by implantation of some of the embryos, the selection for a characteristic such as sex may be allowed. However, the position on IVF for the sole purpose of sex selection is not at all clear. Rabbinic authorities who allow IVF, and presumably other technologies, are doing so in support of the religious obligation to procreate. Some authorities forbid IVF completely, and others preserve the halakhic imperative, which is to maintain natural marital relations. The attempt to conceive a child of the desired sex has been described as, "simply too frivolous a halakhic concern" (25). Neither religion supports sex selection for the sole purpose of choosing the sex of offspring.

Several studies have focused on the perspectives of professional care providers, especially on those who provide clinical genetics services. One study evaluated the attitudes toward sex selection of members of the American Society of Human Genetics, the International Fetal Medicine and Surgery Society, the Society of Perinatal Obstetricians, and selected ethicists and clergy with experience in biomedical issues (26). The majority of respondents in all groups considered sex selection ethically unacceptable. Agreement on this position was stronger regarding the use of sex selection in the second and third trimesters than in the first. The authors of this study indicate that one reason for opposition is a belief that gender is not a disease, and therefore this practice is a form of eugenics and ought not to be a part of health care.

Wertz and Fletcher used hypothetical cases to assess ethical decision making by medical geneticists (27). One case involved a choice about whether to perform prenatal diagnosis for sex selection. The responses were fairly equally split three ways among those who would offer the procedure, those who would refuse the procedure, and those who would refer the patient to another facility that does offer the procedure. Respect for patient autonomy was the reason given by 68 percent of those who stated they would offer the procedure. One interesting finding was that women were twice as likely as men to state that they would perform the procedure in respect for patient autonomy. Burke interviewed genetic counselors to determine attitudes towards fetal sex identification and abortion (28). All but one of the 32 genetic counselors who responded strongly opposed the use of prenatal diagnosis for sex selection. Burke noted that this position imposes stress upon the genetic counselors who also support patient autonomy through the ideal of nondirective counseling and almost universally uphold a patient's right to an abortion. In a similar study of genetic counselor attitudes, Pencarinha, Bell, Edwards, and Best found comparable results (29). When presented with a hypothetical case, 38.3 percent of genetic counselors responding would perform genetic counseling for sex selection while 18.3 percent would refuse to be involved in a such a case and 43.4 percent would refer the couple to another center. Many of the respondents who would refuse the request for these services defended their position, "it is not a medical indication for testing and because prenatal diagnostic services are a limited resource." Genetic counselors who would offer the procedure maintain that the patient has the right to choose and support the patient's decision out of their duty to respect patient autonomy.

Wertz compared the views of geneticists in the United States with those in other European countries (30). As a group, geneticists in the United States were more willing to perform prenatal diagnosis for sex selection, or offer a referral for such services, than geneticists from any other country. Only 4 percent of the participants saw sex selection as having social consequences. Wertz reported that the participants focused on the particular family involved in the case, not on society as a whole. In an extension of this study using the same hypothetical cases, attitudes of genetics service providers from 30 provinces in China were studied (31). The majority, 89 percent, of the participants supported the Chinese laws on termination of pregnancy for genetic abnormalities and for population control and family planning considerations. However, more than half opposed the use of prenatal diagnosis for sex selection.

In summary, most cultural and religious groups surveyed, as well as health care professionals, opposed the use of reproductive technologies for the sole purposes of selecting the desired sex. The opposition was most strong when the result was the abortion of a fetus of the undesired sex.

CASE STUDY

It is sometimes helpful to use a real case to begin to think about how to develop an ethical position involved in a particular issue. The following is a true case that occurred at the University of Minnesota perinatal clinic in the mid-1980s.

AJ was a 35-year-old woman 15 weeks into her fourth pregnancy. She came to clinic with her husband seeking prenatal testing. They were both East Indian and had resided in the United States for about three years. An ultrasound study performed previous to this visit revealed that this was a twin pregnancy. AJ and her husband stated that they have had three previous pregnancies resulting in two normal healthy girls and one son, born in India, who died of a heart defect and many other birth defects. The cause of the anomalies was reported to be unknown and medical records were not available. The couple expressed the concern that if they had another son, he would also be affected because their girls were born healthy. They were also concerned about their age-related risk for having

a child with a chromosome abnormality such as Down syndrome.

The genetic counseling session involved a discussion of age-related risks for chromosome abnormalities, the amniocentesis procedure, and the risks and limitations of the testing. Also discussed was the possibility that their son may have had a chromosome abnormality which, if it were to recur, would be detected with the amniocentesis. He also could have had an undefined genetic birth defect that could recur, but the ability to make a prenatal diagnosis would be limited to what could be seen by a level II ultrasound study.

The couple expressed a strong interest in the amniocentesis and the level II ultrasound study. The results of these studies were normal for both fetuses. When results were called to the couple, they stated that they wanted to know the sex of the fetuses. Both fetuses were female. These results were given to the couple at about 18 weeks gestation. Although they both previously expressed concern if one or both babies were male, they did not appear to be relieved with the test results, and the conversation was short.

About three weeks following the results discussion, AJ called the genetic counselor again. She was calling from a local abortion clinic. She was clearly distressed, and she wanted to know if the results of the chromosome studies could possibly incorrect. We discussed the fact that it was possible but very unlikely. Laboratory errors are rare, and the ultrasound study at the time of the amniocentesis agreed with the chromosome analysis regarding the sex of the fetuses and no abnormalities were seen.

AJ then revealed that she had no real choice in her decision to terminate the pregnancy. Although she very much wanted these babies, she had to terminate this pregnancy. She feared that her husband would leave her and her two daughters if she decided to continue the pregnancy. She had no formal education, no money, and no skills necessary for making a living. She had no way to support herself or her children without her husband. She had struggled with the possibility of having to make this decision since the time she learned of her pregnancy. She then confessed that there had never been a son with birth defects. They had invented this story to explain their interest in the sex of the fetuses, fearing that the clinic would not supply them with this information. Her role in the family was to provide a son and so far she had failed. In reviewing this case we wondered, had she and her husband come to clinic requesting prenatal diagnosis solely for sex selection, what would be the ethical arguments that would help to make a choice for or against providing this service?

ETHICAL ARGUMENTS FOR AND AGAINST SEX SELECTION

Many arguments have been offered in opposition to the use of reproductive technologies for the sole purpose of having a child of a desired sex. A majority point to the injustice of sex discrimination and the value of women. Other arguments in opposition include, but are not limited to, treating children as commodities, inappropriate use of medical technology, setting gender in the same category as disease, the unbalancing of the sex ratio, and lack of respect for human life.

A few organizations have taken a position on this subject. In 1996 the committee on ethics of the American College of Obstetricians and Gynecologists published a position paper specifically on sex selection (32). The committee approved of the use of sex selection for the prevention of sex-linked genetic disorders but strongly rejected the practice of sex selection on demand for the sole purpose of having a child of the desired sex. The main argument given in defense of this position was that they felt this practice, "may reflect and encourage sex discrimination." The committee was concerned that physicians meeting these requests, "may ultimately support sexist practices." The Council on Ethical and Judicial Affairs of the American Medical Association in their position paper on ethical issues related to prenatal genetic testing also opposed the use of sex selection except when it is employed to prevent or treat genetic disease (33). The Council considered selection for sex as, "the most evident example of the discriminatory potential of selection for benign genetic traits." They go on to say that this procedure encourages the value of one sex over the other and places sex in the same category as disease. They argue that the practice of sex selection may result in harms to society including discrimination and the view that children are products. They view sex selection as a form of eugenics. The Turkish General Directorate of Mother and Child Health and Family Planning, analyzed the technical and ethical issues of sex selection (34). This report was explicit in that it emphasized that a baby should never be considered as a technological product. It stated that, "parents should not or any other authorities should not have a right to choose any physical or behavioural features of the baby unless an associated medical problem exists." The Ethical Committee of the Turkish Medical Association agreed that gender should never be treated as a disorder.

In 1985 H.B. Holmes wrote an extensive review of the available technology and ethical arguments and came to the conclusion that sex selection is the practice of eugenics (35). She is careful to acknowledge that for women in countries where females are not valued, the decision to have male children may be a correct moral choice. Given the present social practices, women in these countries are attempting to, "maximize their own and their family's happiness and minimize the suffering of little girls." What is needed, she argues, is social change so that women are valued. She concludes that when people design their children through choosing particular characteristics, they practice eugenics. She states that, "No human is wise enough to choose the kinds of people who ought to perpetuate our species."

Grazi and Wolowelsky examine the issue of sex selection in relation to contemporary Jewish law and ethics (36). They conclude that although new reproductive technology represents an opportunity for alleviating pain and suffering, it should not be used for "frivolous considerations." Rabbinic authorities, not the couple, reserve the right to decide under which situations these technologies can be used. Choosing the desired sex would likely be considered frivolous by most authorities.

Shrivastav writes on this subject from the perspective of the United Arab Emirates (37). It is noted that most citizens of the United Arab Emirates are of the Muslim faith, and since abortions are unacceptable to Muslims, sex selection followed by abortion would not be allowed according to the faith. However, gender preselection may be acceptable to followers of Islam as it would not contravene the Sharia law where IVF and embryo biopsy to rule out disorders are permitted. Despite this, Shirvastav considers that the fact that male offspring are preferred and writes that, "as far as society in the United Arab Emirates is concerned, potential availability of techniques for selecting the gender of offspring will encourage couples to alter the sex ratio of their offspring in favour of boys. Without these techniques, they would probably accept whatever nature has in store for them!"

The issue of unbalancing the normal gender distribution in society by allowing couples to choose the desired sex is a major concern. Seibel, Seibel, and Zilberberstein address this issue and offer a unique solution (38). They first state that they consider using prenatal diagnosis followed by abortion on the basis of sex to be morally unacceptable and that this practice could ultimately lead to an unbalanced sex ratio in society. They go on to say that preimplantation diagnosis for the sole purpose of sex selection, where the unused embryos are destroyed, appears to be an inappropriate use of technology. However, they feel that this practice could be used in an ethically acceptable manner by taking the embryos of the undesired sex and donating them to infertile couples. They argue that if couples were synchronized, this would result in gender distribution. Shenfield, in another letter to the same journal, refutes a previously held position that nature will soon redress the balance of the sex ratio (39). The problem with this position, Shenfield contends, is that it implies the acceptance of the superiority of one sex over the other. Shenfield goes on to assert that the practice of sex selection, "would be detrimental to both sexes to be brought up in a society which acknowledged, by a selective practice, that personal freedom may be obtained at the cost of one's gender identity being constantly assaulted by the implicit disapproval entailed when it becomes a serious handicap worthy of termination."

Steinbacher (40) describes the advantages of the firstborn as being more intelligent, achievement oriented, and successful than second born. Since those who are already privileged will be the ones to utilize sex-selection techniques, supporting selection means enhancing the disparity between men and women globally. "Fewer firstborn females, a higher [male] sex ratio at birth, more poor women in developed countries and elimination of women in the third world are indeed devastating outcomes of sex preselection for women" (40, p. 190). To prevent women's lives from being controlled by technologies, she says that women must first have "voice and vote" when policies are made about all methods of preselection.

Baird's ethical evaluation separates the three approaches: sex-selective abortion, sex-selective implantation, and sex-selective insemination (41). Baird submits that all of these approaches raise ethical concerns about the values of a society of our choice. The concern about sex-selective abortion is lack of respect for human life and dignity. Sex-selective implantation is an invasive and expensive procedure posing risk to women and using medical resources to prevent something that is not a disease. This too, she says, demonstrates a lack of respect for human life. Baird argues that sex-selective insemination does not violate the respect for human life as a life does not yet exist. However, she reasons that this practice reinforces the belief that the sex of the child is important and that families with children of only one sex are less than ideal. Additional concerns about all of these technologies are the possibility that lack of regulation may lead to commercialization of reproduction, the exploitation of the public, and the transformation of children into commodities. Baird argues for regulations and policies to address these concerns.

Botkin examined the broad subject of prenatal screening with respect to policies that would limit parental choice and included the subject of sex selection in his analysis (42). Botkin argues that it would not be justifiable to require that a patient defend her reason for abortion after prenatal screening when abortion is available on demand. However, he does not suggest that parents should have the right to request the use of prenatal screening for any and all purposes, nor that all physicians are obliged to provide all services that patients request. Botkin argues that broad policies limiting parental choice are not workable without a social consensus on the relative values involved, but rather, "physicians should be strongly encouraged to establish and uphold personal moral standards with respect to prenatal screening," thereby respecting the autonomy of both patients and physicians. He believes that, "limits to parental choice may be more appropriately applied through the moral values of individual physicians in their provision of diagnostic services."

Arguments supporting the use of sex selection mostly address parental rights and freedom of choice. Other arguments in support include, but are not limited to, sex selection is merely an extension of other assisted reproductive technologies, it is the least harmful option in some countries, preselection would reduce the number of gender-based abortions and the incidence of infanticide, it would slow population growth, it would eventually result in increased value of women, and it would reduce the number of unwanted children.

In a chapter in *Biomedical Ethics Reviews*, Warren writes in support of the practice of sex selection by refuting the position of Holmes in the same reference (43). She disputes the position that sex selection is a sexist practice. She notes that although some people will only want a child of one sex, many would choose to have a child who is the sex opposite of the child(ren) they already have. Also, in societies where the preference for a son is strong, Warren believes that accusing women of sexism is commensurate with blaming the victim. She reasons that it is not considered wrong to condemn a couple who do not want a child because they are unable to afford to care for it, and so it is also wrong to condemn women in certain societies who decide not to have daughters. She also maintains that many of the arguments made in opposition are based on speculation about the possible

long-term consequences or about how people might behave if this practice was widely accepted. She says that it is wrong to condemn something based on such speculation. In a later publication, she states that sex selection is not always a form of gendercide in that, "if it were inherently wrong to alter sex ratios, then it would be wrong to seek cures for heart disease, breast cancer, and other lethal illnesses which afflict primarily members of one sex" (44). She also holds that, "sexism and its potential for harm are very much a function of how it is done, why it is done , and the social context."Warren argues it would be wrong to condemn a practice outside of its social context.

Anand Kumar offers this social context from the perspective of the culture in India (45). In India, sons are considered an asset, while daughters are considered a great burden. The son preference is so strong that female infanticide is a prevalent occurrence, and although abortion and infanticide are illegal, legislative measures have failed to produce any change. Kumar submits that the real ethical choice lies between the prevention and the perpetuation of feticide, infanticide, and homicide of females. Kumar notes that social change is a long process and asks the question, "Can we afford to wait until these social changes occur and in the meantime silently witness female deaths at all stages of life?" In light of what is now taking place, one argument is that reliable methods of sex preselection would offer the least harm to this society.

In response to the position that that sex selection will lead to an altered sex ratio, Lilford points out that in many countries there is no real preference for one sex over the other but rather a preference for a balanced family (46). He then goes on to say that in countries where there is a strong preference for boys, sex selection may slow the population growth, and ultimately, the demand for girls would increase thereby eventually changing the direction of preference. In the same publication, Lilford refutes the claim that sex selection is a form of discrimination against women but rather a preference for a particular sex may be the result of discrimination. He states that, "peoples' choice for a particular sex is a mirror of their society."

Mahoney separates sex selection followed by abortion from sex preselection. He rationalizes that availability of sex preselection would decrease the number of gender-based abortions performed and the incidence of infanticide (47). Mahoney points out that in Great Britain abortion is legal due to serious social pressure, so it is reasonable that people have access to it for any reason including selecting the sex of the child. This argument has been used by others including Egozcue who wrote, "Sex selection: why not?" in the journal, *Human Reproduction* (48). He discusses the use of sex selection on embryos where those embryos of the unwanted sex are discarded. Egozcue submits that this should not be a problem in those countries where abortion is available on demand as this would be an extension of other assisted reproductive technologies. Also Egozcue states that planned parenthood organizations have always supported the view that every child should be a wanted child, and therefore a child of the desired sex very much is a wanted child. Smith, in a letter to the *British Medical Journal*, also discusses the issue of wanted children (49). He points out that in many families unwanted children are abused. Smith argues that even if sex selection results in an altered sex ratio, the scarcer sex would be valued over time, and since fewer children would be born, it would slow the population growth. In the same letter, Smith discusses the justice problems in attempting to regulate the practice of sex selection. He says, "the rich and connected can usually gain access to any technological innovation that they want."

Pennings addressed sex selection for balancing families and proposed ethical guidelines (50). He proposed that sex selection should not be allowed for the first child nor when there is already a balance in a family. In response to the Pennings proposal, Dawson and Trounson ask the question, "Who will enforce these guidelines?" Although they do not support sex selection, they do point out that the United Nations Declaration of Human Rights states that each individual has the right and freedom to form a family. The application of this declaration to the issue of sex selection is not clear. However, they feel that the Pennings proposal represents, "a violation of the right to freely form a family given in the Declaration of Human rights and, given that the appropriate technology is available." In the end, however, they argue that sex selection is not a responsible use of technology (51).

Lilford challenges the argument that technology should only be used for medical purposes (46). He submits that it is difficult to determine the difference between a medical and a nonmedical mission. He states that, "the important factors in human life are those of suffering and happiness, and the eradication of disease is merely a means towards these ends. If medical technology can produce these ends without eradicating a disease, then it is equally worthwhile."

One final major argument in support of sex selection is that of respect for individual autonomy. Kaye and LaPuma emphasize this position from the perspective of the clinical geneticist (52). They assert that although clinical geneticists are not themselves ethically neutral on this subject, "the best interests of the patient, not of society or the human race, should determine diagnosis and treatment." They strongly maintain that the overriding ethical principle is that of beneficence for the individual patient. Shulman, of the Genetics and IVF Institute in Fairfax Virginia, concludes that it is fundamental to free societies for responsible individuals to have the freedom to differ and to make personal choices based on their own convictions (53). Stephens demonstrates the strong belief that patient autonomy must be the guiding principle when he writes, "It is my opinion that the only issue that is the sole responsibility of physicians really is support of the patient's (and it is usually women who are burdened with this responsibility) right to exercise her (or, in the instance of a couple in a counseling situation, their) reproductive options, regardless of the indication, regardless of the personal moral or ethics standpoint held by the practicing physician" (54). Such wholehearted support of patient autonomy is most prevalent among geneticists in the United States (27).

Finally, an examination of the arguments surrounding sex selection would be incomplete without looking at the

work of John C. Fletcher who has written extensively on this topic and revised his ethical position over time. His initial arguments opposed sex selection because (*1*) sex is not a disease, (*2*) abortion for sex choice could contribute to social inequality between the sexes because of a preference for male offspring, (*3*) sex choice is a "frivolous" and indefensible reason for abortion, and (*4*) amniocentesis is a scarce resource in light of the total number of at-risk pregnancies (55). In reviewing his writings over the years, it is interesting to see how he has re-evaluated his position and his reasons behind his position many times. This speaks to the extreme complexity of this issue.

In this 1980 essay, Fletcher assumes that the main reason for discouraging sex selection is the belief of physicians that performing prenatal diagnosis that such abortions are morally unjustifiable (55). Although this is still his personal view, he believes that the legal rules on abortion defined by the U.S. Supreme Court supersede the clinician's personal moral views. Since a woman need not state reasons for abortion under any circumstances, sex selection ought not be subjected to public scrutiny. He concludes that, "it is inconsistent to support an abortion law that protects the absolute right of women to decide, and at the same time to block access to information about the fetus because one thinks that an abortion may be foolishly sought on the basis of that information" (55, p. 16). Fletcher continues to believe that physicians have a right to state their own moral views and to describe risk factors of prenatal testing, including "an unknown risk of insult to other numbers of the family and to wider society." However, if a couple continues to request the information for sex selection, Fletcher stated that the physician may not legally or morally refuse to provide it if he or she wants to "keep faith" with the moral intent of the law.

In a later re-evaluation of this issue, Fletcher, along with Wertz, considered all of the arguments given in support of sex selection and came to the conclusion that the medical profession as a whole has not demonstrated any responsibility in this arena. They state, "We hold that a very strong normative case exists against sex selection that transcends cultural boundaries, especially based on claims of equal worth of both sexes and justice in social life" (56). They strongly suggest that the medical community take a stand against sex selection. Fletcher and Wertz believe that by doing this, the medical community will ultimately be protecting important reproductive choices by demonstrating that they are able to set the standards for practice. With these standards in place, government intervention would not be necessary and reproductive choices involving medical decisions would not be lost (56). In a more recent study, Fletcher, along with Wertz, evaluated the attitudes of medical geneticists about sex selection in 19 nations. They found that in many nations women do not have access to prenatal testing to detect birth defects either because of the cost or because of the scarcity of such medical technology. They conclude that it is unfair to use these limited resources for nonmedical reasons. They state a concern that sex selection could be, "the first step on a 'slippery slope' toward cosmetic choices for height, weight, eye or hair color." Again, they call for the medical profession to abandon its nonjudgmental stance and set a standard of care with regard to sex selection (57).

SUMMARY

With few exceptions, positions of professional societies and governmental agencies oppose the practice of sex selection. Yet, sex selection continues to occur in practice, and the debate about the morality of the practice continues among ethicists and practitioners. Why is that?

There are two major reasons. One is that sex selection is only one of the possible traits that one can select prenatally. Most prenatal testing is accepted today because it serves the interests of people who want to make decisions based on health information. Even for many who are comfortable with prenatal testing and autonomous choice, however, there is concern about the intrinsic value of human life as it is created. Choosing the sex of their offspring represents the first real and available choice for parents who want to select a child with traits that fit their vision of an ideal family. The specter of sex selection could be the first step down a slippery slope to the "brave new world" of designer children. We worry about whether allowing choice of the sex of the child will open the door to the use of genetic technologies to select other traits for more "trivial" reasons than avoiding disease (55).

Second, sex selection also is a paradigm case for considering what values really ought to guide health care policy. The same tensions, between issues of justice for many and respect for individual autonomy, exist in determining the appropriate use and distribution of health care technologies more generally. A majority of those opposing sex selection address broad societal justice concerns. Some say that using such measures will continue, and even enhance, gender discrimination. Others believe that health technologies ought to serve the needs of improving health, and ought not be squandered in support of individual or societal determination of what human traits are valuable and worthy.

On the other side of the argument, are those who strongly support the individual's right to self-governance and the professionals' obligation to respect that autonomy. Individual (or couple) autonomy undergirds most arguments for allowing sex selection, since prohibiting the practice would be limiting autonomy in reproductive decision making. Genetic and reproductive technologies, particularly sex selection, make the professional obligation to respect autonomy more complex, however. Because decisions about how to use them are intrinsically about families they raise the question of who is the patient, and thus whose interests ought to be served by clinicians and policy makers. Is the primary obligation to the mother, the father, both parents, the child or potential child, existing children, or to the societies in which reproductive and genetic applications are made available?

Both clinicians and policy makers must ponder the right and appropriate use of health care technologies and from the framework that guide how these decisions are made. Sex selection is but one example of this challenge. Examining how positions on sex selection are cast may provide some insight into how other ethical challenges in health care will be addressed as well.

ACKNOWLEDGMENTS

The authors would like to thank the Center for Bioethics at the University of Minnesota for its support. In addition, we would like to extend special appreciation to Robert Koepp, a Center research assistant, for his assistance.

BIBLIOGRAPHY

1. B.L. Gledhill, *Semin. Reprod. Endocrinol.* **6**(4), 385–395 (1988).
2. B.E. Reubinoff and J.G. Schenker, *Fertil. Steril.* **66**(3), 343–350 (1996).
3. A. Botchan et al., *J. Androl.* **18**(2), 107–108 (1997).
4. S.A. Carson, *Fertil. Steril.* **50**(1), 16–19 (1988).
5. R.J. Ericsson, *Fertil. Steril.* **51**(2), 368–369 (1989).
6. J.H. Check, D. Katsoff, and A. Bollendorf, *Fertil. Steril.* **61**(6), 1181–1182 (1994).
7. J.H. Check and D. Katsoff, *Hum. Reprod.* **8**(2), 211–214 (1993).
8. J.J. Beernik, W.P. Dmowski, and R.J. Ericsson, *Fertil. Steril.* **59**(2), 382–386 (1993).
9. S.P. Flaherty et al., *Hum. Reprod.* **12**(5), 938–942 (1997).
10. M. Greier, J.L. Young, and D. Kessler, *Fertil. Steril.* **53**(6), 1111–1113 (1990).
11. R.G. Edwards and H.K. Beard, *Hum. Reprod.* **10**(4), 977–978 (1995).
12. S.P. Flaherty and C.D. Matthews, *Mol. Hum. Reprod.* **2**(12), 937–942 (1996).
13. I. Findlay et al., *J. Assist. Reprod. Genet.* **13**(2), 96–103 (1996).
14. J.C. Harper, *J. Assist. Reprod. Genet.* **13**(2), 90–95 (1996).
15. J.T. Queenan, *Semin. Perinatol.* **11**(3), 264–267 (1987).
16. C. Ruegsegger Veit and R. Jewelewicz, *Fertil. Steril.* **49**(6), 937–940 (1988).
17. T.M. Martequ, *Br. Med. J.* **306**, 1704–1705 (1993).
18. H. Statham et al., *Lancet* **341**, 564–565 (1993).
19. A. Green, C. Wray, and B. Balsuch, *Psychol. Rep.* **73**(1), 169–170 (1993).
20. P. Liu and G.A. Rose, *Hum. Reprod.* **10**(4), 968–971 (1995).
21. M.A. Khatamee et al., *Int. J. Fertil.* **34**(5), 353–354 (1989).
22. P. Baird, *J. Assist. Reprod. Genet.* **12**(8), 491–498 (1995).
23. T. Sheldon, *Br. Med. J.* **311**, 10–11 (1995).
24. S.K. Khanna, *Soc. Sci. Med.* **44**(2), 171–180 (1997).
25. R.V. Grazi, B. Wolowelsky, and R. Jewelewicz, *Gynecol. Obstet. Invest.* **37**, 217–225 (1994).
26. M.I. Evans et al., *Am. J. Obstet. Gynecol.* **164**(4), 1092–1099 (1991).
27. D.C. Wertz and J.C. Fletcher, *Am. J. Med. Genet.* **29**, 815–827 (1988).
28. B.M. Burke, *Soc. Sci. Med.* **43**(11), 1263–1269 (1992).
29. D.F. Pencarinha et al., *J. Genet. Couns.* **1**(1), 19–30 (1992).
30. D.C. Wertz, *Clin. Obstet. Gynecol.* **36**(3), 521–531 (1993).
31. X. Mao and D.C. Wertz, *Clin. Genet.* **52**(2), 100–109 (1997).
32. American College of Obstetricians and Gynecologists (ACOG), *Int. J. Gynecol. Obstet.* **56**, 199–202 (1997).
33. Council on Ethical and Judicial Affairs, American Medical Association, *Arch. Fam. Med.* **3**, 633–642 (1994).
34. Kalaca and A. Akin, *Hum. Reprod.* **10**(7), 1631–1632 (1995).
35. H.B. Holmes, in J.M. Humber and R.F. Almeder, eds., *Biomedical Ethics Reviews*, Humana Press, Clifton, NJ, 1985, pp. 39–71.
36. R.V. Grazi and J.B. Wolowelsky, *J. Assist. Reprod. Genet.* **9**(4), 318–322 (1992).
37. P. Shrivastav, *Hum. Reprod.* **10**(5), 1319–1320 (1995).
38. M.M. Seibel, S.G. Seibel, and M. Zilberstein, *Hum. Reprod.* **9**(4), 569–570 (1994).
39. Shenfield, *Hum. Reprod.* **9**(4), 569–570 (1994).
40. R. Steinbacher, in H.B. Holmes, B.B. Haskins, and M. Gross, eds., *The Custom-Made Child? Women Centered Perspectives*, Humana Press, Clifton, NJ, 1981, pp. 187–191.
41. P. Baird, *Annu. Rev. Med.* **47**, 107–116 (1996).
42. J.R. Botkin, *Obstet. Gynecol.* **75**(5), 875–880 (1990).
43. M.A. Warren, in J.M. Humber and R.F. Almeder, eds., *Biomedical Ethics Reviews*, Humana Press, Clifton, NJ, 1985, pp. 73–89.
44. M.A. Warren, *Bioethics* **1**(2), 189–198 (1987).
45. C. Anand Kumar, *Hum. Reprod.* **10**(5), 1319 (1995).
46. R.J. Lilford, *Hum. Reprod.* **10**(4), 762–764 (1995).
47. J. Mahoney, in P. Byrne, ed., *Medicine, Medical Ethics and the Value of Life*, Wiley, New York, 1990, pp. 141–157.
48. J. Egozcue, *Hum. Reprod.* **8**(11), 1777 (1993).
49. T. Smith, *Br. Med. J.* **307**, 451 (1993).
50. G. Pennings, *Hum. Reprod.* **11**(11), 2339–2343 (1996).
51. K. Dawson and A. Trounson, *Hum. Reprod.* **11**(12), 2577–2578 (1996).
52. C.I. Kaye and J. LaPuma, *Hastings Cent. Rep.* **20**(4), 40–41 (1990).
53. J.D. Shulman, *Hum. Reprod.* **8**(10), 1541 (1993).
54. J. Stephens, *Am. J. Obstet. Gynecol.* **166**(3), 1024–1025 (1992).
55. J.C. Fletcher, *Hastings Cent. Rep.* **10**(1), 15–20 (1980).
56. D.C. Wertz and J.C. Fletcher, *Hastings Cent. Rep.* **19**(3), 21–27 (1989).
57. D.C. Wertz and J.C. Fletcher, *Soc. Sci. Med.* **37**(11), 1359–1366 (1993).

See other entries GENETIC COUNSELING; see also REPRODUCTION, ETHICS entries; REPRODUCTION, LAW, REGULATION OF REPRODUCTIVE TECHNOLOGIES.

REPRODUCTION, ETHICS, THE ETHICS OF REPRODUCTIVE GENETIC COUNSELING: NONDIRECTIVENESS

BARBARA BOWLES BIESECKER
National Human Genome Research Institute
Bethesda, Maryland

OUTLINE

INTRODUCTION

In reproductive genetic counseling, nondirectiveness may refer to an ethic of practice or to the process itself. Different aspects of genetic counseling have been described as nondirective; the communication style, the offering of genetic testing, or the counseling interaction. These various interpretations of the term *nondirectiveness* have lead to confusion about the goals and practice of reproductive genetic counseling. As well, it has diluted conversation about important issues surrounding the personal nature of reproductive choice involving genetic risk. As an ethical principle, nondirectiveness suggests that pregnant women and their partners ought to be supported to make autonomous decisions about prenatal testing and their reproductive outcomes without the direct influence of the counselor. The personal autonomy of the client facing the genetic reproductive decision is paramount. Nondirectiveness should be used exclusively to describe an ethical principle in reproductive genetic counseling. Although it is not evident always how this principle translates into the practice of genetic counseling, the process may be discussed as a dialogue of client-centered counseling that is guided by nondirectiveness.

NONDIRECTIVENESS IN GENETIC COUNSELING

Nondirectiveness describes components of a young medically related professional service, called *genetic counseling*. This psychoeducational practice assists people who have concerns about birth defects, genetic conditions, or genetic risk (1–3). Throughout its short history, genetic counseling has been consistently described as nondirective, as opposed to advice giving. Genetic counseling may be the only medically related practice intended to be nondirective. The term has been used to describe not only the ethic of practice but also the goal of genetic counseling, the process, and an outcome. The literature discusses nondirectiveness assuming one of these practice components but frequently fails to distinguish its meaning. For those who strive to understand, to investigate, or to use genetic counseling services, it is unfortunate that the concept is inconsistently portrayed. Even those who practice genetic counseling have confused the meaning and interpretations (4,5).

Several scholarly articles have appeared to address the confusion in the meaning of nondirectiveness (4,6,7). The literature has begun to distinguish the various uses of the term in an effort to achieve some consensus on the goals and process of genetic counseling. Since the literature on nondirectiveness is discrepant, this chapter will delineate uses of the term and compare their implications. The success of the practice of genetic counseling depends on continued efforts to define and strive towards nondirectiveness assuming the profession can agree on what it is, that it is central to the process and that it can be achieved.

HISTORY OF NONDIRECTIVENESS

The original introduction of nondirectiveness into the genetic counseling literature remains elusive. Sheldon Reed, a medical geneticist who coined the phrase "genetic counseling" in 1947, spoke of a nondirectivelike practice but only used the term later in his writings after it had appeared in the literature (8). Reed described a genetic social worklike practice of explaining genetic concepts and supporting clients who use the information to make reproductive decisions. In this case the concept of nondirectiveness describes the process of genetic counseling more so than the overarching ethical principle. Some authors claim that nondirectiveness in genetic counseling arose in opposition to the eugenics movement. Resta points out that many of the medical geneticists writing about the process of genetic counseling in the 1950s used the term nondirective but then also described eugenic ideas about the practice (9,10). It is evident from the literature that certain supporters of nondirectiveness were not opposed to eugenic practices. Thus, such claims about nondirectiveness may be unfounded (4,11).

The general source of the term nondirectiveness predated genetic counseling by about two decades. Dr. Carl Rogers, a prominent psychologist, used the word to describe his theory of psychotherapy (12). By 1951, however, Rogers had come to describe his theory and practice as client-centered. This clarification in his terminology acknowledged the presence of directive components to the therapeutic relationship, yet emphasized the focus on the client's expressed needs rather than the explicit direction of the counselor. Rogerian psychotherapy developed prior to the existence of genetic counseling and in parallel to, not in reaction against, the eugenics movement in the United States. It is intriguing to consider why the profession of genetic counseling adopted as its mantra a term that was rejected early on by the field of psychotherapy. Since its introduction into genetic counseling, nondirectiveness has lead researchers, academics, and practitioners astray.

Genetic counseling has sustained the use of the term nondirectiveness despite its ambiguity. Client-centered theory and practice have offered one useful framework (within limits) of thinking about and practicing genetic counseling. Nondirectiveness has been used effectively to describe a client-centered counseling style not unlike a Rogerian approach. Since genetic counseling has evolved as a clinical and atheoretical practice, it has borrowed ideas from its theoretical neighbors. In a different sense, genetic counseling has long recognized the lack of desire or ability to make reproductive decisions for others. It has emphasized autonomy and voluntariness (13,14). Genetic counseling embraces a certain hands-off approach to sensitive issues of life and death that are entwined in reproductive decision making. This has proved to be

a more comfortable stance for genetic counselors than entering into the difficult and sticky terrain of directing people in their childbearing decisions that involve genetic risk. Rather than as a reaction against eugenics, perhaps nondirectiveness has been sustained by an abhorrence of eugenic practices. Some would argue that it might also serve to shield practitioners from confronting difficult aspects of reproductive genetic counseling.

INTERPRETATIONS AND IMPLICATIONS OF NONDIRECTIVENESS

Nondirectiveness as a Guiding Ethical Principle

There have been at least four different, yet overlapping, meanings of nondirectiveness expressed in the literature. Most often, nondirectiveness has been used to mean a desire to uphold the personal nature of reproductive decision making. Nondirectiveness in this sense represents an underlying value or ethical principle of the profession. Genetic counselors in the United States have emphasized the principle of nondirectiveness conceptionally in their code of ethics: "Genetic counselors strive to enable their clients to make informed independent decisions, free of coercion, by providing or illuminating the necessary facts and clarifying the alternatives and anticipated consequences (15, p. 41). Yet as a value it does not readily translate into a way of practice or a specific goal. It is difficult to assess whether an individual or couple has made a "good" personal decision. How does a genetic counselor promote the decision-making process within an ethical framework of nondirectiveness? Clients experience many influences on their reproductive decisions. Exclusively personal or autonomous decision making is difficult, and not necessarily uniformly desirable, to achieve. Yet it is important that providers not assert undue influence on the reproductive outcomes of their clients. This is a blatantly eugenic goal and contradicts the desires of most geneticists and genetic counselors internationally (16). Kessler points out that even when there is an explicit goal to discourage certain reproductive outcomes (e.g., in a country that supports such practice), a significant number of clients ignore the advice (17). It is unclear that it is necessary for professionals to completely withhold advice from clients. Yet much of the international genetics community, and in particular, genetic counselors in the United States, Canada, and the United Kingdom, finds the notion of advising people directly on their reproductive choices to be loathsome. It is difficult to know or to appreciate the values, resources, thought-processes, and ideas of another person sufficiently to provide advice about having or not having children who may be affected with a certain genetic condition. The truth is, most people struggle to understand what choices they would make for themselves, let alone know better for another.

Genetic counselors need to be exquisitely self-aware and not harbor personal opinions of what constitutes a life worth living. If they do, they must be honest with themselves and disclose to clients that they may hold beliefs that children affected with certain genetic conditions should not be born. This differs from stated goals to enhance personal choice for clients. Yet it is more honest than undisclosed potential agendas. Most counselors, who also work with children and families affected with genetic conditions, serve as advocates for those with special needs as children or who are disabled as adults. As a profession they value diversity and often enter the field of genetic counseling concerned about genetic conditions and how society views disability.

Medical genetics services, such as triple screening for neural tube defects and carrier screening for recessive or sex-linked genetic conditions, may have the more or less explicit goal of reducing the number of individuals affected with genetic conditions (18–20). Cost–benefit analyses to justify such programs may be based on an assumption that a significant number of affected pregnancies will be aborted. In this case the genetic counseling that accompanies such practices may have values that are in conflict with the intention to reduce the incidence of genetic conditions. Genetic counseling may strive to help the individual make the best personal decision, yet the goal of the program may be to abort affected fetuses. Genetic counselors may find themselves caught in a dilemma between professional values that emphasize personal autonomy and programs that are justified by social policy to improve the health or well-being of the populace. If counselors uphold a nondirective ethic, then they should not paradoxically endorse genetics services that have a goal of preventing the birth of individuals who will be affected with a genetic condition. Genetic counselors should and do endorse services that emphasize informed and autonomous choice in reproductive decision making. An example is the choice about whether to undergo amniocentesis to determine the chromosomal status of a fetus. Nevertheless, aspects of service provision (e.g., assuming the outcome of the decision to undergo testing by scheduling the amniocentesis to follow the counseling session) do not always promote the genetic counselor's role to ensure personal choice about testing.

A challenging aspect of an ethic of nondirectiveness is not so much the goal to refrain from explicit influence on reproductive decisions, as it is the potential for more subtle and unintended (even unconscious) influence. Such practice may occur when a genetic counselor harbors a belief that a certain reproductive outcome is most desirable for a person or couple. But rather than state the bias outright, the counselor's approach is influenced by her or his beliefs. This would be an ethically directive approach even if the counselor did not intend to provide direction to the client.

When counselors successfully manage to facilitate the client's decision making without influencing the outcome, the process is flexible and difficult to operationalize. Counseling is inherently directive, as is providing education to ensure understanding about the options. Genetic counselors have no standard of practice to consistently uphold an ethic of nondirectiveness. Counselors recognize that the type of information they provide and how they present it may influence decisions (21,22). An ethical principle should translate into an effective mode of practice. White has proposed a counselor–client dialogue as a working description of the process (7). The practice is to facilitate client centered reproductive decision making, within

an ethical framework of nondirectiveness. As a guiding principle this ethic would suggest a process of genetic counseling that emphasizes the values and beliefs of the client, but that tolerates the direction offered by a competent counselor who does not preconceive a decision for the client. In order to further clarify the underlying ethical principles of genetic counseling, the field may need to differentiate itself from other genetic services whose goals (such as abortion of affected pregnancies) are inconsistent with the values of the profession.

Nondirectiveness as a Guiding Policy on Genetic Testing

Nondirectiveness also has been used to describe the concept of not denying access to genetic testing. This definition relates to genetics health policy and access to services. It is a practical one, although it has overlap with the previous definition in its intention to uphold reproductive freedoms. In this case genetic counselors are reticent to deny access to any genetic test that an individual or couple may request (even one that puts a pregnancy at risk) provided that there is understanding about the risks and benefits of the test and its potential outcomes.

Historically much of genetic counseling has addressed risk for serious conditions, with the exception of certain mild birth differences (such as a cleft lip) and sex chromosome "anomalies" (such as Turner syndrome). Counselors offer genetic testing to determine whether a fetus will be affected (prenatal) or whether a couple may be at increased risk for having an affected child (carrier). One survey has suggested that the majority of genetic counselors, internationally, offer prenatal diagnosis (or a direct referral) for sex selection (23). This is worrisome when one considers that genetic counseling originated from a desire to help people grapple with difficult dilemmas about serious genetic conditions or birth defects. Genetic counselors, as represented by the U.S. professional society (NSGC), uphold a moral right to reproductive freedom (24). The majority of practitioners believe that if a woman (or couple) has consented to prenatal testing by considering the relative balance of risks and benefits, she should be offered the opportunity to determine the sex of the fetus, even if she desires to abort a fetus of undesired sex. Such a finding bodes poorly for the future of genetic testing, when prenatal tests may be offered for physical or personality traits. Will genetic counselors, in the name of nondirectiveness, offer prenatal testing for anything a couple desires as long as they are informed?

In this regard, the meaning of nondirectiveness has caused the profession of genetic counselors to be passive about taking a stand on what tests ought to be offered. There have not yet been professional guidelines written by U.S. genetic counselors discouraging certain prenatal testing. Within NSGC there are position statements and a resolution on genetic testing or screening, for instance, one exists on prenatal and childhood testing for adult-onset disorders (25). It states that while such testing is discouraged, each case should be considered individually and counselors should decide whether to offer parents testing of their children on a case by case basis. This leaves the judgment of reasons, fitness, and values of the client up to the counselor. While inherently flexible and accounting for individual differences, it neglects to take a clear stand and puts counselors in the position of practicing inconsistently. It leads to confusion for the profession. In response, members of the genetic counseling community published a substantial position statement on the genetic testing of children for adult onset conditions to more clearly state a testing policy (26). While it is unlikely that there are many moral absolutes in reproductive decision making, an insistence on nondirectiveness has stymied the process of policy making in prenatal genetic testing. With the promised onslaught of new genetic tests, reproductive counselors seem to be prepared to offer testing for any indication. In the name of nondirectiveness, genetic counselors have avoided their professional and moral obligations to take a stand on the appropriateness of certain types of prenatal testing.

The approach to reproductive genetic testing, "anything goes as long as the individual has had pre-test counseling," predicts that counselors will play less of a role in establishing genetics health policy. Yet genetic counselors may be one of the most important groups of professionals to be involved in helping to establish guidelines or polices about what testing may not be an appropriate use of resources or may be morally reprehensible (27). Do genetic counselors advocate for the use of prenatal testing to potentially abort fetuses found to be at somewhat increased risk for adult-onset cancer, for instance? Worse yet, for a slightly lower projected adult height? The role of testing gatekeeper may be an important one for genetic counselors in the future. Yet nondirectiveness has been misinterpreted to imply that any genetic tests that are technically feasible should be offered. In the name of nondirectiveness, counselors refrain from judging the choices of their clients. In doing so, genetic counselors may be washing their hands of the responsibility to offer morally, not to mention economically, responsible reproductive testing options.

Rather than interpreting that nondirectiveness holds no opinion on genetic testing, reproductive genetic counselors ought to offer genetic testing only for serious conditions that may significantly impede an individual's ability to achieve quality of life. While there is no consensus on what constitutes a serious genetic condition (28), this should not dissuade genetic counselors and other providers from establishing responsible genetic testing services and genetics health policy (29,30). This misunderstanding of nondirectiveness has led to a significant lost opportunity and an ongoing need for the professional practice of reproductive counseling.

Nondirectiveness as a Style of Communication

In contrast, nondirectiveness has been construed as a style of communication within the practice of genetic counseling (31,32). Genetic counseling has been described as a value-neutral encounter despite awareness that any human relationship is value laden. The mis-notion of value neutrality has further confused the issue of nondirectiveness (33). In communicating genetic information within genetic counseling, there are many directive components. In an educational relationship, the person with

the information has more power and there is an inequality to the relationship (1). The way the information is conveyed and the amount of information given may be quite directive. Genetic counselors lead, guide, and even advise their clients. Each of these is a directive process. Genetic counselors strive to give complete and balanced information, but it is human nature to be inconsistent and influenced by individual experiences. This might be described as directive practice as well.

This interpretation of nondirectiveness implies information should be conveyed in a nonleading way. Studies that have been conducted to assess use of directive language have concluded that the process is directive (30). While such outcome studies are necessary and valuable for determining what happens in genetic counseling, they seek to document a nondirective psychoeducational practice. It is an unattainable paradox.

In a desire to use nondirective language to communicate, genetic counselors may seek to use words that are ambiguous. Such avoidance of direct language may not be useful to clients who are often seeking not only information but also advice on how to use it or how to make meaning of it. The irony of the use of vague language is that expert communication of complex genetics information is often heralded as a prominent goal of genetic counseling. A nondirective intent has guided counselors into inexplicit use of language that could otherwise make genetic concepts and their implications more obvious to clients. This use of nondirective communication has lead to process studies that have shed light on this perplexing notion of genetic counseling (31,34). Conclusions have been drawn that counselors are directive in a manner that implies they are undermining a guiding ethical principle. In fact they are merely communicating as professionals do, using language that is often directive and in a manner that may be directive. While there is merit in research toward understanding how an ethical principle such as nondirectiveness translates into practice, the mode of communication is only one component of a complex dialogue within a relationship of influence.

Without guidance on the adaptation of an ethic of nondirectiveness into practice, counselors have assumed a nonjudgmental approach that also involves noncommittal or evasive communication around difficult issues. This minimization of an ethical principle has led to one outcome counselors seek: clients who have not been explicitly directed in their reproductive decisions. But it has also lead to not providing clear messages about the implications of the information, and perhaps even not facilitating "good" reproductive decisions. For instance, Wertz and her colleagues found that in the majority of prenatal genetic counseling sessions they studied, abortion was not mentioned (35). Since it is the only intervention a couple could choose to take for the vast majority of conditions tested for, it should be prominent in discussion about the potential usefulness of prenatal testing. If nondirectiveness had been uniformly adopted as an ethical principle that supports a client-centered counseling process, rather than a communication style, word choice and tone would be considered less significant than the components of dialogue within a therapeutic relationship.

Nondirectiveness as a Theoretical Basis of Counseling Practice

This point segues into a further interpretation of nondirectiveness, the intent to provide client-centered counseling. This definition is not dissimilar to that of Rogers's theory of psychotherapy. In this regard, nondirectiveness represents a reasonable and responsible goal for genetic counseling. It heralds the role of the client as central and as a goal, can be achieved (1,36). In this venue nondirectiveness provides a model for genetic counseling that can uphold an ethic of personal reproductive decision making. But nondirectiveness should not be used to describe both the process of counseling as well as the underlying ethical framework or the existing confusion will pervade.

In describing genetic counseling as a psychoeducational process, the psychological or therapeutic goal is to explore the meaning that the genetic information has for clients. This is a client-centered approach that focuses on client values, beliefs, ideas, and desires. The process by which it is achieved varies but the client's agenda and needs are paramount. As Rogers previously discovered, the term nondirectiveness in this case compounds the confusion, since many strategies used by the counselor might be described as directive. Yet they are executed with the client in mind. For instance, the counselor may help the client to set an agenda to explore implications of the information in a way that is personal, useful, and lends itself to decision making. The counselor may be directive in helping a client determine what may be reasonable to try and accomplish in one or two sessions. While these behaviors are directive, they do not override the needs of the client. Rather, they represent the counselor's expertise that may be used to enhance the effectiveness of genetic counseling. The client's needs are the ones addressed, but the client is not left to talk randomly without focus on the problem or issue at hand. Without such structure, a session would never become therapeutic. This is only one example of directive practices of counselors that do not undermine a client-centered approach.

In this more appropriate use of nondirectiveness the term remains problematic and should be replaced with client-centered genetic counseling. In much the same way nondirectiveness did not accurately depict the therapeutic approach proposed by Rogers, it has lead genetic counselors to largely ignore the need to engage actively with clients in order to address their concerns. Transcripts of genetic counseling sessions indicate that counselors practice inadequate counseling skills to accomplish even a minimally client-centered approach (1,5,31). The mantra of nondirectiveness may have caused counselors to hesitate over using their own best judgment about people's ability to make good decisions for themselves, to grow from difficult experiences, and to cope and adjust. In the name of nondirectiveness, many counseling opportunities have been lost in genetic counseling. An active dialogue about options, alternatives, resources, strengths, and outcomes within a therapeutic relationship may best help clients (3,6,7). Such a dialogue is likely to have many directive statements in it but does not direct the client toward a certain outcome in a coercive or even persuasive way.

Although the various uses of the term nondirectiveness have been problematic for the profession of genetic counseling, the pervasiveness of the concept of personal autonomy in genetic reproductive decision making sets the practice apart from the majority of medical services. The ethical principles of autonomy and beneficence in reproductive decision making, the need for thorough informed consent for genetic testing, and the value of human diversity and the lives of those affected with disabilities are crucial to reproductive genetic counseling. And the few outcome studies that have been conducted suggest that reproductive genetic counseling clients are satisfied with the service. They like their genetic counselor and are grateful for the time spent teaching them genetic principles (37). From a process and outcomes perspective, clients are likely to be best served by a psychoeducational approach that includes a client-centered or cognitive, theoretically based practice. This therapeutic process and its desired outcomes of self-determination, feelings of personal control, and restored self-esteem have yet to be studied (38). To achieve them, a counselor may be directive and to conduct process studies that investigate how directive she or he is seems counterproductive. Studies are needed on the most effective therapeutic approaches in genetic counseling, to observe how successful they are in achieving desired outcomes. Research would be facilitated if genetic counselors embraced such a therapeutic approach, and if consensus could be achieved on the goal of restoring feelings of personal power to clients and on outcomes of the process that can be systematically measured.

NONDIRECTIVENESS IN PRACTICE

In the most common reproductive genetic counseling example of a couple learning that their fetus is affected with Down syndrome (due to an extra chromosome 21), genetic counseling is the process through which the couple can determine what the condition may mean for their lives. Down syndrome cannot be "repaired," although some of the symptoms can be treated. The child will be mentally retarded, although to what degree is unknown. The couple may continue the pregnancy as planned or have an abortion. This is an agonizing decision even for couples who initially feel confident about what they would do in such circumstances. In facing the situation couples often take into account their expectations of parenthood, family life, economic resources, previous experiences with persons who have Down syndrome, opinions of family and friends, spiritual beliefs, social influence and expectations, and so on. Decisions about a pregnancy are complex, deeply personal, and irreversible. Important aspects of the decision may even be intangible or elusive to the couple themselves.

A genetic counselor in this situation seeks to establish an empathic connection or a therapeutic bond with the couple in order to help them make personal meaning of the information about Down syndrome. The counselor may strive to identify resources useful to the couple in making the decision so that they can live with their decision (one way or the other) in the years to come. A therapeutic approach focuses on enhancing self-determination and perceived personal control. Couples are helped to recognize that they have the strength to make such a difficult decision and that they have made other decisions successfully in the past. The counselor works toward facilitating the decision-making power of the couple in addressing their needs and concerns. This process may be described by some as nondirective counseling. Yet it is more appropriately described as client-centered and personally empowering.

The example of a client or couple asking the counselor what he or she would do in the same circumstances is often used to illustrate nondirectiveness in genetic counseling. Common responses by genetic counselors may be:

- "I am not in your situation so I couldn't possibly know what I would do."
- "Other people in your situation have chosen to continue the pregnancy, while others have had an abortion."
- "There are no right answers, I am here to help you make the best decision for yourself."
- "I will support any decision you make."

Evasive answers such as these do not address the concerns of the client. The client is asking for advice because she has not received the help she needs to make her decision. It is unlikely that she is literally handing over responsibility for the decision to the counselor (although a minority of clients may do so). Nor is she likely to mimic the choice of the counselor in order to solve her dilemma. However, all too often counselors neglect to work toward exploring and understanding where the client's anxiety and concerns come from in an effort to best help her with the decision. Kessler reminds genetic counselors that if the clients are frequently asking this question, there is something fundamentally flawed about the process (4). There are many respectful and considerate ways to address this question without abandoning the client in a time of great need. They challenge genetic counselors to fully experience with clients some of the hardest decisions of their lives. It takes a lot of hard work and *direction* on behalf of the counselor. A nondirective mode of practice misinterpreted is a missed counseling opportunity and at its worse an abandonment of a client in need of help.

SUMMARY

Nondirectiveness is a term to be reserved for an ethical principle of practice in reproductive genetic counseling. It emphasizes the importance of autonomy in genetic reproductive decision making. As a guiding principle, nondirectiveness provides a moral framework for providing client-centered counseling. Reduced to merely how a counselor communicates or to a lack of health policy on the use of genetic tests, nondirectiveness is an ineffectual concept. Its counterpart, direction, is an essential aspect of effective client centered counseling that supports informed reproductive choices involving genetic risk.

BIBLIOGRAPHY

1. S. Kessler, *J. Genet. Counsel.* **6**(2), 287–295 (1997).
2. R. Kenen and A.C.M. Smith, *J. Genet. Counsel.* **4**, 115–124 (1995).
3. B. Biesecker, *Kennedy Inst. Ethics J.* **8**, 145–160 (1998).
4. S. Kessler, *Am. J. Med. Genet.* **72**, 164–171 (1997).
5. F. Brunger and A. Lippman, *J. Genet. Counsel.* **4**, 151–167 (1995).
6. S. Suter, *Univ. of Chicago Law School Roundtable* **3**(2), 473–489 (1996).
7. M. White, *J. Genet. Counsel.* **6**(3), 297–313 (1997).
8. S.C. Reed, *Soc. Biol.* **21**, 332–339 (1974).
9. R. Resta, *J. Genet. Counsel.* **6**(2), 255–257 (1997).
10. D. Paul, *Controlling Human Heredity 1865 to the Present*, Humanities Press, Atlantic Highlands, NJ, 1995.
11. B. Fine, in D. Bartels, B. LeRoy, and A. Caplan, eds., *Prescribing Our Future: Ethical Challenges in Genetic Counseling*, Aldine de Gruyer, New York, 1993, pp. 101–117.
12. C. Rogers, *Counseling and Psychotherapy*, Houghton Mifflin, Boston, MA, 1942.
13. M. Yarborough, J. Scott, and L. Dixon, *Theor. Med.* **10**(2), 139–149 (1989).
14. B. LeRoy, in D. Bartels, B. LeRoy, and A. Caplan, eds., *Prescribing our Future: Ethical Challenges in Genetic Counseling*, Aldine de Gruyer, New York, 1993, pp. 39–54.
15. National Society of Genetic Counselors, *J. Genet. Counsel.* **1**(1), 41 (1992).
16. D. Wertz and J. Fletcher, *Am. J. Hum. Genet.* **42**, 592–600 (1988).
17. A. Czeizel, J. Metneki, and M. Osztovics, *J. Med. Genet.* **18**, 91–98 (1981).
18. P. Rowley, S. Loader, and R. Kaplan, *Am. J. Hum. Genet.* **63**(4), 1160–1174 (1998).
19. D. Asch et al., *Med. Decis. Making* **18**(2), 202–212 (1998).
20. D. Asch et al., *Am. J. Public Health* **86**(5), 684–690 (1996).
21. G. Loeben, T. Marteau, and B. Wilfond, *Am. J. Hum. Genet.* **63**, 1181–1189 (1998).
22. N. Press and C. Browner, *Soc. Sci. Med.* **45**(7), 979–989 (1997).
23. D. Wertz and J. Fletcher, *Soc. Sci. Med.* **46**(2), 255–273 (1998).
24. National Society of Genetic Counselors, Resolution Adopted 1987.
25. National Society of Genetic Counselors, Resolution Adopted 1995.
26. W. McKinnon et al., *JAMA* **278**(15), 1217–1220 (1997).
27. K. Nolan, *Hastings Center Rep.* **22**(4), S2–S4 (1992).
28. B. Knoppers et al., *Am. J. Hum. Genet.* **57**(4, suppl.), A296 (1995).
29. J. Botkin, *Hastings Center Rep.* **25**(5), 32–39 (1995).
30. E. Parens and A. Asch, *Hastings Center Rep.* (September–October) spec. suppl. 1–22 (1999).
31. S. Michie et al., *Am. J. Hum. Genet.* **60**, 40–47 (1997).
32. B. Bernhardt, *Am. J. Hum. Genet.* **60**, 17–20 (1997).
33. A. Caplan, in B. LeRoy, D. Bartels, and A. Caplan, eds., *Prescribing our Future: Ethical Challenges in Genetic Counseling*, Walter de Gruyter, New York, 1993, pp. 149–168.
34. M. Rose, J. Benkendorf, and M. Prince, *J. Genet. Counsel.* **7**(6), 497–498 (1998).
35. D. Wertz, *J. Genet. Counsel.* **7**(6), 499–500 (1998).
36. G. Wolff and C. Jung, *J. Genet. Counsel.* **4**(1), 2–25 (1995).
37. B. Bernhardt, B. Biesecker, and C. Mastromarino, *Am. J. Med. Genet.*, in press.
38. S. Kessler, *J. Genet. Counsel.* **8**(6), 333–343 (2000).

See other entries GENETIC COUNSELING; see also REPRODUCTION, ETHICS entries; REPRODUCTION, LAW, REGULATION OF REPRODUCTIVE TECHNOLOGIES.

REPRODUCTION, LAW, IS INFERTILITY A DISABILITY?

LAURA F. ROTHSTEIN
Louis D. Brandeis School of Law
University of Louisville
Louisville, Kentucky

OUTLINE

INTRODUCTION

Infertility is defined as the inability to conceive a child within a period of one year. There are many causes of infertility, and it may be attributable to either partner or a combination of factors related to both partners. Infertility can be costly and time-consuming to treat, and success is not guaranteed or even probable in many cases. The costs to infertile individuals and couples can involve money, time, and physical and mental health.

Working women and men who are infertile want to keep their jobs, even if they require leave time or scheduling changes for fertility treatment. The spouses of infertile partners also may require workplace accommodations to participate in fertility treatment. Both individuals want to have health insurance coverage that provides reimbursement for costly fertility treatments. For these reasons it is important whether infertility is considered a disability under federal or state disability discrimination law.

The Americans with Disabilities Act (ADA) of 1990 (1), the Rehabilitation Act of 1973 (2), and many state laws prohibit discrimination against individuals with disabilities by employers and providers of services, which may include health insurance. It is essential to

determine whether a particular condition is a disability before applying the nondiscrimination and reasonable accommodation mandates of various statutes. It has yet to be determined whether infertility is considered a disability.

THE AMERICANS WITH DISABILITIES ACT

Major Statutory Provisions

Employment and Health Insurance Providers. Most employers are subject to one or more federal laws protecting individuals with disabilities from discrimination on the basis of their disabilities. The Americans with Disabilities Act of 1990 applies to private employers with 15 or more employees and to employees of state and local governmental agencies. The Rehabilitation Act of 1973 protects employees of federal agencies, most federal contractors, and recipients of federal financial assistance. Virtually all states have statutes covering public and private employers, although they vary in the number of employees necessary for an employer to be covered.

The application of disability discrimination law to health insurance providers is less clear. Title III of the Americans with Disabilities Act prohibits discrimination by 12 categories of private providers of programs and services to the public. While the weight of opinion is that Title III applies to health insurance providers, this has not been definitively decided. In addition, where health insurance is a benefit of employment, discriminatory treatment in an employer-provided health insurance program would be subject to the employment discrimination prohibitions.

Even if health insurance is covered by disability discrimination laws, insurance companies may be permitted to limit or exclude coverage for certain treatments in appropriate circumstances. The legality of such limitations and exclusions has yet to be clearly defined in the context of infertility treatment (3–6).

Nondiscrimination and Reasonable Accommodation. The major mandate of disability discrimination law is to prohibit discrimination against individuals with disabilities who are otherwise qualified. Lawmakers have recognized that most discrimination is not intentional, particularly in the context of disabilities. For that reason, facially neutral policies and practices that have a disparate impact on individuals with disabilities are subject to challenge as well, although not all will be found to be impermissible. For example, requiring that an employee have a drivers' license could have a disparate impact, and this might be challenged if driving is not an essential function of the position.

In addition to prohibiting discrimination, these laws also require employers to provide reasonable accommodations to known disabilities. Employers are not required to provide accommodation if it would constitute an undue hardship to do so. Undue hardship means significant difficulty or expense. Neither are employers required to lower performance standards or to make fundamental alterations to the program. Employees must be able to perform essential functions of the job if reasonable accommodations are provided, although the employer generally has the burden of showing that a particular function is essential.

Reasonable accommodations in the employment context might be the removal of architectural barriers, acquisition or modification of equipment, and other accommodations such as interpreters or readers. For an employee with fertility problems, the accommodations that might be sought would include job restructuring, part-time or modified work schedules, and reassignment to a vacant position.

Definition of Who Is Protected

Substantial Impairment, Regarded as, Record of. When most people think of disability discrimination laws, they think of individuals who are wheelchair users or those with visual or hearing impairments. Nevertheless, these laws cover a broad range of conditions and require not only nondiscrimination but reasonable accommodation. Whether infertility is to be included in statutory coverage is not clear on the face of the statutory language or the regulations.

Federal discrimination laws and many state discrimination laws define those to be protected similarly (7). Although some have urged that the definition be categorical and that specific impairments be listed to determine coverage, Congress specifically declined to do so. Instead individuals with disabilities are those who have a physical or mental impairment that substantially limits or more major life activities, those who have a record of such an impairment, or those who are regarded as having such an impairment.

The ADA regulations (1630.2 h) define a physical or mental impairment as "any physiological disorder, or condition, cosmetic disfigurement, or anatomical loss affecting one or more of [listed body systems]." These listed body systems include the reproductive system.

Major life activities are defined in the regulations as functions such as caring for oneself, performing manual tasks, walking, seeing, hearing, speaking, breathing, learning, and working. These listed activities are not intended to be all inclusive.

Substantially limited refers to being "unable to perform a major life activity that the average person in the general population can perform" or being "significantly restricted as to the condition, manner or duration under which an individual can perform a particular major life activity as compared to the condition, manner, or duration under which the average person in the general population can perform that same major life activity."

The requirement that the impairment be one that is substantially limiting is an important issue when considering infertility. This is because of the natural physiological changes that occur during the aging process which affect fertility, without an impact from disease, injury, or other condition that would affect fertility during the normal life cycle. This raises the question whether a woman who is in the average menopausal or postmenopausal age range would be considered disabled because she is no longer fertile or whether the definition

only covers women in their twenties or thirties who have substantial difficulty conceiving.

Statutory language and the regulations adopted pursuant to statutes are essential starting points for determining what is prohibited and what definitions apply in a particular policy context. Consideration is also generally given to interpretations provided by federal and state agencies charged with enforcement or implementation of statutes. In the case of employment discrimination, the primary agency is the Equal Employment Opportunity Commission (EEOC).

The EEOC was the agency that promulgated the regulations for the employment portion of the ADA, Title I. These regulations include an interpretive appendix. In addition the EEOC has issued a number of separate interpretive guidelines on various aspects of Title I. Not all of the EEOC interpretations have received complete acceptance by the courts and commentators.

The EEOC has published a memorandum providing guidance as to the definition of disability (8). In that memorandum the EEOC did not clarify specifically whether fertility is a disability. It has been argued, however, that the EEOC memorandum indicates an intent that procreation is a major life activity. In this regard, since infertility substantially limits the ability to procreate, infertility thus should be considered a disability under the ADA (9). EEOC does indicate that the ADA should be read broadly. The EEOC's discussion of human immunodeficiency virus (HIV) has been argued to support a determination that infertility is a disability. The EEOC has indicated that even someone with HIV who is asymptomatic would be covered because of the impact of the virus on procreation. Although not all courts initially accepted this interpretation (10), this interpretation was applied by the Supreme Court in 1998 in *Bragdon v. Abbott*, which is discussed later in the section on judicial interpretations (10).

The Department of Justice (DOJ), an agency with major implementation responsibility for the Rehabilitation Act, similarly supports protection for individuals with HIV, even those who are asymptomatic (11). The DOJ position is that a person with HIV cannot procreate without significant fears about the impact of the virus on the child. Again, this reasoning was adopted by the Supreme Court in *Bragdon*.

It has been argued that the DOJ logic on application to individuals with HIV should extend to infertility. The argument is:

> If an asymptomatic HIV-infected individual is protected under the Act because the potential to pass the virus onto a biological child constitutes a substantial limitation of the major life activity of procreation, then an infertile person, whose physical impairment substantially limits his or her ability to procreate in the first instance, likewise should be afforded the protection of the Act (12).

Courts have reached a wide range of conclusions about the coverage of various conditions. In cases involving sensory or mobility impairments, the decisions generally turn on the severity of the condition and the nature of the employment. Cases that are more problematic involve medical conditions, such as cancer, diabetes, obesity, and heart disease. Infertility is one of these problematic conditions.

Associational Disability. In addition to protecting individuals who are themselves impaired, the ADA (and arguably the Rehabilitation Act) also protects individuals from discrimination based on their association with someone with a disability. For example, it would be impermissible for an employer to refuse to hire someone because he or she had a child who is mentally retarded.

While the individual associated with someone with a disability is protected from discrimination, federal law does not require that reasonable accommodations be provided based on "associational disability." For example, while it would be impermissible to fire an employee because it was learned that the employee's daughter had suffered severe brain damage in an automobile accident, the employer is not required under the ADA or the Rehabilitation Act to provide an accommodation of allowing the employee time off to take the daughter for medical treatment or rehabilitation (14). The Family and Medical Leave Act (15), however, may provide relief to the employee in such a case, but nondiscrimination statutes will not.

This is significant with respect to infertility. Even if it were decided that infertility is a disability, accommodations would only be required for the partner with the medical condition, not for the other partner, whose presence may be necessary for certain infertility treatments.

OTHER APPLICABLE STATUTES

The importance of finding protection under disability discrimination statutes is highlighted when viewed in the context of other laws that might provide some protection for individuals and couples with fertility problems. As noted below, while these statutes are of some help, they do not provide the same level of substantive protection that would be available under the ADA or the Rehabilitation Act.

Pregnancy Discrimination Act of 1978

The Pregnancy Discrimination Act (PDA) of 1978 is a amendment to Title VII of the Civil Rights Act of 1964 (16). This statute prohibits employers from discrimination on the basis of "pregnancy, childbirth, or related medical conditions." The PDA does not require reasonable accommodation, so even if it were applied to an individual who is infertile, it is unlikely to be an avenue for the type of remedy being sought, namely accommodations in the work schedule and coverage of fertility treatment by an health insurance provider. It would only provide protection against an employer terminating employment or otherwise adversely treating an employee because of such a condition.

Several courts have found that infertility is a pregnancy-related condition under the PDA. One case involved an employee whose employment was allegedly

terminated because of her use of sick leave and vacation days to undergo fertility treatment. The court held that such action was subject to review under both the PDA (because infertility is a pregnancy-related condition) and the ADA (because infertility was determined to be a disability) (17).

Family and Medical Leave Act of 1993

While the PDA is unlikely to be a statutory basis for a leave of absence, the Family and Medical Leave Act (FMLA) of 1993 (18) does provide for such a leave in appropriate circumstances. The FMLA applies to employers with 50 or more employees, and it requires employers to provide up to 12 weeks of unpaid leave in a 12-month period of time. The leave is required only for the birth, adoption, or placement for foster care of a child; for care of a child, spouse, or parent with a *serious* health condition; or for the employee's own *serious* health condition that results in the employee's inability to perform the job. The term *serious health condition* is defined as one that involves inpatient care or continuing treatment by a health care provider. Neither the statute nor interpretations of the statute have discussed the potential applicability of the FMLA to infertility.

Health Insurance Portability and Accountability Act of 1996

The Health Insurance Portability and Accountability Act (HIPAA) of 1996, also known as Kennedy-Kassebaum (19), applies to employer-provided group health plans and group health insurance issued by private providers. HIPAA was intended to allow individuals to change group health insurance coverage without being unduly penalized. After initial eligibility with the first group health insurance plan, a transfer to subsequent plans should not adversely affect the individual. The covered employers and insurers may not deny coverage, or discriminate in eligibility, enrollment, or premium rates based on preexisting conditions. For the individual with infertility, the only benefit would be that if the individual is covered for fertility treatments by a health insurance plan subject to HIPAA, and the individual changes jobs, there would be no preexisting condition exclusion and no waiting period for coverage if fertility treatment is covered by the new employer's health insurance plan.

State Laws

Many states have statutes that are similar to PDA, FMLA, and the ADA. In general, state law interpretation often mirrors federal statutory applicability and interpretation, although there are a few notable exceptions. No state law clearly protects individuals with infertility problems.

JUDICIAL INTERPRETATION

The courts have addressed whether infertility is a disability in several cases. Two early cases have been subject to substantial commentary. Unfortunately, they have reached opposite conclusions, and there is not yet a definitive resolution of this issue as a result.

Both cases involved individuals who had been employed for some time. In *Pacourek v. Inland Steel Co., Inc.* (20) an employee with 10 years of service was dismissed because of her absences related to fertility treatments. The court determined that unexplained fertility is a physical impairment under the ADA. It further decided that reproduction is a major life activity, based on inference from EEOC interpretation and other judicial decisions, and that infertility is substantially limiting to this major life activity. Therefore infertility is a disability under the ADA.

The court relied on an earlier federal appellate court decision, *McWright v. Alexander* (21) in which the court had indicated that the Rehabilitation Act protected individuals with physiological disorders affecting the reproductive system. In *McWright*, the individual was seeking leave time to care for an adopted baby. Ms. McWright was unable to bear children as a result of childhood polio.

Zatarain v. WDSU-Television, Inc. (22) also involved a long-time employee. Ms. Zatarain was a television news anchor whose fertility treatments were initially accommodated. Eventually, however, her contract was not renewed after negotiations involving accommodations to her treatment. The court rejected her ADA claim, deciding that reproduction is not a major life activity. The court's reasoning was that other examples of major life activities in ADA regulations (e.g., walking, seeing, speaking, breathing, learning, and working) are done throughout the day, every day. Because one does not reproduce throughout the day, every day, this is not a major life activity. Neither is she substantially limited in working because she is not "significantly restricted in the ability to perform either a class of jobs or a broad range of jobs in various classes" because of her condition (23). The reasoning in *Zatarain* has been criticized as not being an appropriate interpretation of EEOC guidance (24).

It is noteworthy that these two cases addressing infertility as a disability, cases that result in conflicting definitions, had been decided at the time EEOC issued its 1995 interpretive guidance. The failure of the EEOC to specifically clarify its position on infertility is therefore troubling.

While *Pacourek* and *Zatarain* are the first major cases addressing infertility as a disability, a case decided after these cases is the first to reach a federal appellate court level. In *Krauel v. Iowa Methodist Medical Center* (25), the court considered the denial of health insurance coverage for a surgical procedure for a woman with endometriosis, a condition affecting her fertility. The court held that infertility is not a disability because it does not interfere with performing her job duties as a respiratory therapist, and that it would be a "considerable stretch of federal law" to treat it as a disability (25).

In one of the few cases involving health insurance coverage, rather than termination of employment, a court in the same jurisdiction as the *Pacourek* decision decided that a police officer with an ovarian dysfunction and infertility was protected as disabled under the ADA (26). The claim was that the employer had violated the ADA in denying health insurance coverage for fertility treatments.

The same court also determined that an employee with an incompetent cervix, which compromised her ability to carry a fetus to term, was protected as disabled under the ADA (27).

In 1998 the Supreme Court, in *Bragdon v. Abbott*, answered some of the questions about whether infertility should be treated as a disability under federal discrimination laws (28). The case involved a plaintiff who was a woman of child-bearing age who was HIV positive but asymptomatic. When she sought treatment from a dentist, he examined her in his office but indicated that he would only fill her cavity in a hospital because of her HIV status, and that she would have to bear the additional costs of hospital treatment. She brought suit under Title III of the ADA, claiming discrimination on the basis of disability by a private provider of a public accommodation. The Supreme Court addressed the issue of whether being HIV positive, but asymptomatic, is a disability under the ADA. The Court held that for this plaintiff it is.

The Court first determined that reproduction is a major life activity, by relying on the plain meaning of the statute and congressional intent to construe the statute to be consistent with regulations under the Rehabilitation Act. The second part of the test is whether one is "substantially limited" in that major life activity. Again, the Court held that Ms. Abbott is substantially limited because it affected her decision to conceive because of the significant risk to the partner as well as to the child, and this plaintiff had provided unchallenged testimony that her HIV infection controlled her decision not to have a child. So for Sidney Abbott, her HIV status was a substantial limitation to a major life activity.

Applying this analysis to infertility, it would seem clear that just as in the case of HIV, reproduction would be a major life activity for purposes of determining whether someone who is infertile is protected. The resolution of the question becomes more difficult, however, in applying the second part of the test, that is, whether for a particular individual, infertility is a substantial limitation to a major life activity. At first the answer might seem to be that it clearly is. It is not entirely clear, however, that all individuals who are infertile will be protected, just as the *Bragdon* Court did not hold that all individuals with HIV are automatically protected. Similarly it is not entirely clear that all individuals who are infertile will be automatically covered under the definition.

PROBABLE FUTURE DIRECTIONS

The initial split of opinion by the courts has been reinforced by other subsequent decisions. At the time of this writing, there have been several cases in which the court have determined that infertility is a disability. Most of these decisions, however, have been made by the same federal court in Illinois, so the number of decisions does not necessarily indicate the weight of authority.

The Supreme Court has, however, seemingly resolved the split to some degree in the *Bragdon* decision. The Court has at least decided that reproduction is a major life activity. This decision could certainly be extended to determine that infertility is a substantial impairment to a major life activity. What is unresolved is whether all individuals who are infertile will be covered.

As was previously noted, substantially limited means being "unable to perform a major life activity that the average person in the general population can perform" or being "significantly restricted as to the condition, manner or duration under which an individual can perform a major life activity as compared to the condition, manner, or duration under which the average person in the general population can perform that same major life activity." This may be interpreted to apply only to individuals who are in the normal age range for child bearing or fertility. This will thus be different for males and females.

Courts applying this analysis to cases involving infertility may examine whether the individual who is infertile is within the normal age range for fertility. Thus the 63-year-old may not be protected as disabled, while a 25-year-old woman would be. The fact that men are generally considered to be fertile under normal conditions throughout their adult lives raises some interesting disparate treatment issues. Would the 63-year-old man be considered disabled if he became impotent as a result of prostate cancer treatment?

Even if it is definitively decided that infertility is a disability under discrimination statutes, that is only the first step in receiving statutory protection. The individual must also be able to carry out the essential functions of the position with or without reasonable accommodations. The employer will generally have the burden of demonstrating what the essential functions of the job are and proving that accommodations such as schedule changes are unduly burdensome or fundamentally alter the program or lower standards. This will necessarily involve an individualized determination. And even if it is reasonable to accommodate an employee who is infertile, employers will not be required to make accommodations for employees where it is the partner with the infertility problem requiring the employee's presence for fertility treatment.

Finally, assuming that infertility is considered a disability and assuming that the accommodations sought are reasonable, there remain unresolved policy questions as to whether employers should be required to accommodate, and insurance companies to provide, health insurance coverage for an individual on an indefinite or undefined basis. Should the protections be extended to the 63-year-old woman who is seeking to conceive a child or to the woman who has already given birth to septuplets and seeks additional pregnancies? How will the application of disability discrimination law apply to individuals seeking access to health insurance coverage for drugs such as Viagra? These questions remain for the policy makers regardless of the direction taken by the courts and regulatory agencies in interpreting existing law.

BIBLIOGRAPHY

1. Americans with Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 327; codified at 42 U.S.C. §§12101 *et seq*. (1994).
2. Rehabilitation Act of 1973, 29 U.S.C. §§706, 791, 793–794 (1994).
3. D. Dallman, *William & Mary Law Rev.* **38**, 371–415 (1996).

4. B. Gilbert, *Defense Counsel J.* 42–57 (January 1996).
5. D. Millsap, *Houston Law Rev.* **32**, 1411–1450 (1996).
6. D. Millsap, *Am. J. Law Med.* **22**, 50–84 (1996).
7. Americans with Disabilities Act Regulations, 29 CFR Part 1630 (1997).
8. EEOC Compliance Manual (BNA) §902 (March 14, 1995).
9. S.M. Tomkowicz, *Syracuse Law Rev.* **46**, 1051–1092 (1996).
10. *Runnebaum v. NationsBank of Maryland*, 123 F.3d 156 (4th Cir. 1997).
11. Department of Justice, Rehabilitation Act Memorandum on the Application of Section 504 of the Rehabilitation Act to HIV-Infected Persons, from Douglas W. Kmiec (September 27, 1988).
12. S.M. Tomkowicz, *Syracuse Law Rev.* **46**, 1051–1092, 1071 (1996).
13. Interpretive Guidance to Title I of the Americans with Disabilities Act, 50 Fed. Reg. 35742 (July 26, 1991).
14. *Tyndall v. National Education Centers*, 31 F.3d 209 (4th Cir. 1994).
15. Family and Medical Leave Act, 29 U.S.C. §§2601–2654 (1994).
16. Pregnancy Discrimination Act of 1978, 42 U.S.C., §2000e(k).
17. *Landfair v. Sheahan*, 911 F. Supp. 323 (N.D. Ill. 1995).
18. Family and Medical Leave Act, 29 U.S.C. §§2601–2654 (1994).
19. Health Insurance Portability and Accountability Act of 1996, 42 U.S.C. §300gg (1997).
20. *Pacourek v. Inland Steel, Co., Inc.*, 916 F. Supp. 797 (N.D. Ill. 1996),
21. *McWright v. Alexander*, 982 F.2d 222 (7th Cir. 1992).
22. *Zatarain v. WDSU-Television, Inc.*, 881 F. Supp. 240 (E.D. La. 1995).
23. 29 CFR 1630.2(j)(3)(i).
24. S.M. Tomkowicz, *Syracuse Law Rev.* **46**, 1051–1092 (1996).
25. *Krauel v. Iowa Methodist Medical Center*, 95 F.3d 674 (8th Cir. 1996).
26. *Bielicki v. City of Chicago*, 10 NDLR ¶28 (N.D. Ill. 1997).
27. *Soodman v. Wildman, Harrold, Allen & Dixon*, 9 NDLR ¶200 (N.D. Ill. 1997).
28. *Bragdon v. Abbott*, 1998 U.S. LEXIS 4212 (1998).

See other entries DISABILITY AND BIOTECHNOLOGY; REPRODUCTION, ETHICS, PRENATAL TESTING, AND THE DISABILITY RIGHTS CRITIQUE; REPRODUCTION, LAW, REGULATION OF REPRODUCTIVE TECHNOLOGIES; REPRODUCTION, LAW, WRONGFUL BIRTH, AND WRONGFUL LIFE ACTIONS.

REPRODUCTION, LAW, REGULATION OF REPRODUCTIVE TECHNOLOGIES

KAYHAN P. PARSI
Institute for Ethics, American Medical Association
Chicago, Illinois

OUTLINE

INTRODUCTION

Almost as fascinating as the reproductive technologies are the cultural responses to the new methods of controlling, enhancing, or limiting individuals' abilities to procreate. Writers have often invoked Aldous Huxley's *Brave New World* as a predictor of things to come, as if a single novel can faithfully capture the complexity and richness of reproductive technology. In reality, reproductive technologies vary a great deal, from the low tech, such as surrogacy and artificial insemination, to more highly sophisticated methods, such as intracytoplasmic sperm injection (ICSI). Regulation of these technologies by professional associations, states, and countries also varies considerably. While these reproductive technologies have created interesting and complicated abstract ethical issues, this article will only examine the regulatory responses to such technologies. Although regulation is often associated with the prohibition of specific activities, it is also enabling by allowing key parties to execute particular agreements in the area of reproductive technologies. This enabling aspect of regulation is in accordance with a liberal notion of procreative rights.

Commentators such as John Robertson generally support an individual's ability to choose a procreation method. Robertson believes that Americans have a procreative liberty interest. The very title of his book, *Children of Choice*, reflects a very American attitude: Choice is an inherently good thing, and this applies to reproductive technologies as well (1). Robertson's perspective epitomizes the viewpoint that reproductive technologies enhance and expand an individual's ability to make choices regarding procreation. Robertson acknowledges that although the technologies may produce ambivalence about their use, limiting persons' reproductive freedoms would curtail one of the most fundamental aspects of our lives. Other commentators, such as Dorothy Roberts, observe a darker side to certain kinds of reproductive technologies. Instead of focusing on the standard set of reproductive technologies, such as in vitro fertilization (IVF) or cryopreservation, Roberts concentrates on technologies that limit the ability to procreate (2). Roberts's concerns focus on utilization of drugs, such as Depo-Provera and Norplant, and sterilization of poor people. She is critical of the general acceptance of reproductive technologies to

enhance wealthy people's ability to procreate, while governments employ the aforementioned methods to curtail poor — especially black — women's ability to procreate. As exemplified by Robertson's and Roberts's viewpoints, goals of reproductive technology are quite diverse.

U.S. CASE LAW

In the United States, a number of cases provide a basis for Robertson's notion of procreative liberty. One of the earliest cases to assert a constitutional right to procreate was *Skinner v. Oklahoma* (3). This case overturned the now infamous sterilization case *Buck v. Bell* (4). In *Skinner*, the Supreme Court declared unconstitutional the Oklahoma Habitual Criminal Sterilization Act, which allowed for the sterilization of habitual criminals (2). The Court's language accorded reproductive freedom a high level of deference: "Oklahoma deprives certain individuals of a right which is basic to the perpetuation of a race-the right to have offspring ... we are dealing here with legislation which involves one of the basic civil rights of man." *Griswold v. Connecticut* (6) was another landmark case where the Supreme Court further expanded the notion of procreative liberty. The Court struck down a Connecticut statute that prohibited the distribution of contraceptives to married couples. This statute, the Court held, violated married couples' constitutional privacy rights. Justice Douglas wrote that "the First Amendment has a penumbra where privacy is protected from governmental intrusion. ...We deal with a right of privacy older than the [Bill of Rights]. Marriage ... is an association for as noble a purpose as any involved in our prior decisions" (5). In *Eisenstadt v. Baird* (6), the Court expanded the privacy right beyond *Griswold's* realm of marriage. In his opinion, Justice Brennan wrote:

> If the right of privacy means anything, it is the right of the individual, married or single, to be free from unwarranted governmental intrusion into matters so fundamentally affecting a person as the decision whether to bear or beget a child (7).

The Court again addressed the issue of contraception in *Carey v. Population Services International* (7), by striking down a New York statute that criminalized distribution of contraceptives to minors under 16, prohibited nonpharmacists from distributing contraceptives to people over 16, and banned any advertising or displaying of contraceptives. The *Carey* Court held that:

> [I]t is clear that among the decisions that an individual may make without unjustified government interference are personal decisions relating to marriage...; procreation...; family relationships...; and child rearing and education. The decision whether or not to beget or bear a child is at the very heart of the cluster of constitutionally protected choices (7).

After Carey, procreative liberty rights continued to grow. The now-famous Supreme Court case of *Roe v. Wade* (8) expanded privacy rights to permit women in collaboration with their physician to have an abortion. In *Planned Parenthood v. Casey* (9), the Supreme Court, based on its "liberty" rights reasoning, argued that abortion is an essential liberty. The liberty argument in *Planned Parenthood* focused more on control of one's own body, as opposed to limiting the government's ability to control reproduction, which was the language used in the aforementioned privacy cases (10).

The foregoing cases strongly suggest a basic negative right to procreative decision making and influence a variety of reproductive technologies — such as surrogacy, IVF, artificial insemination, and contraception. The remainder of this article will explain how different entities regulate various reproductive technologies; however, the discussion does not evaluate an exhaustive list of all possible reproductive technologies. First, a description of the guidelines promulgated by certain professional groups, such as the American Medical Association's (AMA) Council on Ethical and Judicial Affairs (CEJA) and the American Society of Reproductive Medicine (ASRM), will be given. The article will then examine state responses in the form of case law and legislation. It will then look at different approaches by countries and then finally different approaches by international bodies (e.g., the Council of Europe) to these technologies. A variety of approaches are represented, from the U.S. free-market approach to the more heavily regulated and centralized UK approach.

PROFESSIONAL GROUP GUIDELINES

Professional guidelines are a useful place to start when examining regulatory aspects of reproductive technologies. Although they do not have the force of law, they often do inform legal cases. Moreover, in the absence of any national legal or ethical consensus regarding these technologies in the United States, professional guidelines provide physicians, researchers, and ethicists with some guidance. The AMA's Code of Medical Ethics addresses the following reproductive technology issues: artificial insemination, IVF, freezing pre-embryos, pre-embryo splitting, and surrogacy.

Artificial Insemination

The AMA Code (11) makes the following requirements for recipients of artificial insemination: counseling, informed consent (e.g., risks, benefits, and alternative treatments), and information regarding the conceived child's legal status. The Code stipulates that sex selection is only allowed to avoid an inheritable sex-linked disease. Posthumous use of frozen sperm, according to the instructions of the decedent, is allowed (see the later California case *Hecht*). The Code also requires rigorous screening of potential donors for infectious or inheritable diseases, recommends the use of frozen semen (to ensure freedom of HIV infection), and advises physicians to use the professional guidelines set out by the ASRM, the Centers for Disease Control and Prevention (CDC), and the Food and Drug Administration (FDA). Physicians are also required to maintain permanent records of the sperm donors, reflecting health and genetic information that is both identifying and nonidentifying. The Code recommends obtaining the consent of the husband if he

will become the father through artificial insemination by anonymous donor. Unlike certain European guidelines, the Code does not prohibit single or lesbian women from obtaining artificial insemination by anonymous donor. Last, the Code admonishes compensating donors beyond incidental expenses such as time.

Regarding IVF, the Code prohibits using fertilized ova that will later be implanted to be used for laboratory research. Those fertilized ova that will not be implanted may be used for research purposes, but only in accordance with the Code's fetal research guidelines. The following guidelines are offered as aids to physicians when they are engaged in fetal research:

1. Physicians may participate in fetal research when their activities are part of a competently designed program, under accepted standards of scientific research, to produce data that are scientifically valid and significant.
2. If appropriate, properly performed clinical studies on animals and nongravid humans should precede any particular fetal research project.
3. In fetal research projects, the investigator should demonstrate the same care and concern for the fetus as a physician providing fetal care or treatment in a nonresearch setting.
4. All valid federal or state legal requirements should be followed.
5. There should be no monetary payment to obtain any fetal material for fetal research projects.
6. Competent peer review committees, review boards, or advisory boards should be available, when appropriate, to protect against the possible abuses that could arise in such research.
7. Research on "dead fetus," macerated fetal material, fetal cells, fetal tissue, or fetal organs should be in accord with state laws on autopsy and state laws on organ transplantation or anatomical gifts.
8. In fetal research primarily for treatment of the fetus:
   a. Voluntary and informed consent, in writing, should be given by the gravid woman, acting in the best interest of the fetus.
   b. Alternative treatment or methods of care, if any, should be carefully evaluated and fully explained. If simpler and safer treatment is available, it should be pursued.
9. In research primarily for treatment of the gravid female:
   a. Voluntary and informed consent, in writing, should be given by the patient.
   b. Alternative treatment or methods of care should be carefully evaluated and fully explained to the patient. If simpler and safer treatment is available, it should be pursued.
   c. If possible, the risk to the fetus should be the least possible, consistent with the gravid female's need for treatment.
10. In fetal research involving a fetus in utero, primarily for the accumulation of scientific knowledge:
    a. Voluntary and informed consent, in writing, should be given by the gravid woman under circumstances in which a prudent and informed adult would reasonably be expected to give such consent.
    b. The risk to the fetus imposed by the research should be the least possible.
    c. The purpose of research is the production of data and knowledge that are scientifically significant and that cannot otherwise be obtained.
    d. In this area of research, it is especially important to emphasize that care and concern for the fetus should be demonstrated (12).

The Code makes the gamete providers the primary decision makers when exerting control over a frozen pre-embryo. The providers are prohibited from selling their gametes, but they are allowed to donate them to others. Research is also prohibited on a pre-embryo if it will later be implanted in a woman. Interestingly, however, the Code allows pre-embryos to thaw and deteriorate. With regard to use of pre-embryos, the Code requires consent of both providers and encourages the use of agreements between providers in case the couple divorces. The Code permits pre-embryo splitting with the agreement of both gamete providers. Pre-embryo splitting allows a greater chance for conception while diminishing the number of potentially painful procedures to procure eggs (13).

Surrogacy

Surrogacy is one of the oldest forms of reproductive technologies. It is referred to in the Bible when Abraham's servant Hagar bears Abraham's child to be raised by him and his wife Sarah (14). It is the focus of such recent dystopian novels as *The Handmaid's Tale* by Margaret Atwood. And it has raised a number of ethical and legal concerns in recent years. Accordingly the AMA Code cites certain concerns regarding surrogacy agreements: commodification of children, exploitation of poor women, and interference with the natural maternal–child bond. It also raises potential psychological problems and the possibility that the mother may want to have an abortion or even refuse giving up the child. The intended parents may not even want the child if it happens to be born with disabilities. Despite major criticisms of surrogacy contracts, the Code permits them with certain safeguards. For instance, the birth mother should have:

> the right to void the contract within a reasonable period of time after the birth of the child. If the contract is voided, custody of the child should be determined according to the child's best interests. In gestational surrogacy, in which the surrogate mother has no genetic tie to the fetus, the justification for allowing the surrogate mother to void the contract becomes less clear. Gestational surrogacy contracts should be strictly enforceable (i.e., not voidable by either party) (15).

Cryopreservation

ASRM outlines a number of concerns with regard to cryopreservation. For instance, it encourage couples who are considering storage to put into writing what they wish to happen to their stored embryos in the following instances: death, divorce, separation, failure to pay storage charges, inability to agree on disposition in the future, or lack of contact with the program. ASRM requires that the consent form allow disposal of embryos if the couple lose contact with the program after some period of time and if they do not provide the program with key contact information, such as addresses and phone numbers. The ASRM guidelines allow a couple to revise their initial directions for embryo disposition by drafting a new set of written directions. Moreover, ASRM does not make rigid requirements with regards to embryos that lack written directions for disposal. In the absence of clear legal guidelines, programs may want to store indefinitely or accept the risk of liability by disposing of embryos after attempts to contact the couple have failed and a lengthy period of time has elapsed. ASRM approaches the issue of embryo preservation and storage pragmatically, in that a couple cannot claim an injury if they have failed to provide written directions, have lost contact with the program, and have not provided current address and phone information. Simply put, ASRM does not believe that programs have the ethical obligation to store embryos indefinitely. ASRM does state, however, that a period of five years should elapse and that "diligent effort" by phone and registered mail to contact the couple should be made by the program. Last, ASRM does not allow abandoned embryos to be used for research or donation to another couple without appropriate consent. The embryos must be thawed and allowed to deteriorate (16).

Embryo Splitting

Neither AMA nor ASRM prohibits embryo splitting (a technique whereby in vitro fertilized pre-embryos are split to create genetically identical siblings). Both organizations argue that splitting provides certain benefits: splitting provides a greater number of embryos, increasing the chances of a successful pregnancy and may prevent additional invasive procedures to retrieve more embryos, which are painful and costly. Neither AMA nor ASRM finds persuasive ethical concerns about using split embryos as a source of organs or tissues for an existing child, or the sale of stored embryos with desirable genomes based on the appearance of characteristics of existing children. For instance, the AMA Code recommends a complete prohibition of the sale of pre-embryos. Both the Code and ASRM acknowledge that offspring with identical genomes may be born at different times. ASRM raises the issue of "personal identity and the meaning of being a twin that require further investigation before it can be determined that such transfers are ethical." It recommends that such scenarios can be avoided "by transferring all genetically identical embryos in the same cycle." The Code permits couples to transplant genetically identical siblings in order to harvest their tissue for a needy sibling. The Code discounts any accrual of psychological harm and argues that the sibling may indeed gain psychological benefits by saving his or her sibling. In sum, both the ASRM and the Code take a consequentialist approach in that they believe that the benefits outweigh any costs in allowing embryo splitting (13,18).

Use of Fetal Oocytes in Assisted Reproduction

ASRM sees a number of problems with fetal oocyte donation: emotional harms to children, informed consent dilemmas, and an impersonalizing influence in assisted reproduction. Because of these concerns ASRM does not believe this technology should be pursued (18).

Posthumous Reproduction

Similar to surrogacy, posthumous births have an ancient lineage. These kinds of births routinely occurred when a woman conceived and her husband or partner died before she delivered the child. This child was commonly considered the legal heir of the father. ASRM, however, recognizes that posthumous reproduction became an issue when semen could be frozen and later implanted after the death of the donor. Although the AMA Code only makes a fleeting commentary regarding posthumous reproduction, ASRM has extensive commentary. ASRM cites a variety of scenarios where such reproduction may take place: A widow may retrieve a dead husband's sperm to procreate, sperm may be retrieved from terminally ill spouses or partners (employing techniques such as stimulated ejaculation, microsurgical epididymal sperm aspiration, MESA, or testicular sperm extraction, TSE), or sperm from a dead anonymous donor may be used. Moreover, a man facing radiation therapy or chemotherapy may want to store his semen for later (possibly posthumous) use.

Cryopreserving ova for posthumous procreation also poses certain concerns, but the inability to successfully freeze ova (as compared to semen and pre-embryos) has limited this particular technology. ASRM permits the designation of frozen gametes or embryos to be used in posthumous procreation as long as the key parties involved are fully informed. However, the absence of clear and written instructions would preclude posthumous reproduction. Moreover, ASRM believes that the requests of a living spouse should not override the express wishes of the deceased spouse.

ASRM also permits a husband to use his deceased wife's ova for implantation in a surrogate. Although the surrogate would not be considered a traditional surrogate, the ASRM requires that the surrogate be made aware of the circumstances and informed that she would be involved in a posthumous pregnancy. The rearing parents may lack genetic ties to the dead donor and may not be involved in the gestating pregnancy. They should, however, be made aware of the deceased status of the donors of gametes and embryos. Although the ASRM guidelines are supportive of posthumous reproduction, they do cite some reservations:

> . . .when reproduction takes place as a consequence of a loving relationship in which both partners were desirous of children, but a pregnancy is frustrated by the death of one partner, posthumous reproduction would ordinarily be well accepted both socially and culturally. . .. There is less certainty of the

impact on the child and more caution should be exercised for posthumous reproduction that occurs with the use of donated gametes from unrelated individuals who are not living and may have been deceased for several years, as may occur with the use of commercial banks as a source for sperm, frozen embryos, or ova (19).

STATE LAW AND REGULATIONS

Reflecting its organization, American state law that regulates reproductive technologies is a hodgepodge of case law, statutes, and administrative regulations. Certain issues, however, have been addressed extensively. One of them is surrogacy. One of the early surrogacy cases was *In The Matter of Baby M* (20). In this case, William Stern entered into a surrogacy contract with Mary Beth Whitehead. Elizabeth, Stern's wife, was infertile, and the Sterns were hoping to be able to have a baby with the assistance of Whitehead. The contract was for the sum of $10,000. In early 1985, Mr. Stern and the Whiteheads executed a surrogacy agreement. To avoid revealing the nature of the agreement, the child's birth certificate listed the Whiteheads as parents. It became clear very quickly that Mrs. Whitehead did not wish to part with her baby. Despite initial misgivings, Mrs. Whitehead gave the baby up to the Sterns. Whitehead underwent a depression and threatened suicide. The Sterns, therefore, returned the baby to Whitehead for a short visit. Thereupon, Whitehead left with the baby to Florida for four months. The Sterns's complaint, in addition to seeking possession and ultimately custody of the child, sought enforcement of the surrogacy contract. Pursuant to the contract, the Stern's asked that the child be permanently placed in their custody, that Mrs. Whitehead's parental rights be terminated, and that Mrs. Stern be allowed to adopt the child.

The Supreme Court of New Jersey invalidated the agreement and named Whitehead the mother of the child. The judge in this case noted certain ethical problems, such as commodification, with regard to surrogacy.

Because of commodification concerns, a number of states do prohibit enforcement of surrogacy agreements. For instance, Arizona, the District of Columbia, Michigan, and Utah prohibit all surrogacy agreements, whereas Kentucky and Louisiana prohibit commercial surrogacy agreements. Some states permit unpaid surrogacy agreements: Florida, Nevada, New Hampshire and Virginia. Florida, New Hampshire and Virginia require that the intended mother be infertile (21).

Another important surrogacy case was *Johnson v. Calvert* (22). In this case Mark and Crispina Calvert entered into a contract with Anna Johnson for Johnson to carry the Calverts's fertilized embryo (Crispina Calvert had underwent a hysterectomy but was able to produce ova). The Calverts agreed to pay Johnson $10,000; in return, Johnson would relinquish her parental rights to the child in favor of the Calverts. Near the end of the pregnancy, Johnson demanded immediate payment or she would not deliver the child after birth to the Calverts. The Johnsons sued, seeking a declaration that they were the legal parents of the child. Calvert countersued. The case wound its way through the California judicial system, until the Supreme Court of California heard it in 1993. The California court took a very different approach than the New Jersey court in *Baby M*. Here, the court was not concerned about potential exploitation or commodification. Moreover it found such concerns to be paternalistic and condescending toward women. The court held that the woman who intended the birth of a child that she intended to rear was the natural mother under California law.

Another issue that has been addressed by state case law is embryo storage and disposal. Perhaps the most famous case involving frozen embryos is the Tennessee case *Davis v. Davis* (23). In this case, a couple undergoing a divorce were arguing over the disposition of their frozen pre-embryos. The Davises never executed a disposition agreement regarding their embryos. Mary Sue Davis wished to achieve a pregnancy after their divorce, whereas Junior Davis wanted to avoid becoming a father entirely. In the absence of any statutory authority, the Tennessee Supreme Court stated that the pre-embryos occupy an intermediate status between property and persons. The court ruled that the fate of the frozen pre-embryos should be decided by "the party wishing to avoid procreation" if the other party has a reasonable possibility of achieving parenthood by other means and the parties have not made an agreement regarding their disposition. A later New York case, *Kass v. Kass* recognized the legal enforceability of a disposition agreement regarding frozen embryos. The court ruled that the couples' agreement controlled the fate of their embryos (24). Last, a couple of state statutes provide some guidance: Florida law requires that couples execute disposition agreements when undergoing a reproductive technology procedure (25) and Louisiana law defines the embryo as a "juridical person," limiting the ability of progenitors to dispose of their embryos (26).

Finally a case that raised the issue of posthumous use of sperm is *Hecht v. Superior Court of LA County* (27). In this case William Kane "willed" a vial of his own sperm to his girlfriend Deborah Hecht. He thereupon took his own life. Despite the fact that Kane's intent was clear, his ex-wife and children challenged his bequeathal. The probate judge initially ordered that the sperm be divided according to the original property settlement. The Court of Appeals later ruled that the remaining vials of sperm, which were in the custody of the administrator, be delivered to Hecht. The court argued that Kane had a legitimate property interest in his own sperm. Analogous to the Davis court, the Hecht court argued that the sperm occupied an interim category of property because of its potential for life. The court's reasoning in this regard was not dissimilar from the policy guidelines of ASRM.

FEDERAL GUIDELINES

Although federal regulation of assisted reproductive technologies is very weak in the United States, Congress has made some attempts to impose some regulation in this area. For instance, in the early 1990s, Congress passed the Fertility Clinic Success Rate and Certification Act of 1992: A Model Program for the Certification of Embryo Laboratories (28). This act required the Secretary

of the Department of Health and Human Services, through CDC, to develop a model program for the certification of embryo laboratories. This program was to be carried out voluntarily by interested states (28).

INTERNATIONAL APPROACHES

In addition to the United States, other countries, such as Canada, the UK and Australia, have all addressed the legal and regulatory aspects of reproductive technologies. Despite Canada's tradition of strong social solidarity reflected in its national health insurance system, no federal guidelines have been created to govern the practice of IVF, artificial insemination, as well as egg and sperm donation. Similarly, although no specific law prohibits surrogacy contracts, they would not stand up in court because they violate Canadian contract and family law principles. Recommendations do exist, however, for storage of gametes and embryos — 10 years and 5 years, respectively (29). The Law Reform Commission's 1992 Working Paper recommends that "[t]he commercialization of donated gametes and embryos must be prohibited outright. Allowing gametes and embryos in the consumer market would constitute a direct assault on human dignity" (30). After many years and millions of dollars, Canada's Royal Commission on New Reproductive Technologies issued its voluminous report *Proceed with Care*. The report's recommendations were essentially prohibitive in nature. In the aftermath of the Commission's issuance of *Proceed with Care*, a liberal government bill, C-47 (The Reproductive and Genetic Technologies Act), was on its way to passage in 1996 but died because a federal election was called (31). This bill would have prohibited the following practices and procedures: commercially exchanging sperm and eggs, cloning, fusing animal and human zygotes, implanting a human embryo into an animal (and vice versa), altering the genetic structure of the germ line, retrieving an ovum or sperm from a cadaver with the intention of using it in a live recipient, maintaining a human embryo outside of a human body, fertilizing an ovum for purposes of research, and commercial surrogacy (32).

In England, the focus shifted from the status of the fetus to the status of the embryo in the aftermath of the birth of Louise Brown, the first "test tube" baby. Public alarm about the untrammeled growth of reproductive technologies motivated Parliament to explore legislative measures (33). The Committee of Inquiry into Human Fertilization and Embryology (the Warnock Committee) took its charge into examining these issues seriously. The Warnock Committee's work in the reproductive health arena left an indelible stamp. The Committee recognized a special status for the embryo but permitted embryonic research up to the fourteenth day after fertilization. The committee permitted research on excess embryos, whether or not the embryos were intentionally developed for research. Britain enacted legislation concerning the reproductive technologies in 1990 with the Human Fertilization and Embryology Act 1990. The Act defines an embryo as "a live human embryo where fertilisation is complete." Moreover, the Act states that "fertilisation is not complete until the appearance of a two cell zygote." The Act also outlines a number of prohibited practices, as outlined below:

> (1) No person shall —
> (a) bring about the creation of an embryo, or
> (b) keep or use an embryo, except in pursuance of a licence.
> (2) No person shall place in a woman —
> (a) a live embryo other than a human embryo, or
> (b) any live gametes other than human gametes.
> (3) A licence cannot authorise —
> (a) keeping or using an embryo after appearance of the primitive streak,
> (b) placing an embryo in any animal,
> (c) keeping or using an embryo in any circumstances in which regulations prohibit its keeping or use, or
> (d) replacing a nucleus of a cell of an embryo with a nucleus taken from a cell of any person, embryo or subsequent development of an embryo.
> (4) For the purposes of subsection (3)(a) above, the primitive streak is to be taken to have appeared in an embryo not later than the end of the period of 14 days beginning with the day when the gametes are mixed, not counting any time during which the embryo is stored (34).

The Act addressed four of the treatments available: artificial insemination using donated gametes, egg donation, embryo donation, and IVF. The act also contained explicit statutory regulations of embryo research, which is permitted until the appearance of the primitive streak ("taken to have appeared in the embryo not later than the end of the period of 14 days beginning with the day when the gametes are mixed") (35). The Act prohibits the creation of hybrids using human gametes, the cloning of embryos by nucleus substitution to produce genetically identical individuals, and genetic engineering to change the structure of an embryo. Despite these regulations commentators have called the Act a "'radical laissez faire' approach; a system of regulated private ordering" (36). Hence sex selection is not explicitly prohibited, as evidenced by an Essex woman who selected her child's sex in 1994 at the London Gender Clinic (36). Moreover, the National Health Service announced in 1994 that it would offer fertility treatment to lesbian couples, a departure from the heterosexual requirements seen in other countries (37).

The Human Fertilisation and Embryology Act of 1990 also amended the Abortion Act of 1967 by reducing the limit for legal abortion to 24 weeks. One commentator noted that the 1988 Alton bill sought to reduce the legal limit for an abortion even further, to 18 weeks (38). Despite the fact that English abortion law was codified in the Abortion Act of 1967, judges were still left with some amount of discretion in their rulings. Moreover certain judges would interpret English legal tradition as affording the fetus a fairly high level of protection. Judge Denning in *Royal College of Nursing* compared the fetus in utero to a child and argued that English law recognized a criminal cause of action if a fetus was intentionally killed. These mixed views toward the status of the fetus, however, did not undermine the United Kingdom's effort in passing the

Human Fertilisation and Embryology Authority. Among the countries surveyed here, the Authority is a rare instance where the government has enough consensus to regulate reproductive technologies on a national level.

In Australia, the Rios case in the early 1980s triggered an interest in the status of embryos ex utero. In this case the Rioses had participated in IVF procedures in Melbourne. They produced three embryos from Ms. Rios's eggs and the sperm of an anonymous donor. One was implanted and the other two were frozen for storage. The Rioses both died in a plane crash, prompting a son from Ms. Rios's previous marriage to declare his share of his stepfather's estate. The embryos were left frozen, following the Australian Waller Committee's declaration that the embryos had no status of their own to be independently unfrozen and implanted in another surrogate mother.

The first official Australian pronouncements on the legality of IVF were issued by the National Health and Medical Research Council in 1982. The Council's guidelines permitted IVF, but with certain constraints: The recipient had to be in an "accepted family relationship," embryos could not be kept beyond the normal implantation time schedule, and cryopreservation could be maintained no longer than usual reproductive need. Cloning was strictly prohibited (39). Each state has modified the Council's rules. For instance, in South Australia and Western Australia, artificial insemination procedures are available to married couples and to couples in de facto relationships of a certain length. In Victoria, access to artificial fertilization procedures is generally limited to married couples. In Western Australia, only married couples or heterosexual de facto couples who have lived together for a total of five years out of the previous six are eligible to be treated with IVF (40).

As far as storage issues are concerned, South Australia prohibits the storage of embryos for more than 10 years (40). In Victoria, there are no statutory time limits on storage, but it is an offense to freeze an embryo unless it is done with the intention of subsequently implanting it in a woman's womb. In Western Australia, storage of reproductive material is prohibited unless undertaken in accordance with a license or exemption. Ova that are being fertilized or embryos must not be stored unless the primary intention of the storage is their "probable future implantation." In any event, they must not be stored for more than three years (40).

In South Australia, Victoria, and Western Australia the rights of control and disposal of gametes and embryos are set down in legislation. In South Australia, the Reproductive Technology Act 1988 (SA) provides that the code of ethical practice will make provision for decisions to be made for the use or disposal of stored embryos (41). Such decisions must be able to be reviewed at least every 12 months. The maximum period of storage for an embryo is 10 years.

In Victoria, under the Infertility (Medical Procedures) Act 1984 (Vic), if an embryo cannot be implanted in a woman's womb due to the woman's death or injury, the embryo may be given to another woman for use in a procedure permitted under the Act in accordance with the consent of the gamete donors. If the donors have died or cannot be located, the Minister will direct the hospital to make the embryo available for use in a procedure permitted under the Act (42). In Western Australia, under the Human Reproductive Technology Act 1991 (WA), the providers of gametes have all rights of use or disposition of the gametes as if they were personal property until the gametes have been used, although the gametes may not be sold. If gametes are donated to a licensee, rights of control and disposition vest in the licensee who, subject to the consent of the donor, may only use the gametes for in vitro fertilization for a person named or chosen by means specified in the consent, for artificial insemination, for approved research, or for diagnostic procedures. If the gametes are not used for one of these procedures, they must be permitted to succumb subject to the rights of control that may be vested in the recipient couple. Where a donor gives conditional consent for use of gametes, provided that the gametes have not been used, the rights in relation to the gametes revert to the donor if the condition is violated. In Western Australia, consent must be given by a person before his or her gametes or any fertilizing ovum or embryo are used or stored and such use or storage must be in accordance with the consent (43). In Victoria, consent must be given by persons who donate sperm or ova or embryos. Victoria permits embryo experimentation up to 14 days after fertilization.

In South Australia, research using human reproductive material will be governed by the code of ethical practice to be formulated by the South Australian Council on Reproductive Technology. The Code will prohibit development of a human embryo outside the body "beyond the stage of development at which implantation would normally occur." Research using human reproductive material may be undertaken in accordance with a license (40).

In anticipation of such cases as the recent *Kass* case in New York, Australian lawyers at a meeting of the executive of the Law Council's family law section in 1992 proposed a novel rule. Their proposal would require all couples who participate in assisted reproductive technology (ART) to execute a disposition agreement for the future fate of their frozen embryos (44). As Michael Watt, a Melbourne barrister who represented the mother in Australia's first dispute over frozen embryos said: "Trying to squeeze it into custody raises the definition of when does life begin. If it's custody, you have to decide whether an embryo is a child. If it is a human life, it isn't property.... At the moment, every country simply says parties have joint property in the embryos, or an equal say in the future disposal and general responsibility for their future disposal, and there is no deadlock-breaking provision in any legislation. ...It avoids having to have definitions on when life begins" (44).

In early 1995 the Victorian state cabinet approved laws permitting noncommercial surrogacy. The revised laws would make commercial surrogacy a crime (45). In the Australian Capital Territory, the Substitute Parent Agreements Act 1994 prohibits commercial surrogacy. Queensland prohibits commercial surrogacy and the publication of advertisements for surrogacy services. In South Australia, surrogacy contracts and procuration contracts are illegal and void. In Tasmania, the Surrogacy Contracts

Act 1993 (Tas) makes it an offense to introduce potential parties to a surrogacy contract, arrange or negotiate a surrogacy contract, or to give or receive valuable consideration in connection with a surrogacy contract (40). In Victoria, commercial surrogacy is prohibited, and it is an offence to publish or advertise surrogacy services. Even in states without specific legislative provisions governing surrogacy, it appears likely that surrogacy contracts would be unenforceable on the grounds of public policy.

In 1996 the Supreme Court of Tasmania heard the case of *In the Matter of Estate of the Late K* (46). Here the issue was whether the product of the ova of a widow and the semen of her deceased husband are children of the deceased upon being born alive. The deceased died intestate, leaving behind three children from a previous marriage. Justice Slicer in his opinion looked at cases from other common law countries (*Paton v. Trustees of the British Pregnancy Advisory Service, Roe v. Wade, R v. Morgentaler*), as well as in Australia. He determined the following:

- A foetus is not recognized, by the law, as a person in the full legal sense.
- The law has long recognised foetal rights contingent upon a legal personality being acquired upon its subsequent birth alive.
- A child, *en ventre sa mere*, is not a human being. To be human a child must have quitted its mother in a living state.
- A child so born is by a legal fiction treated as having been living at an earlier point of time and as if by being so treated the child would receive a benefit to which it would have been entitled if actually born at that earlier time (46).

Slicer concluded that at the time of the decedent's death there were no human offspring in existence. He asked whether the law should distinguish between a child *en ventre sa mere* and a sibling who was a frozen embryo? He stated that the New South Wales Law Reform Commission 1988 recommended that "children conceived posthumously as a result of IVF procedures and children born from stored embryos should be able to make a claim against the estates of their genetic parents under the Family Provisions Act 1982." Slicer ultimately concluded "[t]hat a child, being the product of his father's semen and mother's ovum, implanted in the mother's womb subsequent to the death of his father is, upon birth, entitled to a right of inheritance afforded by law."

Australia reflects the Anglo-American tradition by respecting certain individual rights (e.g., abortion). Yet Australian law does take an activist approach toward issues such as IVF, surrogacy, and embryo experimentation. Although the rules are primarily procedural, certain substantive values are being promoted. For instance, family stability as a societal goal is embedded in the rules concerning artificial reproduction. A definite preference is given to heterosexual unions or de facto marriages that have existed for a number of years. The law, then, gives legal sanction to a cultural norm. Moreover, the 1989 report concerning the status of the embryo clearly compares its status to that of a living person. Such a statement in the United States could easily be in the position paper of a "pro-life" organization.

Compared to the other common law countries surveyed, Australian courts have not had to grapple with issues of fetal status nearly as frequently. Yet Australia has been at the forefront in creating innovative legislation with regard to assisted reproductive technologies. Moreover the myriad rules that have been created suggest that embryonic and fetal life is accorded some respect in Australia.

The Council of Europe, a multinational body dedicated to human rights and the rule of law among its member states, has issued its own regulations regarding reproductive technologies. Two articles in the Convention on Human Rights and Biomedicine prohibits two kinds of activities. Article 14 prohibits the use of reproductive technologies to choose a child's sex. The Article makes an exception, however, to avoid a serious hereditary-linked disease. Article 18 of the Convention prohibits the creation of embryos for research purposes. Moreover, although it allows research on embryos in vitro, the Article states that the embryo will be afforded "adequate protection" (47).

Last, a variety of non-English-speaking countries have attempted to address the regulation of reproductive technologies. They have ranged from legislative bills in Argentina that are prohibitive in nature, to bans on embryo research in Norway, to limits on oocyte freezing in Denmark, to French legislation that limits artificial insemination (AI) to heterosexual couples and prohibits embryo experimentation, to German bans on surrogacy contracts. Although this entry has focused on English-speaking countries, reproductive technologies are being used in a number of settings throughout the world. Regulatory responses are as varied as the countries themselves (48).

CONCLUSION

Among English-speaking countries, a great variety of approaches exist in regulating reproductive technologies. The United States has the most laissez-faire approach, with very little federal regulation. Most regulation is left to the private sector or the states; professional organizations such as the AMA and the ASRM have issued the most comprehensive guidelines regarding reproductive technologies. The United States seems to reflect Robertson's view of a strong sense of procreative liberty. Other English-speaking countries are more hesitant to grant such broad negative rights and they have all adopted to varying degrees a greater amount of formal regulation, whether in the form of the Authority in the UK or among the laws of the different states in Australia. In the late 1990s, however, certain international organizations, such as the Council of Europe, have attempted to formalize certain kinds of prohibitions with regards to some of the reproductive technologies available.

ACKNOWLEDGMENT

I would like to thank Patty Sokol and Amy Darby for their editorial assistance in preparing this article.

Unless explicity stated, the views expressed herein do not necessarily reflect official AMA policy.

BIBLIOGRAPHY

1. J.A. Robertson, *Children of Choice*, Princeton University Press, Princeton, NJ, 1994.
2. D. Roberts, *Killing the Black Body*, Vintage Books, New York, 1997.
3. *Skinner v. Oklahoma*, 62 U.S. 1110 (1942).
4. *Buck v. Bell*, 274 U.S. 200 (1927).
5. *Griswold v. Connecticut*, 381 U.S. 479 (1965).
6. *Eisenstadt v. Baird*, 92 U.S. 1029 (1972).
7. *Carey v. Population Services International*, 426 U.S. 918 (1976).
8. *Roe v. Wade*, 410 U.S. 113 (1973).
9. *Planned Parenthood v. Casey*, 502 U.S. 1056 (1992).
10. A. Charo, in *Encyclopedia of Bioethics*, 2241–2247, Simon & Schuster Macmillan, New York, 1995.
11. *Code of Medical Ethics*, Council on Ethical and Judicial Affairs, American Medical Association, Chicago, 1998–1999, Available at: *http://www.ama-assn.org/ethic/ceja.htm*
12. *Code*, Opinion 2.10.
13. *Code*, Opinion 2.145.
14. B. Furrow et al., *Bioethics: Health Care Law and Ethics*, West, St. Paul, MN, 1997.
15. *Code*, Opinion 2.18.
16. American Society for Reproductive Medicine (ASRM), ASRM Ethics Committee, date accessed: 1/26/2000. Available at: *http://www.asrm.org/current/press/abandon.html*
17. American Society for Reproductive Medicine (ASRM), ASRM Ethics Committee, Available at: *http://www.org/current/press/embsplit.html*
18. American Society for Reproductive Medicine (ASRM), ASRM Ethics Committee, Available at: *http://www.asrm.org/current/press/fetalegg.html*
19. American Society for Reproductive Medicine (ASRM), ASRM Ethics Committee, Available at: *http://www.asrm.org/current/press/posthum.html*
20. *In the Matter of Baby M*, 537 A.2d 1227 (1988).
21. H.D. Gabriel and E.B. Davis, *Loyola Law Rev.* **45**, 221 (1999).
22. *Johnson v. Calvert*, 851 P2d 776 (1993).
23. *Davis v. Davis*, 842 S.W.2d 588 (1992).
24. *Kass v. Kass*, 235 A.D.2d 150 (1997).
25. Fl. Stat. Ann. 742.17, 1997.
26. La. Rev. Stat. Ann. 9:123, 9:126, 1991.
27. *Hecht v. Superior Court of LA County*, 20 Cal. Rptr. 2d 275 (Cal. Ct. App. 1993).
28. Pub. L. 102–493, 42 U.S.C. 263a-1 et seq.
29. B.M. Knoppers and S. LeBris, *Current Opin. Obstetrics Gynecol.* **5**, 630–635 (1993).
30. *Commonwealth Law Bull.*, 981 (1992).
31. M. McDonald, *Toronto Sun*, February 25 (1998).
32. Bill C-47, 2nd Session, 35th Parliament, The House of Commons of Canada, 45 Elizabeth II, 1996–97.
33. K. McNorrie, in S. McLaren, ed., *Law Reform and Human Reproduction*, Dartmouth Publishing, Hanover, NH, 1991.
34. *Blackstone's Guide*, 190.
35. D. Morgan and L. Nielson, *J. Law, Med. Ethics* **21**(1), 31 (1993).
36. C. Hall, *The Independent*, March 16 (1994).
37. D. Fletcher, *Daily Telegraph*, July 2 (1994).
38. C. Morris, *UCLA Women's Law J.* **47**(8), 76 (1997).
39. P. Kasimba and P. Singer, *J. Med. Philos.* **14**, 403 (1989).
40. B. Bennett, *Reproductive Technology, Laws of Australia, Title* **20**, 91 (1995).
41. *Commonwealth Law Bull.*, January 1989.
42. The Reproductive Technology Act 1988 (SA).
43. The Human Reproductive Technology Act 1991 (WA).
44. P. Innes, *Melbourne Age*, 10 March 16 (1922).
45. M. Forbes, *Melbourne Age*, 1 April 9 (1995).
46. *In the Matter of Estate of the Late K and In the Matter of the Administration and Probate Act 1935*, No. M25/1996; Judgment No. A16/1996 (Supreme Court of Tasmania, 1996).
47. Council of Europe, Convention on Human Rights and Biomedicine, Oviedo, 04, IV, 1997, Available at: *http://www.coe.fr/eng/legaltxt/164e.htm*
48. E.A. Pitrolo, *Hous. J. Int. Law* **19**, 147 (1996).

See other REPRODUCTION entries.

REPRODUCTION, LAW, WRONGFUL BIRTH, AND WRONGFUL LIFE ACTIONS

JEFFREY R. BOTKIN
University of Utah School of Medicine
Salt Lake City, Utah

OUTLINE

INTRODUCTION

Wrongful life and wrongful birth are two closely related medical malpractice actions that have arisen since the 1973 *Roe v. Wade* decision. Both actions typically are brought against health care providers after the birth of a child with congenital malformations or a genetic disease. Wrongful *birth* actions refer to suits by the parents who claim harm from the birth of an impaired child. The claim is that had the parents been adequately informed of their reproductive risk, they would have taken measures to prevent the pregnancy or birth of the affected child. Wrongful *life* claims are brought in similar clinical circumstances; however, these claims arise from the child who claims harm from birth in an impaired condition. The child claims that but for the negligence of the health care provider, she would not have been born to suffer with her condition. Clearly the wrongful life claim poses a complex philosophical challenge. It is important to emphasize that the claims do not allege that the defendant *caused* the impairment through negligent actions. Rather the claims are based on allegations of inadequate or incorrect

information that would have permitted the parents to avoid pregnancy or to detect the abnormality prenatally and terminate the pregnancy. (Physicians who are alleged to have caused a congenital malformation through, say, the prescription of a teratogenic drug, are liable under more traditional tort actions.)

The wrongful life and wrongful birth suits have become increasingly common since *Roe v. Wade* for two reasons. First, *Roe v. Wade* established constitutional protection for abortion decisions through the first two trimesters of pregnancy, and, second, technology has offered an expanding array of tests and procedures to evaluate the health of the fetus. In light of these rights and choices, health care providers are seen to have parallel obligations to offer testing in a variety of clinical circumstances, and to adequately warn couples who have an increased risk of bearing a child with a heritable condition or congenital malformation. Failure to provide timely, accurate information according to the standard of care may leave providers liable under wrongful life and/or wrongful birth suits.

The rapid pace of research in human genetics and fetal imaging means that an ever larger number of conditions will be amenable to prenatal diagnosis in the future. Rare conditions, late-onset diseases and relatively mild health conditions may be identifiable early in a pregnancy. In addition future behavioral traits and normal physical characteristics may be predictable to some degree in an embryo or fetus. A clear challenge for the health professions, and for society more broadly, is to articulate the standards for prenatal diagnosis. How much information should prospective parents have access to about the biologic nature of their future children? The wrongful life and wrongful birth suits raise fundamental legal and philosophic issues about reproductive choice in an emerging era of powerful genetic technologies.

As will be discussed below, the wrongful birth suits have been widely successful in the U.S. court system, while the wrongful life claim has met with limited support. The wrongful life and wrongful birth claims should be distinguished from "wrongful pregnancy" suits in which parents claim damages for the birth of a health child following an alleged negligently performed sterilization or abortion procedure. Wrongful pregnancy suits will not be discussed in this article. This discussion will focus on the medical background of the suits, their legal history and the philosophic issues inherent in these claims.

CLINICAL CIRCUMSTANCES GIVING RISE TO WRONGFUL LIFE AND WRONGFUL BIRTH CLAIMS

Wrongful life and wrongful birth claims can arise from a variety of clinical circumstances. In the majority of cases to date, the claims have resulted from allegedly inadequate information provided to pregnant women about risks to their child. The list of conditions prompting wrongful birth or wrongful life suits is included in Table 1. The usual condition is a pregnant woman of an "advanced maternal age," that is, 35 years or older at the anticipated time of delivery, who is not warned of the increased risk of bearing a child with Trisomy 21 (Down syndrome) and other aneuploidy syndromes (e.g., Trisomy 18 and Trisomy 13). It currently is the standard of care to warn older pregnant women of their increased risk and to offer prenatal diagnosis. Without such a warning, and on the birth of a child with Trisomy 21, the obstetrician would be at risk for a wrongful birth suit by the parents and a wrongful life suit by the child. The suits are based on the assertion that had the warning been provided, prenatal diagnosis would have been pursued, the child's condition would have been detected and the pregnancy would have been terminated.

Table 1. Conditions Prompting Wrongful Birth and Wrongful Life Suits

Down syndrome	Fetal hydantoin syndrome
Congenital rubella syndrome	Leber's congenital amaurosis
Spina bifida	Infantile polycystic kidney disease
Tay Sachs disease	Duchenne muscular dystrophy
Cystic fibrosis	Anhidrotic Ectodermal dysplasia
Sickle cell anemia	Pelizaeus-Merzbacher syndrome
Larsen syndrome	Albinism
Retinoblastoma	"Hydrocephalus and multiple congenital defects"
Neurofibromatosis	No arms and other anomalies
Hemophilia B	"Severely deformed and retarded child"

The second single most common condition giving rise to these suits after Down syndrome is congenital rubella syndrome. Congenital rubella syndrome is due to a prenatal maternal infection with the rubella virus that can cause serious impairments and congenital malformations in the child. If a physician fails to assess the risk of the pregnant woman to rubella or fails to diagnose an active rubella infection, the physician is at risk for suit upon the birth of the impaired child. Again, the allegation here is not that the physician should have prevented the rubella infection, but that she should have provided information sufficient to allow prevention of the birth of the impaired child.

Obstetricians have been subject to the majority of wrongful life and birth suits to date. To the extent that these suits have been successful, obstetricians clearly have an obligation to assess pregnant women for the risk of congenital and hereditary abnormalities and to provide them information accordingly. It is important to note that these suits do not claim that the physician should have provided prenatal diagnosis or pregnancy termination. Providers are free to follow their own ethical standards in the provision of services and there are a substantial number of obstetricians, for example, who are "pro-life" and do not provide these services. In such circumstances physicians still are required to provide risk information according to the standard of care and information about testing options, such that patients can pursue services with other providers, if they wish. Suits may arise either from the failure to provide sufficient information, or the provision of inaccurate information. Clearly, the risk to the child need not be genetic in origin since infections

like rubella and other teratogenic infections and agents can cause impairments as well that may be amenable to prenatal diagnosis.

Prenatal care providers other than obstetricians have been subject to wrongful birth and life cases in the context of prenatal diagnosis. Ultrasound imaging of the fetus has become virtually routine in pregnancy. Failure of the radiologist to accurately diagnose a congenital malformation may give rise to these claims. Similarly failure to accurately perform prenatal tests may give rise to suits against laboratories that process the clinical specimens.

While prenatal care has given rise to most of these suits, care providers in other fields of medicine are not immune, and their liability will increase as genetic information increases. The case of *Schroeder v. Perkel* (1) provides an important example. In this case a child with cystic fibrosis was not diagnosed until seven years of age, despite suggestive symptoms for a number of years. Prior to the diagnosis of the child, a sibling was born who also was affected with cystic fibrosis. The parents claimed, successfully, that had an accurate and timely diagnosis of the first child been made, the parents would have been warned of their reproductive risk and they would have taken measures to prevent the birth of a second affected child. Cases also have arisen due to the provision of inadequate genetic information once an accurate diagnosis of a genetic condition was made. In the case of *Ellis v. Sherman* (2), a surgeon correctly diagnosed neurofibromatosis in an adult patient but failed to indicate that this is a hereditary condition. The patient subsequently fathered an affected child and sued the surgeon for failure to provide sufficient information to allow him to make an informed reproductive decision.

The success of this wrongful birth claim has implications for virtually all clinicians caring for patients who may have a genetic component to their illness. Typically physicians focus almost exclusively on the welfare of the individual patient in making decisions about testing and other evaluations. Genetic conditions may be part of a "differential diagnosis" for a patient (the list of possible conditions that might explain the symptoms), but unless there is a specific reason to pursue the diagnosis, other conditions may be considered first. Since genetic conditions are often difficult or impossible to treat, the diagnosis of genetic conditions may be delayed by the physician's desire to initially pursue treatable conditions. There often is little direct benefit to the patient in making a prompt diagnosis of a genetic condition. However, wrongful life and wrongful birth suits illustrate that there may be benefits to the *family* by making a prompt genetic diagnosis. Such a diagnosis will alert family members to their reproductive risks and potentially permit them to avoid the birth of an affected child. Therefore these legal claims herald a significant change in the responsibilities of health care providers from a narrow focus on the health of individual patient to a broader focus on the reproductive interests of the patient and the patient's family members.

LEGAL HISTORY OF WRONGFUL LIFE

Wrongful life and wrongful birth are tort actions or, more specifically, malpractice actions. As such, a successful claim against a health care provider requires that the plaintiff show (*1*) a duty existed on behalf of the provider to the plaintiff, (*2*) a breach of duty occurred (i.e., negligent conduct occurred), and (*3*) the plaintiff was harmed as a result of the negligence. In addition to these three elements, courts may consider broader public policy issues as they attempt to reach a just conclusion. The first case under the wrongful life term is notable in this regard. In *Zapeda v. Zapeda* (3), an illegitimate child brought suit in an Illinois appellate court against his father who had seduced his mother into intimate relations with a promise of marriage. The child sued his father for the harms associated with illegitimacy. The court was willing to recognize the duty, breach of duty, and harm but was unwilling to invite the flood of suits that might arise from children in similar circumstances.

The wrongful life claim has met with limited success in the U.S. judicial system. To date, five state courts have recognized the wrongful life claim (4), while 19 have rejected this tort. The primary stumbling block for the wrongful life claim has been the notion inherent in the suits that a child would prefer nonexistence to existence in an impaired condition. Recall that existence without the condition was never a possibility for these children, so the choice on behalf of the child was existence with impairments or nonexistence through contraception or pregnancy termination. The children in whose name these suits are brought must assert that, but for the negligence of the defendant, they would not exist. In response to this dilemma, most courts have adopted the reasoning first articulated in the New York case of *Becker v. Schwartz* (5), in which two basic problems with wrongful life suits were identified:

> The first, in a sense the more fundamental, is that it does not appear that the infants suffered any legally cognizable injury.... Whether it is better never to have been born at all than to have been born with even gross deficiencies is a mystery more properly to be left to the philosophers and the theologians. Surely the law can assert no competency to resolve the issue....Not only is there to be found no predicate at common law or in statutory enactment for judicial recognition of the birth of a defective child as an injury to the child; the implications of any such proposition are staggering.

The second problem identified by the Becker court was the inability to calculate damages on any reasonable basis.

In contrast, the courts that have recognized the wrongful life claims have been willing largely to overlook the philosophical problems inherent in the claim and to support the suits based on the medical needs of the child and/or the public policy advantages of deterring negligent medical care. A California court (6) in 1980 concluded:

> The reality of the "wrongful life" concept is that such a plaintiff both exists and suffers, due to the negligence of others. It is neither necessary nor just to retreat into meditation on the

> mysteries of life. We need not be concerned with the fact that had the defendant not been negligent, the plaintiff might not have come into existence at all.

Similarly the New Jersey Supreme Court in 1984 (7), was unwilling to allow the problematic logic of the wrongful life suits to stand in the way of what it judged to be a just outcome for the case. In the case at hand, the parents had been barred from bringing a wrongful birth suit on their own behalf due to the statute of limitations. Since the statute of limitations for suits by children is much longer than for adults, the only available route for the family to receive compensation for the alleged negligence was the wrongful life suit. The court stated that the child should not be denied adequate medical care simply because the parents were unable to sue on their own behalf.

For more than a decade, other states consistently declined to recognize the wrongful life claim. In recent years, however, courts in both Massachusetts and Connecticut have supported the tort. A 1997 decision by the Connecticut Superior Court (8), quotes a 1983 decision: "There is nothing illogical in a plaintiff saying 'I'd rather not be suffering as I am, but since your wrongful conduct preserved my life, I am going to take advantage of my regrettable existence to sue you.'" A Massachusetts court (9) was faced with a case in which a suit was brought against a physician who failed to report abnormalities on a fetal ultrasound and to repeat the examination. The parents of the child, born with heart and bowel abnormalities, relinquished the child for adoption. The court concluded:

> . . .Corey's parents are not entitled to recover against the defendant for the ongoing extraordinary costs that Corey will incur because of his defect (due to the fact that they are no longer his legal guardians or official parents). Nor will Corey's adoptive parents be entitled to recover, since they defendant owed them no duty. Therefore, this Court must consider whether Corey should have this cause of action since no one else can recover the extraordinary costs. . . .In this situation, it appears fair . . . to require the negligent Doctor to pick up these costs if negligence is proven.

Therefore in order to assure adequate care to a child with disabilities, some courts have been willing to recognize wrongful life claims without explicitly declaring that life with disability can be worse than nonexistence. Some commentators have noted that the incentive to recognize the wrongful life claim in selected courts would decrease if the U.S. health care system assured better services for all children with significant health care needs.

The New York court in *Becker v. Schwartz* deferred to the philosophers and theologians on the basic question of whether existence confers a harm on some children. A range of opinions have been offered on this question from bioethicists, theologians, and physicians. John Lorber (10), a British surgeon, wrote in 1975 of the deliberate nontreatment of some severely affected children with spina bifida:

> There are ethicists and moralists, as well as doctors, who consider that life must be maintained at any cost, because any life is better than no life. It may be legitimate to adhere to such principles within their own family, but is it not right to enforce such a philosophy on others who do not hold with it. To my knowledge none of the world's great religions or religious leaders believe that a severely defective innocent newborn infant would be worse off in heaven or wherever they believe their souls will go after death. Is it therefore humane to inflict an immense amount of suffering on such infants and on their families to ensure that they reach this heaven or haven in the end? . . .[Quoting De Lange, a neurosurgeon] "Large numbers of spina bifida children kept alive by early closure of the defect . . . are now adolescents, most of them painfully aware of the deficiencies. Some of us feel their presence not as a tribute to a medical achievement, but as an accusation against misuse of medical power."

Margery Shaw, a geneticist and attorney, argues that fetal abuse, through knowingly bringing a child to birth with a genetic condition, should be made analogous to child abuse in the law. She would sanction not only wrongful life suits against negligent physicians but similar suits against parents.

> . . .[P]arents should be held accountable to their children if they knowingly and willfully choose to transmit deleterious genes or if the mother waives her right to an abortion if, after prenatal testing, a fetus is discovered to be seriously deformed or mentally defective. They have added to the burdens of the other family members, they have incurred a cost to society, and, most importantly, they have caused needless suffering in their child.

Indeed, the wrongful life claim raises this curious question of the parent's responsibility for the birth of an affected child. When prenatal diagnosis detects a fetus with a genetic condition or congenital malformation, some parents choose to continue the pregnancy. Also parents at risk for bearing a child with a genetic condition may choose to forgo prenatal diagnosis and accept the risk of an affected child. As argued by Shaw, might the affected child have a wrongful life claim against the parents? The State of California was concerned enough about this possibility after the success of a wrongful life claim in the case of *Curlender v. Bioscience* that it passed legislation barring suits by children against parents for the harm of their existence (11).

In contrast to these authors, Bopp et al. (12) argue from a "right to life" perspective that one of the very foundations of modern law and civilized society is that all human life has enormous intrinsic value.

> . . .[W]rongful birth/life claims . . . require a new legal theory, in that life itself is considered a wrong, and death is preferred over life with disabilities. By deviating from the general principle, historically found in civilized law, that life, even with disabilities, is valuable and that only wrongful death is compensable, wrongful birth/life actions are a radical departure from fundamental legal philosophy.

Similarly authors writing from a disabilities rights perspective assert that it is simply wrong that those with disabilities lead lives of hopeless despair, devoid of the values that all others experience in their lives (13,14).

The greatest difficulties for those with impairments, it is claimed, are often not due to the condition per se, but to the discriminatory attitudes and barriers in society. Wrongful life suits (and wrongful birth) are seen by many of these authors as reflective of an inaccurate and inappropriate attitude in society toward life with a disability.

Finally, some bioethicists claim that the assertion that the life with impairments is worse than nonexistence is only justifiable for a few extremely severe conditions (15,16). From the perspective of the child, even the most rudimentary awareness and existence might be sufficient to experience a life of value. According to these authors, the kinds of conditions for which wrongful life suits have been brought, such as Down syndrome or congenital rubella syndrome, would not be justified from the perspective of the child.

LEGAL HISTORY OF WRONGFUL BIRTH

To date 26 states and the District of Columbia (17) and 3 federal courts (18) have recognized a cause of action for wrongful birth. One state has enacted legislation recognizing the validity of wrongful birth suits (19). In contrast, five appellate courts have rejected the claim (20) and six states have enacted legislation barring wrongful birth suits (21). Legislative bans have been prompted primarily by the philosophy that the birth of a child, even a child with significant impairments, should not be considered a harm to either the child or the parents. Two state laws banning wrongful birth have been upheld as constitutional (22). Although the national trend is clearly toward the recognition of the claim, wrongful birth remains controversial.

Several courts and scholars argue that the wrongful birth concept is an extension of the constitutionally protected right to privacy in abortion decisions (23). The claim is that abortion decisions are dependent on information about the welfare of the fetus. Therefore reproductive choice is limited if inadequate prenatal diagnostic information is provided. It is argued that the harm in these cases is not the birth of the impaired child, but the infringement on free choice in reproductive decisions.

In contrast, other commentators and courts argue that there is no basis for wrongful birth suits under the umbrella of privacy as articulated in *Roe v. Wade* (12,24). The constitutional right of privacy in reproduction and abortion only prevents state interference with abortion decisions, it is argued, and imposes no positive duties on health care providers to provide information about the fetus. Two state courts (Minnesota and Pennsylvania) have examined these arguments and held that the state laws barring wrongful birth suits are constitutional (22). Therefore, to date, the provision of prenatal diagnostic information has not been held to be a protected right under the Constitution.

Other commentators and courts argue that wrongful birth suits fall more appropriately under the patient's right of informed consent (25–27). Informed consent relates specifically to the amount and type of information that health care providers must provide to patients about medical options. It is argued that in the context of the medical condition of pregnancy, couples should be told the risk of a problem for the child in order to decide whether to obtain prenatal diagnosis. Under the current foundation for wrongful birth as recognized by the majority of the courts, physicians are held to the prevailing standard of care for the provision of timely and accurate information about the welfare of the child.

The requirement that the plaintiff (in this case, the parents) demonstrate harm secondary to the negligence of the defendant has not been carefully evaluated by the courts. In many cases, courts presume that the birth of a child with a impairment constitutes a harm. As noted above, many individuals with disabilities strongly contest this notion. In addition pediatricians and hospital ethics committees dealing with pediatric issues are often faced with the dilemma of parents of a severely disabled child who demand full medical support for the child, even when physicians believe such efforts are futile. Therefore whether a disabled child is a harm to the parents is a subjective issue, and this may change for parents between the birth of the child and later life with the child as a unique individual. Bopp et al. have captured the complexity of this issue by inviting us to imagine a woman with an impaired newborn who is driving to court to enter a claim of wrongful birth. An auto accident occurs and the child is killed. The mother now decides to enter a claim of wrongful death against the other driver. This case vignette illustrates the complex mix of benefits and burdens that children bring to families.

The complexity of the issue of harm is reflected in debate over the appropriate calculation of damages in courts recognizing the tort. There are several options that courts have considered that have tried to balance in various ways the benefits and costs of bearing and raising an impaired child. One method of calculation is to award the parents a monetary sum equal to the costs of the continued pregnancy, the delivery, and the medical and other costs incurred by child's impairment. These are seen as the additional costs directly incurred by the claimed negligence of the physician. Courts also may consider an additional award to compensate for the parent's emotional pain and suffering of bearing and raising a child with a disability. A third element that courts have variously considered is an offset to either of these damages for the benefits that a child brings to a family. Therefore the damages for emotional pain might be reduced by the jury's estimate of the child's positive contribution to the family.

Clearly, the concepts of emotional pain from bearing and raising an impaired child and emotional benefits of raising any child are highly value laden. As a result many courts have been unwilling to allow these kinds of calculations (or, in some circumstances, state law does not permit these kinds of awards or offsets). The majority of the courts have permitted damages to be awarded more simply for the medical and other extraordinary costs incurred by the child's unwanted condition (28).

PHILOSOPHICAL AND PUBLIC POLICY ISSUES

The prevalent acceptance of the wrongful birth concept suggests that this kind of malpractice liability will

encourage health care providers to conform to the contemporary standard of care. However, the standard of care in this arena is ambiguous and may be a moving target as new tests become available. At the present time it is the standard of care in the United States to inform women of advanced maternal age of their increased risk of trisomy syndromes, to offer Alpha Feto-Protein (AFP) screening to all pregnant women, and to offer couples from specific racial or ethnic backgrounds tests relevant to those backgrounds — Tay Sachs screening for Ashkenasic Jews, for example, and sickle cell screening for African-Americans. Further it can be concluded that any prenatal tests or procedures that are performed must be performed and communicated in a timely and accurate fashion. Currently many pregnant women of all ages are offered "triple screen" blood testing to detect a fetus with Down syndrome. This testing has a relatively poor predictive value, and it is uncertain at the present time whether courts would find a physician liable for failing to offer such testing.

Another specific test that remains the subject of controversy is screening for cystic fibrosis carriers in the general population. Cystic fibrosis (CF) is an autosomal recessive condition that primarily affects Caucasians of Northern European heritage. The carrier frequency in the Caucasian population is approximately 1 in 25 individuals. Genetic testing for CF is possible, however the sensitivity of the test remains less than 100 per cent. With the identification of the CF gene in 1989, there was prevalent speculation that carrier screening of the general population would be forthcoming. To forestall this development, professional societies promptly and clearly articulated an opinion that it was not the standard of care to offer CF carrier screening in the general population (29,30). To date, CF carrier screening has not be widely offered to pregnant women in the absence of a family history, although it is conceivable that a wrongful birth suit could force the issue and accelerate the adoption of this technology into general screening. With the development of each new major test, society will have to struggle with the question of whether, or when, it becomes the standard of care to offer the technology.

The broader philosophic issue raised by the wrongful birth concept is the limits, if any, that should be placed on prenatal diagnosis. If parents have a right to be informed of reproductive risks, what is the extent of such a right? Imagine a woman who carries a mutation in the BRCA1 gene that confers a lifetime risk of up to 85 per cent for cancer of the breast or ovary. Such a risk begins when a woman is in her thirties and forties. Should BRCA1 prenatal testing be offered and made available to this woman? Are parents harmed if their child has a genetic susceptibility to an adult onset disease like cancer or Huntington disease? At the extreme, can parents claim harm if a child of the "wrong gender" is born after an inaccurate prediction by ultrasound? It is unlikely that a flood of such suits will be brought to court, but the responsibilities articulated by successful suits encourages society to consider the appropriate boundaries for prenatal diagnosis.

A number of scholars and authoritative committees have raised concerns over the use of prenatal diagnosis for "mild" conditions or "trivial" indications. The President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (31) focused primarily on prenatal diagnosis for gender selection of the child, stating:

> The idea that it is morally permissible to terminate pregnancy simply on the grounds that a fetus of that sex is unwanted may also rest on the very dubious notion that virtually any characteristic of an expected child is an appropriate object of appraisal and selection. Taken to an extreme, this attitude treats a child as an artifact and the reproductive process as a chance to design and produce human beings according to parental standards of excellence, which over time are transformed into collective standards. . . .[T]he Commission concludes that although individual physicians are free to follow the dictates of conscience, public policy should discourage the use of amniocentesis for sex selection.

The Committee on Assessing Genetic Risks of the Institute of Medicine (32) took a similar stand, recommending that:

> . . .prenatal diagnosis not be used for minor conditions or characteristics. In particular, the committee felt strongly that the use of fetal diagnosis for determination of fetal sex or use of abortion for the purpose of preferential selection of the sex of the fetus is a misuse of genetic services that is inappropriate and should be discouraged by health professionals. . . .The committee believes this issue warrants careful scrutiny over the next three to five years as the availability of genetic testing becomes more widespread, and especially as simpler, safer technologies for prenatal diagnosis are developed.

A statement by the American Medical Association's Council on Ethical and Judicial Affairs supports limitation of prenatal diagnostic services to more serious conditions. The council suggests: "Selection to avoid genetic disorders would not always be appropriate. . . .[S]election becomes more problematic as the effects of the disease become milder and as they become manifest later in life" (33). Several scholars have taken similar positions. Thomas Murray concludes: "In short, we should not offer to provide prenatally information about traits or afflictions that are not substantial burdens on parent and child. We certainly should not assist couples in a misguided quest for the child that embodies their ideal collection of traits, including gender" (34). Like the President's Commission, other authors have framed the issue of limits of technology use around prenatal sex selection. Wertz and Fletcher state: "[W]e believe that it is important that the medical profession take a stand now against sex selection. A posture of ethical neutrality on this issue could lead to unfortunate precedents in moral thinking about future uses of genetic knowledge. . ." (35).

The courts have, on occasion, addressed the issue of the extent of the physician's obligation in the context of wrongful birth. The Supreme Court of Kansas wrote in a 1990 case: "In recognizing a cause of action for wrongful birth in this state, we assume that the child is severely and permanently handicapped. By handicapped, we mean, in this context, that the child has such gross deformities, not medically correctable, that the child will never be able to function as a normal human being" (36).

Certainly there are scholars and others in the general public who reject the notion of prenatal diagnosis and selective abortion entirely. After considering the rationales for prenatal diagnosis, Leon Kass concludes that there is no convincing "moral justification for the practice of genetic abortion" (37). In contrast, Philip Kitcher (38) writes:

> Couples who test their fetuses for the presence of blue eyes or curly hair and who decided to abort otherwise healthy fetuses when an alternative genotype was present would, to say the very least, have a distorted conception of value. Should states therefore limit the liberty of couples to make reproductive decisions? Not necessarily. Freedom of reproductive choice can be defended provided that the fetus is not taken to be a person with rights and interests that the state has a duty to protect. For the moment, we can accept the idea of the testing supermarket, open to free choice with few restrictions.

Similarly John Robertson argues that the legitimate concerns over unlimited access to prenatal diagnostic information do not warrant infringement on the parent's fundamental procreative liberties (39). As noted above, those from a disabilities rights perspective argue that the whole prenatal diagnostic enterprise largely reflects and reinforces negative stereotypes about living with a disability, or parenting a child with a disability. Drawing lines between types of disabling conditions to declare that some are appropriate for targeting with prenatal testing and that some are not is unacceptable for some disability rights advocates. They contend that such a policy decision sends a hurtful message to those living with disabilities who fall on the wrong side of the line.

The broad acceptance of wrongful birth suits and the diversity of opinion on the appropriate uses of prenatal diagnosis leaves medical care providers without clear guidance at the present time. What should the ethical practitioner offer to prospective parents from the expanding menu of tests? What tests should they provide upon request? Should tests for "mild" and late-onset conditions be made available? In the absence of a policy establishing professional standards on this issue, it may be up to the courts to decide if physicians have failed in their professional obligations when wrongful life or birth suits are brought. Wrongful birth and wrongful life suits will provide guidance to the profession on this issue, but malpractice litigation is a painful and inefficient approach to the development of a standard of care.

Prenatal diagnosis promises to be one of the most complex and controversial topics in medicine and in society generally over the next generation. John Fletcher predicts: "[T]he future of prenatal diagnosis and medical genetics will be ethically more complex than its past and present. Practitioners can expect a 'gathering storm' of issues that require societal involvement and establishment of public policy" (40). Society will benefit from mechanisms to reach some measures of consensus on these complex issues that lie at the interface of technology, philosophy, law, and human reproduction.

BIBLIOGRAPHY

1. *Schroeder v. Perkel* 432 A. 2d 834 (NJ 1981).
2. *Ellis v. Sherman* 515 A 2d1327 (Pa 1986).
3. *Zapteda v. Zapeda* 41 Ill App 2d 240, 190 NE 2d 849 (1963).
4. *Harbeson v. Parke Davis, Inc.*, 98 WA 2d 460, 656 P.2d 483 (1983); *Turpin v. Sortini*, 31 CA 3d 220, 643 P2d 954 (1982); *Procanic v. Cillo* 97 NJ 339, 478 A2d 755 (1984); *Rosen v. Katz* 1996 MA Super. Lexus 618; *Quinn v. Blau*, 1997, CT Super Lexis 3319.
5. *Becker v. Schwartz* 46 NY 2d 401, 386 N.E.2d 807 (1978).
6. *Curlender v. Bioscience* 106 CA App 3d 811, 165 CA Rptr 477 (1980).
7. *Procanic v. Cillo* 97 NJ 339, 478 A2d 755 (1984).
8. *Quinn v. Blau*, 1997, CT Super Lexis 3319. A.M. Capron, in A. Milunski and G.D. Annas, eds., *Genetics and the Law II*, Plenum, New York, pp. 81–93.
9. *Rosen v. Katz* 1996, Mass Super. Lexus 618.
10. J. Lorber, *J. R. Coll. Physicians* **10**, 47–60 (1975).
11. CA Civil Code §43.6 (West 1982).
12. J. Bopp, B.A. Bostrom, and D.A. McKinney, *Duquesne Law Rev.* **27**, 461–515 (1989).
13. A. Asch, in S. Cohen and N. Taub, eds., *Reproductive Laws for the 1990s*, Humana Press, Clifton, NJ, 1989, pp. 69–107.
14. D. Kaplan, in K. Rothenberg and E.J. Thomson, eds., *Women and Prenatal Testing: Facing the Challenges of Genetic Technology*, Ohio State University Press, Columbus, 1994, pp. 49–61.
15. J.R. Botkin, *J. Am. Med. Assoc.* **259**, 1541–1545 (1988).
16. B. Steinbock and R. McClamrock, *Hastings Cent. Rep.* **24**(6), 15–21 (1994).
17. *Bader v. Johnson*, 1997, WL 10352 (Ind. 1997); *Greco v. United States*, 1995, 111 Nev. 405; 893 P 2d 345 (NV 1995); *Flanagan v. Williams* 87 Ohio App. 3d 768; 623 N.E. 2d 185 (OH 1993); *Liddington v. Burns* 916 F. Supp 1127 (OK 1995); *Andalon v. Superior Court*, 162 Cal. App. 3d 600 (1984); *Haymon v. Wilkerson*, 535 A.2d 880 (DC 1987); *Fassoula v. Ramey*, 450 So.2d 822 (FL 1984); *Blake v. Cruz*, 108 253, 698 P.2d 315 (ID 1984); *Goldberg v. Ruskin*, 128 Ill. App. 3d 1029, 471 N.E.2d 530 (IL 1984); *Eisbrenner v. Stanley*, 106 Mich. App. 357, 308 N.W.2d 209 (MI 1981); *Smith v. Cote* 513 A.2d 341 (NH 1986); *Berman v. Allen*, 80 N.J. 421, 404 A.2d 8 (NJ 1979); *Becker v. Schwartz*, 46 N.Y.2d 401, 386 N.E.2d 807 (NY 1978); *Speck v. Finegold* 497 PA 77, 439 A.2d 110 (PA 1981); *Jacobs v. Theimer*, 519 S.W.2d 846 (TX 1975); *Naccash v. Burger*, 223 VA 406, 290 S.E.2d 825 (VA 1982); *Harbeson v. Parke Davis, Inc.*, 98 WA 2d 460, 656 P.2d 483 (WA 1983); *James G. v. Caserta*, 332 S.E.2d 872 (WV 1985); *Gallagher v. Duke Univ. Hosp.*, 638 F.Supp. 979 (NC 1986); *Pitre v. Opelousas General Hosp.*, 519 So.2d 105 (LA 1987); *Linenger v. Eisenbaum*, 764 P.2d 1202 (CO 1988); *Garrison by Garrison v. Medical Center of Delaware, Inc.*, 571 A.2d 786 (DE 1989); *Viccara v. Milunsky*, 406 Mass. 777, 551 N.E. 2d 8 (MA 1990); *Arche v. U.S. Dept. of Army* 798 P.2d 477 (Kan 1990); *Walker by Pizano v. Mart* 790 P.2d 735 (AZ 1990); *Reed v. Campagnolo* 630 A.2d 1145 (MD 1993); *Keel v. Banach* 624 So.2d 1022 (A 1993).
18. *Roback v. United States*, 658 F.2d 471 (7th Cir. 1981); *Gildiner v. Thomas Jefferson Univ. Hosp.*, 451 F.Supp. 692 (E.D. PA 1978); *Phillips v. United States*, 575 F. Supp. 1309 (D.S.C. 1989).
19. Me. Rev. Stat Ann. tit. 24, 2931 (Supp. 1988).
20. *Schork v. Huber*, 648 S.W.2d 861 (KY 1983); *Wilson v. Kuenzi,* 751 S.W.2d 741 (MO 1988); *Azzolino v. Dingfelder*, 315 N.C. 103, 337 S.E.2d 528 (1985), cert. denied, 479 U.S. 835 (1986); *Spencer v. Seikel*, 742 P.2d 1126 (OK 1987); *Atlanta Obstetrics & Gynecology Group v. Abelson*, 398 S.E.2d 557 (GA 1990).

21. Idaho Code 5-334 (Supp. 1986); MO Stat. 145.424 (1987 Supp.); MO Ann. Stat 188.130 (Vernon Supp. 1987); S.D. Codefied Laws Annual 21-55-2 (Supp. 1986); Utah Code Ann. 78-11-24 (1986 Supp.); 42 PA Const.Stat.Ann. 8305 [Purdon, 1990, Supp.].
22. *Hickman v. Group Health Plan, Inc.*, 396 N.W.2d 10 (MO 1986); *Dansby v. Thomas Jefferson University Hospital* 623 A.2d 816 (PA Super. 1993).
23. Note. *Harv. Law Rev.* **100**, 2017–2034 (1987). R.K. Johnston, *University of Missouri at Kansas City Law Rev.* **57**, 337–353 (1989). K.J. Jankowski, *Fordham Univ. Law J.* **17**, 27–62 (1989); *Smith v. Cote*, 513 A.2d 341 (N.H. 1986); *Haymon v. Wilkerson*, 535 A.2d 880 (DC App 1987); *Hickman v. Group Health Plan, Inc.*, 396 N.W.2d 10 (Amdahl, dissenting opinion); *Roback v. U.S.*, 658 F.2d 471.
24. J. Lyons, *Villanova Law Rev.* **34**, 681–696 (1989).
25. J.R. Botkin and M.J. Mehlman, *J. Law Med. Ethics* **22**(1), 21–28 (1994).
26. See *Harbison v. Parke Davis, Inc.* 656 P.2d 483 (Wash 1983).
27. P.S. Fox, *Washington and Lee Law Rev.* **44**, 1331–1356 (1987).
28. See *Bader v. Johnson* 1997 WL 10352 (Ind App) for a recent summary.
29. American Society of Human Genetics, *Am. J. Hum. Genet.* **46**, 393 (1990).
30. Workshop on Population Screening for the Cystic Fibrosis Gene, *N. Engl. J. Med.* **323**, 70–71 (1990).
31. The President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Screening and Counseling for Genetic Conditions*, U.S. Government Printing Office, Washington, DC, 1983, pp. 57–58.
32. Committee on Assessing Genetic Risks, Division of Health Sciences Policy, Institute of Medicine, in L.B. Andrew, J.E. Fullarton, N.A. Holtzman, and A. Motulsky, eds., *Assessing Genetic Risks: Implications for Health and Social Policy*, National Academy Press, Washington, DC, 1994, p. 105.
33. American Medical Association Council on Ethical and Judicial Affairs, *Arch. Fam. Med.* **3**, 633–642 (1994).
34. T.H. Murray, *The Worth of a Child*, University of California Press, Berkeley, 1996, p. 139.
35. D.C. Wertz and J.C. Fletcher, *Hastings Cent. Rep.* **19**(3), 21–27 (1989).
36. *Arche v. U.S. Dept. of Army* 798 P.2d 477 (Kan. 1990).
37. L. Kass, *Toward a More Natural Science: Biology and Human Affairs*, Free Press, New York, 1985, pp. 80–98.
38. P. Kitcher, *The Lives to Come: The Genetic Revolution and Human Possibilities*, Touchstone Books, New York, 1997, p. 85.
39. J.A. Robertson, *Children of Choice: Freedom and the New Reproductive Technologies*, Princeton University Press, Princeton, NJ, 1994.
40. J.C. Fletcher, in A. Milunski, ed., *Genetic Disorders and the Fetus: Diagnosis, Prevention and Treatment*, Plenum, New York, 1986, pp. 819–859.

See other entries GENE THERAPY, ETHICS, GENE THERAPY FOR FETUSES AND EMBRYOS; REPRODUCTION, ETHICS, PRENATAL TESTING, AND THE DISABILITY RIGHTS CRITIQUE; REPRODUCTION, LAW, IS INFERTILITY A DISABILITY?; REPRODUCTION, LAW, REGULATION OF REPRODUCTIVE TECHNOLOGIES.

RESEARCH ON ANIMALS, ETHICS, AND THE MORAL STATUS OF ANIMALS

LILLY-MARLENE RUSSOW
Purdue University
West Lafayette, Indiana

OUTLINE

INTRODUCTION

Animal research involves not only scientific but ethical issues. Indeed, these are not two separate, or separable, issues; they are inextricably interwoven. The ethical side of this issue concerns researchers, regulators, concerned citizens, and society at large. Even though questions about animal experimentation are not the exclusive domain of philosophers and ethicists, a careful look at relevant philosophical arguments and moral theories can give us a better understanding of the nature of the debate. The following discussion takes up the issues listed in the title in reverse order.

THE MORAL STATUS OF ANIMALS

The term "moral status" refers to the place of something with respect our ethical reasoning. An alternative way of phrasing the question of moral status is to ask whether something has moral standing. To say that something has no moral standing is simply to say that it does not enter directly into proper ethical deliberations at all. Such things may figure into deliberations indirectly, as a source of concern for those who do have moral standing. My property, for example, does not have moral standing, but *I* can be benefited or harmed, my rights can be respected or violated, depending on how others treat my property. Thus, we might speak of other people's indirect duties to treat my property in certain ways. Traditionally this has been the status accorded to animals by theologians, philosophers, and legal and social practice (1–3). We have duties not to be cruel

to animals not because they are worthy of our direct moral concern, but because how we treat animals affects humans — because the animals are their property, because cruelty to animals might lead to mistreatment of humans, and so on.

There are two other possible positions on the moral status of animals, one on either side of the 'indirect duties' view just sketched. One might hold that if we set aside questions of property, human concerns for animal welfare are simply misguided, and should not be part of a well-grounded ethical theory at all. This view is admittedly uncommon, but not unheard of (4–6). As one noted neuroscientist states: "I believe that the inclusion of lower animals in our ethical system is philosophically meaningless and operationally impossible and that, consequently, antivivisectionist theory and practice have no moral or ethical basis" (4, p. 169) The other is to accord animals moral standing, to say that we have direct duties toward them independently of how our actions affect other humans, and that they ought to figure directly in our ethical deliberations. It should be noted that moral standing need not be egalitarian: There is a large gap between saying that animals have moral standing, and saying that their moral standing is in any way equal to that of a human being. This position on the moral status of animals encompasses a wide range of positions, from radical "liberation" theories such as those defended by Peter Singer and Tom Regan (7,8) to those that defend the use of animals for research and agriculture but still argue that we have a direct moral obligation to minimize animal pain and suffering (9,10). In all these cases we must distinguish the question of whether X is a moral agent — the sort of thing which has duties, moral obligations, can be praised or blamed — from the question of whether X has moral standing — whether it is the sort of thing toward which moral agents have direct duties. The term "moral patient" is often used to designate this latter category. While some ethical theories hold that only moral agents can be moral patients, most theories entail two different sets of criteria for the two groups. Thus it will not do to reject the idea that animals have moral standing (are moral patients) simply by noting that "they don't respect *our* rights," or "they don't have any compunction about killing *us*."

Most people who are not deeply involved in the animal rights debate tend to endorse a "middle-of-the-road" position regarding our moral obligations to animals. Thus, it is widely accepted that it is morally wrong to cause an animal pain or suffering without a compelling overriding good reason. Although there is much less agreement about what sorts of things constitute a good reason for overriding this injunction, most people would also agree that the suffering of the animal itself, not just the side effects on humans, is morally relevant. On the other hand, most of these same people would object quite strongly to the view that animals are in any way entitled to the same sorts of moral protection owed to other human beings; they would hold for example, that much less justification is needed for causing an animal to suffer than is required for the same amount of human suffering, that we are justified in using animals in ways that we ought not use severely retarded orphaned infants, and that injunctions against killing animals are much less stringent and restrictive than similar injunctions against killing humans.

The position (or constellation of positions, for there are many variations within the boundaries sketched) described in the previous paragraph would apparently be endorsed by the vast majority of Americans and Europeans. It is decidedly not, however, the dominant tendency in philosophical writings on ethics and animals. Most of the contemporary philosophical writings argue some variant of the claim that the position described above sanctions much that is morally reprehensible, and that when correctly understood, our moral obligations regarding animals would disallow currently accepted practices such as meat-eating, much of the research in which animals are used, and more (7,8,11–13).

At this point, let us introduce some labels for the sake of convenience. Let us call the sort of position just described "the liberationist view." It is more popularly called "the animal rights position," but this is inaccurate, since at least some of the defenders of the position — notably Peter Singer — do not appeal to rights-talk in their arguments; although the word "rights" crops up occasionally in (7), in his more careful philosophical arguments, he explicitly rejects the notion of animal rights (14,15). However, for better or worse, the label "animal rights" is the one that has been most commonly associated with this strong position; problems with the label will be discussed more fully in a subsequent section. The first sort of position described above admits that there is a basis for direct moral obligations to animals, but holds that it is different, at least in part, from the basis for our obligations to humans; for that reason, it shall be labelled the "differential view." Finally, there is the view that we have no direct obligations or duties to animals, that even the duty to avoid cruel treatment of animals is based on our moral obligations to humans. This shall be refered to this as the "humanist position."

Clearly, these are less-than-ideal terms for views that have been coarsely defined. The arguments within each category differ in important respects that must be considered more carefully. Nonetheless, the labels do identify some clearly distinct trends, and it will be convenient to be able to refer to those trends in a shorthand way. Details and specifics will find their proper place later. It should also be noted that these views cut across the theoretical commitments we will examine presently: One can, for example find utilitarians in all three categories (7,16,17), and while some deontologists (8) subscribe to a liberationist view, others are definitely in the humanist camp (18).

HISTORY AND THEORY

If the "differential view" is the one most people hold, how did the animal liberationist position gain such prominence? At least in the United States, the answer can be traced back to Peter Singer, and the publication of *Animal Liberation* (19).

Animal Liberation

Singer follows Jeremy Bentham and other proponents of "utilitarianism" to support his opposition to animal research or raising animals for meat. Utilitarianism is the view that an action is right if and only if it produces a better balance of benefits and harms than available alternative actions. That is, our ethical evaluation looks only at the consequences of our actions. Utilitarians may disagree about what constitutes a benefit but two widely accepted candidates are pleasure and satisfaction of preferences and interests. Singer argued that sentience — the ability to feel pleasure and pain, and hence to have interests — should be the basis of our ethical assessment of any action, including animal research. The examples he used to support his claim that animal research does not produce the best possible balance of benefits vs. harms have come under serious attack, both for the sketchiness of the descriptions, and for questions of scholarship (20). Nonetheless, they served to make many more people aware of the sorts of research that had been conducted. Even if the specific examples cited are problematic, the theory does give a way of evaluating animal research, both as a general practice, and specific research projects. This approach of trying to weigh the (expected or possible) benefits against the harms, usually to the animals used, is reflected in the current regulations, albeit in an attenuated form.

The Argument from Marginal Cases

More important from the standpoint of theory, Singer launched the first sustained attack on the differential view, arguing that any attempt to justify the use of animals in research would also justify the use of some humans. This argument, which has come to be known as the argument from marginal cases, has become one of the chief weapons in arguments against animal research. It is also intended to support the claim that "speciesism" — a term coined by Richard Ryder (21) but popularized by Singer — is wrong in the same way that racism or sexism is wrong.

The argument is deceptively simple. If we want to justify differential treatment of animals and humans, we must cite a morally relevant difference. However, the differences that have been proposed — such as self-consciousness, autonomy, or the ability to act as moral agents — do not apply to all human beings. Some severely retarded humans, those in a permanent vegetative state, or, to cite the most dramatic example, anencephalic infants, do not possess those qualities now, may never have them, or have possessed them in the past. Thus any argument that attempts to justify animal research on the basis of one of these properties would also justify using these "marginal case" humans for the same procedures. The term "marginal case" may have originally intended to refer to both animals and humans who are at the margins of whatever line we attempt to draw, but it has come to be used to refer to those humans who are severely retarded or otherwise so impaired that they are not capable of very basic perceptions, emotions, and understanding. Since this term is offensive to many, the following discussion will break with tradition and use the term "misfortuned humans" rather than "marginal cases."

There are three possible responses to this argument. The first, best associated with a "humanist" view of the moral standing of animals, is to attempt to identify a morally relevant property that does separate all humans, even "misfortuned humans," from animals. The only candidate for such a property is membership in the human species; anything else, such as the capacity for reason or moral agency, is not going to apply to severely misfortuned humans. This is often disparaged as "speciesism," but has been defended occasionally (22). The second response, defended by R.G. Frey (23), is to argue that there are important differences between normal adults and animals, but also agree that some humans will lack the morally relevant qualities, and accept the other horn of the dilemma: Some research on misfortuned humans would be justified, perhaps even better than animal research. This would be consistent with (although not entailed by) the "differential" view. The third, dictated by the "liberationist" view, is to accept Singer's conclusion, and still insist that research on misfortuned humans would be immoral, but that entails that we are not morally justified in doing research on sentient creatures that meet or exceed the conditions that protect misfortuned humans. Let us look at each of these in a bit more detail.

If one wants to pursue the first strategy and cast about for a difference that might be morally relevant, it is clear, as already noted, that the standard properties — such as autonomy, rationality, the ability to engage in moral reasoning and mutual respect — will not do. Some people have tried to generate a longer list, and even suggest that we ought to adopt a "cluster" approach: none of the qualities individually are necessary, but any small subset of them is sufficient to justify preferential treatment for humans (24). These are exactly the properties that misfortuned humans lack. There are three other possibilities that have been suggested: (*1*) arguments from potentiality: marginal cases have (or did at one point have) the potential for the morally relevant property, even though they do not actually have it, (*2*) defenses of "speciesism," or arguments that attempt to justify giving preferences to individuals just because they are members of our own species, and (*3*) appeals to side effects, or how misfortuned humans must be afforded special protection for the good of others, not necessarily for the misfortuned human. The first of these will give us either too little or too much. Most advocates of this view restrict their attention to the potential that a misfortuned human has now (25), but it is simply false that most of them have the potential in question. An anencephalic baby or someone who is brain dead does not now have such a potential, and appeals to the fact that medicine might someday be able to help such cases is simply otiose; such resources are not available now. Others extend the notion of potentiality even further. They say that the individual had the potential at some point but lost it due to misfortune (26), or a slight variation, that misfortuned humans actually have some moral status in virtue of the fact that they could have possessed some property (27). This may be true, except for those defects that are genetically fixed at conception and manifest during fetal development. However, this casts the net of potentiality too wide; every germ cell (indeed, since

the advent of Dolly the sheep, perhaps every human cell) has, under the right conditions, the potential to develop into a normal human being. The idea that every human sperm should be accorded a higher moral status than a chimpanzee, or even a laboratory rat, appears to be a reductio ad absurdum of this approach. Finally, those who seek a morally relevant difference between misfortuned humans and animals might point to the so-called side effects of treating such humans as potential research subjects (28). The side effects include the emotions of the parents, other interested observers, possible changes in attitudes toward the medical profession, and weakening of the familial bond. The problem with these side effects is that cultural factors are the main determinant of how strong or deleterious these effects will be. If our obligations towards misfortuned humans rest solely on these factors, they will embody a "cultural relativist" approach to ethics, an approach for which ethical theorists have long had persuasive refutations.

Consider then the second response to the argument from misfortuned humans, championed by R.G. Frey (16,23). Frey argues that the value of life depends on the quality of life, and the quality of life is determined by the opportunities for rich experiences; he like Singer, is a utilitarian. He also agrees that on this criterion, some animals will fare better — rank higher in terms of quality, and hence value of life — than some misfortuned humans. At the same time he rejects the liberationist claim that animal research is universally unjustified. While he agrees that some work is frivolous, trivial, or simply bad science, he cites numerous cases in which he argues that the benefits outweigh the harms. His solution is to raise the stakes: We can justify such beneficial research, but only if we would be willing to do the same research on a misfortuned human who falls lower on the quality/value of life scale than the animal proposed as the research subject. Although his view is perfectly consistent, most of his critics continue to cast about for another way to avoid grasping this particular horn of the dilemma.

The third option, as already noted, is the one preferred by Singer, Regan, and other critics of animal research. The conclusion here is that since we would not sanction most research on misfortuned humans (although there are some exceptions), we should not allow the same sort of research on animals who match or exceed those humans with respect to morally relevant criteria. This poses a very basic challenge to defenders of animal research, in that it seems to rest solely on a demand for logical consistency, rather than allegiance to one or another arcane-sounding ethical theory. Although Singer situated the argument from misfortuned humans in a utilitarian context, the argument is by no means limited to that context. Tom Regan (8) and James Rachels (29) utilize the same sort of argument to argue that any argument that establishes that all humans have the right to life, liberty, or respect will also apply to animals. Indeed, at least one philosopher, Evelyn Pluhar, makes the argument from misfortuned humans so central to her argument that no specific commitment to a broader moral theory is required (12).

Rights-Based Arguments

Although Singer's arguments are the most generally straightforward, accessible, and therefore best known by the general public, other philosophers have given more complex and sophisticated arguments to examine the contention that animal research (as well as large scale animal agriculture) should be abolished. In doing so, surprisingly, they employ a theory most usually associated with "human rights." Some of these are still utilitarian in flavor, but others adopt a more "deontological" approach.

The history of deonotology, as well as most current versions of deontological theories, emphasize the unique status of rational, autonomous, human agents. This is in marked contrast with the utilitarian tradition, which from its first formal articulation by Jeremy Bentham, has frequently acknowledged the inclusion of animals (30). In contrast to utilitarianism, deontology insists that some actions may be right even if the consequences are not good, or as good as they could be, while other actions are ethically wrong even though they would produce good consequences. In short, consequences are not the only factor in moral evaluation. Examples typically cited are lying (which would be wrong even if you and I were both happier if I lied to you), keeping promises, and justice. The term "deontology" is often used interchangeably with "right-based theory," since the concept of rights has often been used as a guard against sacrificing what is right (e.g., respect for individual autonomy in matters of religion) for what might be good for the group (forcing compliance to the majority choice). It is also sometimes called a "Kantian theory," since the German philosopher Immanuel Kant was its first major proponent. Kant famously intended his theory to apply only to rational beings capable of understanding moral imperatives, and explicitly excluded animals from its scope, and until recently that focus was unquestioned. Indeed, the sorts of rights typically associated with deontology are often called "human rights." As noted above, even contemporary philosophers who develop and defend deontological or rights-based theories rarely grant animals rights. However, that exclusion has been challenged.

Tom Regan, for example, follows Kant in rejecting utilitarianism as an adequate moral theory in favor of a theory that emphasizes rights that cannot be overridden merely because such an override would yield good consequences for everyone affected. Bernard Rollin also defends a rights-based approach, although he is neither strictly Kantian nor an "abolitionist" about either animal research or animal agriculture (31,32). In both arguments, however, the concept of rights is consistent with almost universally accepted social and political philosophy: for example, that a person cannot be killed (perhaps to harvest his heart, lungs, and other tissues and organs) just because more people would benefit from his death than would benefit from his continuing to live. The novelty is extending this concept of rights to animals and that, at least in Regan, is justified by a form of the argument from misfortuned humans.

For Regan, the attribution of rights is closely tied to an obligation to respect another's life and inherent value. He rejects a Schweitzer-style reverence for all life and argues

instead that we ought to attribute equal inherent value to what he calls "subjects of a life," individuals who have —

> beliefs and desires, perception, memory, and a sense of the future, including their own future, an emotional life together with feelings of pleasure and pain; preference and welfare interests; an ability to initiate actions in pursuit of their desires and goals; a psychophysical identity over time; and an individual welfare in the sense that their experiential life fares well or ill for them, logically independently of their utility for others, and logically independently of their being the object of anyone else's interest (8, p. 243).

The argument that all subjects of a life have inherent value, and hence deserve respect, depends in part on a version of the argument from misfortuned humans. Regan does not attempt to draw a sharp line between animals that are subjects of a life and those that are not, preferring to leave that issue open pending a better understanding of various animal's psychological abilities. He does claim that *at least* all normal mammals over one year of age are subjects of a life. From the notion of inherent value and what he terms "the respect principle," he concludes that animals have certain rights that are equal in moral weight to those rights possessed by humans. These include, most basically, the right to respectful treatment and the right not be harmed.

The fact that Regan's argument to show that animals have rights represents a minority position raises an interesting question: Do his arguments show that "traditional" deontological theories are merely myopic or biased, or do those traditional restrictions of rights to human beings have a legitimate basis? Both Regan and Rachels rely heavily on the argument from misfortuned humans, but there are at least two other questions to be raised. They can most easily be seen by setting out a condensed form of Regan's argument:

1. All "subjects of a life" have rights.
2. Many animals — at least those noted above — are subjects of a life.
3. Therefore many animals have rights.

Thus the two obvious questions to raise are (*1*) whether being a subject of a life is in fact the appropriate criterion for attribution of rights, and (*2*) if so, whether most animals really are subjects of a life in the sense defined by Regan. Although it might seem as if the first is the domain of philosophers and the second a question to be answered by psychologists, ethologists, and other philosophers, in reality the two are inextricably intertwined. The debate is too complex to pursue here, but it is ongoing (33,34). Quite a lot is at stake here. No matter what Peter Singer says, utilitarianism will necessarily sanction at least some research on animals. Only if Regan can successfully establish his rights-based view can he even approximate the "total abolitionist" stance toward animal research that he advocates.

Other Philosophical Foundations

Although utilitarianism and deontological theories tend to dominate the theoretical landscape, they are by no means the only options. At least two other approaches to ethical theory, contractualism, and what shall be referred to as Humean ethics, have been used to address animal issues explicitly. A third, virtue theory, is enjoying renewed interest among philosophers, but does not seem to have been used by contemporary philosophers to address issues about the use of animals, even though Aristotle, widely regarded as the founder of virtue theory, had quite a bit to say about animals (35). Finally, several philosophers writing on animals do not fit neatly into any of these categories, either because they have theoretical commitments that combine elements of one or more approach, or because their arguments are not as "theory driven" as, for example, Regan's is. The discussion that follows will not attempt a critical analysis of these approaches but will describe them briefly in order to provide a more complete overview.

Contractualism, as the name implies, views moral obligations as the outcome of an implicit or hypothetical contract among members of a society. Such contracts are assumed to have the form: I agree to refrain from doing x, y, and z to you, and to do a, b, and c, if and only if you agree to refrain from doing x', y', and z' to me, and to do a', b', and c'. The "prime" indicators are meant to indicate that while contracts are essentially reciprocal arrangements, they need not be symmetrical. One can have a contract between employer and employee, or sovereign and subject, where the rights and duties of the participants might differ. The essential feature of contractualism is that all parties agree to abide by the rules; hence only creatures who are able to understand an abstract concept of rules and the intention to follow them can enter into the moral sphere and have direct moral standing. Contractualism is sometimes confused with deontological theories, and some philosophers (36) have combined elements of both. But there are important differences between them, some of which will come out in the following paragraphs, and some of which are too complex to address here. What follows will often speak of "contractors" and "contracts" for the sake of convenience, but it is important to remember that these typically refer to hypothetical, not actual, contracts.

It is probably already obvious that animals will not fare well on a contractualist approach to ethics: they, like children, are judged incapable of understanding the abstract rules and reasoning implied by the whole notion of a contract. Being unable to enter into a contractual agreement, they are also ineligible for the protection such contracts provide.

Clearly, contractualism sees moral duties as holding directly only among creatures capable of understanding and abiding by such an abstract contract. As the contrast between Regan and Kant indicates, the basis for obligation or duties in a deontological theory is not necessarily so restricted: Some versions of deontological theories have room for duties to animals, but no contractualist theory can possibly do that: The best it can muster is a "contract" in which participants choose to include certain restrictions on the way we treat animals because the contractors would prefer not to live in a society which, for example, tolerates blatant cruelty. As we have seen, this sort of protection at best generates an indirect duty

toward animals. Moreover such protection is to a large extent voluntary, an "option" for the contractors, unlike the strict imperatives generated by a deontological theory. These differences entail that contractualists may differ in their descriptions of what indirect duties, if any, might be included in the hypothetical contract, and we do find the expected variations in the views of contractualists who have written about animals. The two most prominent philosophers in this category are Jan Narveson (37,38) and Peter Carruthers (6). Narveson thinks that contractors might well choose to extend some protection to animals, while Carruthers argues quite vehemently that it would be irrational and misguided to agree to such an extension of contractual protection.

One of several problems with contractualism is the fact that the moral status of children and other "disadvantaged" individuals — those lacking the capacity to understand and hence legitimately agree to the sort of contract required for admission into the moral sphere — are covered only at the whim of the contractors. Some humans in this category might be covered by dint of the contractors' self-interest; if, but only if, they think they will care about the welfare of their children, they might choose to include some protection for children in the contract. However, if they believe that they will only care about their healthy, or "normal," or male offspring, or children born in wedlock, only those children will be protected; no moral agent will have any duties to those "beyond the pale." More generally, since morality is *defined by* the hypothetical contract, there is no way of saying that the contract itself is unethical, unfair or unjust, or fails to capture some real moral obligation; one can only raise a legitimate objection by pointing out that no rational agent would agree to some putative contracts. These features of contractualism have led most philosophers to reject it as an adequate moral theory, and if it is inadequate as a general analysis of morality, it surely cannot be appealed to as a basis of deciding questions about moral obligations towards animals.

Implicit in at least some popular invocations of contractualism are two ineffective justifications: (*1*) Since animals do not respect our rights, we do not have to respect theirs, and (*2*) morality is a human invention, and thus applies only to humans. The first is problematic because it assumes that only moral agents can be moral patients, something that needs to be argued for rather than presupposed. The second is even weaker: After all, humans "invented" mathematics and science, but that does not mean we can make them do anything we want nor that they only apply to humans.

The collection of views that may be called *Humean*, in recognition of the British philosopher David Hume, represents another approach to ethical theory in general, and the issue of animals and animal research in particular. The views have this in common: They hold that ethics is not something that requires only abstract, impersonal rationality. Rather, it must also involve emotions, including sympathy or empathy, and avoid the complete detached impartiality and abstractness of overly rationalistic approaches. Hume is famous for claiming "Reason is, and ought to be, the slave of the passions." The term "passion," in Hume's day covered a wide range of emotional and personal attitudes, but the basic unifying theme is that, in contrast to the impersonal nature of deontological theories and utilitarianism (both of which hold that, all other things being equal one has no ethical justification for giving preferential treatment to one's best friend, spouse, or child), personal, concrete relations do matter as a foundational concern in ethics. Mary Midgeley, for example, justifies her conclusions about our duties toward animals by arguing against excessive emphasis on rationalism, and for the importance of emotion in ethical reasoning (39). Annette Baier also emphasizes Humean themes in her discussions of animals (40). This de-emphasis of rationality has also been a prominent theme in feminist discussions of animal experimentation (41). A rejection of rationality and impartiality seems ripe for ridicule, but both Midgeley and Baier argue extensively that traditional ethical theories have overstated the role of very abstract, detached reasoning.

ANIMAL WELFARE AND ANIMAL RIGHTS

Defenders of animal research often try to draw a deep theoretical distinction between animal welfare and animal rights. They often assert that researchers who use animals, or farmers who raise animals for food, are deeply concerned about animal welfare but ought to reject the notion of animal rights. A more careful analysis of these concepts, however, reveals, that the distinction invoked here is far from clear and often inaccurate (31,34,42), and is usually divisive rather than clarificatory. It is divisive because it frames the discussion in terms of just two sides rather than recognizing the whole spectrum of subtle differences, and because the two sides are presented in an "us versus them" tone rather than looking for points of agreement. It is inaccurate because, as we have seen many "liberationists" such as Peter Singer do not advocate animal *rights*. Moreover many philosophers who argue that animals do have rights reject the idea that these rights afford them the sort of total protection that the label is usually taken to represent (43), and they also explain why talk of animal rights does not entail anything like equal treatment (44). Finally, as reported by Rollin, 80 percent of respondents to a recent poll affirmed that animals have rights (31, p. 149). Thus advocating animal rights is neither necessary nor sufficient for holding the extremist position that the label "animal rights" is often taken to represent. The label "animal welfare" is similarly unhelpful. It usually involves a rejection of the humanist view, as defined at the beginning of this article (although some who invoke the "animal welfare, but not animal rights," slogan seem to attribute only indirect moral status to animals); even so, it covers such a broad spectrum of views that it sometimes tends to become mere window dressing. One end of the spectrum, it could imply a position as strong as Singer's: Animals do not have rights, but their welfare deserves equal consideration, and harms to animals must be weighed against potential human benefits (the converse is true, in principle, but that rarely arises as a moral issue). As we have seen, this would require a major re-evaluation of animal research. On the other end

of the spectrum, "animal welfare" is merely an injunction not to harm an animal unnecessarily, unless such harm is dictated by some human interest (including any desire to know, economic considerations, personal taste, and entertainment). But if these labels are unhelpful, what vocabulary should be used? First, we need a more accurate understanding of the term "animal rights," something which is best obtained by going back to the philosophical roots of theories of rights.

There are, of course, philosophical disagreements about how best to understand rights-claims and theories about rights. Accordingly, a brief survey of the available options will be helpful. The most important choice for our understanding here has to do with the force of rights-claims: What do such claims entail or suggest, and how are they different from other sorts of claims about moral obligations? We soon find an important distinction between what may be called "broad" and "narrow" views on the force of rights-claims. Roughly speaking, a broad interpretation of rights-claims sees them as alternative ways of expressing a wide variety of moral obligations. To say that someone has a right, such as to liberty, is merely another way of saying that we (moral agents) have a direct duty not to interfere with her exercise of free movement, free choice, and so on. A narrow interpretation demands something more stringent: Rights provide a foundation for only the more basic obligations, obligations that are much harder legitimately to override.

The basic theoretical difference (admittedly a difference in degree rather than kind) is that "ordinary" duties can and must be balanced and assessed against all sorts of competing demands, interests, and inclinations. We as a community have duties to respect residents' use of their private property, but those duties can and often are assessed against, and sometimes overridden by, other wants and desires: hence zoning restrictions and some environmental legislation. Rights, on the other hand, are not so easily overridden: A person's right to free speech cannot legitimately be overridden no matter how many people dislike what is being said. Rights in this narrow sense are often said to be "basic," "inviolable," "natural," or "inalienable." The justification for this narrower, but stronger, sense of rights is analogous to that in the political arena, in which rights are seen as a way of protecting the minority from the potential abuses of majority rule.

It might be interesting to note in passing that the only legitimate way of characterizing Peter Singer — the author of *Animal Liberation* and the so-called father of the animal rights movement — as "a defender of animal rights" would be to rely on the broadest possible interpretation of rights-claims. All interests, costs, and benefits are to be weighed equally, and none are set aside for special protection. Joel Feinberg (44) and James Rachels (45) have also defended broad versions of an "animal rights thesis." Tom Regan, as we have seen (8), defends a more stringent and radical rights-based position on which animals have rights in the narrow sense.

In using a broad concept of rights, when people say "animals have a right to be treated humanely," they usually mean only that it is morally wrong to treat them inhumanely, and it is wrong because of the harm to the animal, not just the indirect harm caused to other humans. Understood this way, the position is eminently reasonable, and not necessarily based on any confusion, misunderstanding, or radical propaganda. Accordingly it does not represent a position to be opposed but rather a welcome opportunity for dialogue and better understanding.

When considering the claim that attributing rights to animals is "eminently reasonable," it is important to keep in mind that neither the broad nor the narrow concept of rights *entails* any assumption of equality between animals and humans. Consider, first, the narrow account of rights: Rights bestow a special sort of protection that cannot be overridden by appeals to a greater general good. There is nothing inconsistent in claiming that an obligation to allow an individual to express her political views cannot be overridden (i.e., she has a right to free speech), but her license to vote or drive a car may be revoked or denied for the greater good. We routinely deny both privileges to children on the ground that their immaturity would render their driving or voting harmful to society. Similarly one can, without any inconsistency, argue that animals have the right not to be tortured without thereby being committed to the claim that they have the right not to be killed. Even when the same right is ascribed to two different individuals, narrow views must and do recognize that one such special claim may be stronger than another, or that if only one can be respected, one has an objectively stronger status. Thus, two people may both have the right to inherit someone's estate — a right that cannot be overridden by the fact that more good might be done by distributing the wealth to agencies that would further the public good — and we can still, in many cases, decide that one right is stronger than another. Similarly, even if one argues that both a dog and a human have the right to life, one might legitimately conclude that the human's right is the stronger of the two, if one is in a position where only one can be respected.

The broad view of rights is even clearer on the issue of equality. Even a moment's thought will uncover a wealth of examples in which it is wrong to treat one individual in a certain way and perfectly legitimate to treat another in exactly that way. Since the broad understanding of rights-talk would automatically translate such differential judgments into different rights-claims, ascribing some rights to animals cannot possibly entail that they must have all the same rights as humans, or that the rights they do share with humans have equal weight.

What follows from all of this? Since there is often confusion about the meaning of rights-claims, and since concerns for "animal welfare" overlap significantly with at least some interpretations, perhaps these divisive labels ought to be retired in favor of a more precise statement of what is actually being claimed. If this suggestion is too extreme, at least one ought to be careful to interpret them in a more flexible and open-minded way, rather than to polarize the debate.

ANIMALS IN RESEARCH

There are, of course, some sustained defenses of practically all animal research, without restrictions or qualifications

(4,17,26,46,47). Some are quite thoughtful, but others convey the impression of defensiveness. One litmus test is whether such defenses agree that there is room for improvement, and that not all animal research meets the highest ethical standards. Contrariwise, attacks on animal research that claim that no significant gains have been achieved through animal research, or that all such research could readily be replaced by alternatives such as computer models and in vitro testing, also undermine their own credibility. There are, fortunately, well-reasoned and detailed discussions of all aspects of the debate (48–51), carefully stated and well-documented arguments against animal research (52), other works that concentrate on specific controversies such as research involving primates (53,54), useful literature surveys (55,56), and anthologies that try to present a varied selection of views on animal research in particular (57,58) or broader philosophical debates that have direct implications for research (59).

Current Regulatory Structure

The subject of current regulations is complex, and constantly changing. In addition to official documents such as the National Institutes of Health (NIH) Guide (60) and U.S. Department of Agriculture (USDA) regulations, there are more helpful and detailed studies of these regulations (50,61) as well as the reference library maintained by Animal Welfare Information Center (AWIC) and numerous on-line sites, so what follows is only meant as a quick overview.

Since 1985, any institution which receives federal funding and uses vertebrate animals is subject to NIH regulations, as set forth in the *Guide to the Use and Care of Animals*. Institutions that use mammals other than mice, rats, or common agricultural species are subject to USDA regulations. The strictest level of control is the voluntary Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International accreditation. To a large extent, these regulations and guidelines overlap one another, and in some cases even reference one another.

A key feature of current regulations is the demand that each institution covered by the regulations must establish an Institutional Care and Use Committee (IACUC). This committee must have at least five members, including a veterinarian, a "nonscientist" (defined more precisely as someone who does not engage in animal research), and someone who is not affiliated with the institution. This committee is charged with reviewing protocols for all research involving covered species, conducting semiannual reviews of all animal facilities, ensuring that the institution is in compliance with applicable regulations and guidelines, and generally providing mechanisms for the monitoring and control of animal research. Institutions vary widely with regard to the way in which these duties are carried out: how large the committee is, how members are selected, whether reviews are done by the entire committee or a subgroup, the type of information and documentation required from investigators, how much of the meeting is open to the public, and so on (62). USDA regulations also provide for the licensing of dealers in research animals, both those who breed animals specifically for research — "Class A dealers" — and those who buy and sell animals, typically dogs — "Class B dealers." One of the goals of this licensing process is to answer the public's concerns, fanned by an article in *Life* magazine in 1966, about pets being stolen to be sold to research labs.

In addition to IACUC oversight, covered institutions are subject to regular, unannounced inspections by the USDA. Although these inspections are supposed to take place twice a year, staffing and funding shortages make this more of an ideal than a reality, especially at smaller institutions. Facilities with AAALAC accreditation are also inspected regularly by an independent group of reviewers, and NIH has the right to conduct its own independent inspections, if it chooses to do so. One shortcoming of all these formal reviews is that while they can inspect facilities, animals being held for research, records, and IACUC minutes, they are rarely in a position to monitor ongoing research directly. Thus, it is difficult to guarantee that approved protocols, and only such protocols, are being adhered to. While it is reasonable to believe that the majority of responsible scientists do follow the protocols for which they have received approval, abuses do occur.

The regulatory system just described applies to animal research in the United States; the Canadian system is quite similar. Elsewhere in the world, regulation of animal research ranges from stricter controls than exist in the United States to nonexistent (63–65).

The Three R's

It has become commonplace in justifying animal research to allude to the "three R's," originally formulated by Russell and Burch (66): reduce, replace, and refine.

"Reduction" refers to reducing the numbers of animals used, consistent with obtaining significant results. The latter is an important qualification, since reducing the number of animals too far might render the results statistically questionable and hence a total waste of animals; indeed, IACUCs sometimes find themselves recommending an increase in numbers for just this reason. On the other hand, the goal of reduction could be significantly furthered if the scientific community and research journals were to rethink their definitions of significance. Particularly for preliminary research, a significance level of 90 percent rather than the typical 95 percent would dramatically reduce the number of animals needed for any given test (67).

"Replacement" refers to using nonanimal models, dead animals, or "lower" species whenever possible. In vitro studies are commonly cited as an alternative, but currently tend to be most practical for initial screening of variations on known compounds. Similarly, computer simulation is valuable only when we have enough information to construct an accurate computer model, and thus may be more useful for education than for exploratory research (68,69). The criteria for ranking animals in terms of higher/lower also needs closer scrutiny; it can often mask cultural preferences rather than any objective standard. Thus dogs are often deemed "higher" than pigs, and NIH singles them out, along with cats, primates, and

endangered species as worthy of special status, but there is no clear physiological, evolutionary, or psychological evidence for this distinction. In short, while the injunction to "replace" might be a useful maxim, its application is quite problematic.

Thus, while reference to the three R's has become almost obligatory, it is not at all clear that they still provide useful guidance. As just noted, there is much dispute about the practical applicability of the injunctions in specific protocols, and a general lack of theoretical clarity (70). More work on this topic by both scientists and ethicists is clearly needed.

Teaching and Testing

The strongest arguments in favor of research on animals point to the furthering of basic scientific knowledge or biomedical advances. These arguments are much less convincing when applied to the use of animals in teaching and testing (71). The dividing line between any two of these areas may be fuzzy. Is the lab work of a first-year graduate student teaching or research? When in the process of developing a new drug do we switch from research to testing? Nonetheless, looking at paradigm cases of teaching and testing is necessary in order to have a complete picture of the use of research animals.

When animals are used in the classroom, they are typically described as serving one or more of three purposes: (*1*) as an illustration of a process, event, or state; (*2*) as part of a project designed to help students learn proper research design and practice; (*3*) to allow students (e.g., veterinary students) to learn proper surgical or other techniques.

"Illustration" can range from observation without intervention — keeping a gerbil in an elementary school classroom, or watching a tadpole develop into a frog — to demonstrations of acute medical or pharmacological emergencies — infecting a dog with distemper so that veterinary students can observe the progression of the disease, or dosing a rat with cyanide so that pharmacy students can see the symptoms, and the efficacy of various antidotes. Dissection, particularly at the secondary school level, often falls into this category, as does the use of animals in most science fair projects. Those who use such illustrations often defend them by claiming that a live demonstration or actual dissection is a more vivid and effective teaching tool than a textbook illustration or other alternative; it engages the students more. However, the results of those procedures which are invasive and often painful are almost always known ahead of time, unlike the case of basic research. Videotapes, computer programs, and textbook illustrations can all contain the same information. In the face of the suffering and death that more invasive illustrations cause, appeals to students' interest or attention spans seem trivial. Moreover, such demonstrations can have, either deliberately or inadvertently, a desensitizing effect, conveying to students that animals are mere research tools whose suffering should not overly concern us. This desensitization can develop even in the apparently benign case of having an animal in an elementary school classroom. If children are allowed to handle, observe, or otherwise interact with animals without due understanding of the stress this might cause, or if the issue of animal care over weekends and vacations is trivialized, they may easily pick up a casual, noncaring or nonrespectful attitude toward animals.

As noted above, the dividing line between basic research and testing can easily become blurred. Testing is also easy to trivialize: It is easier to dismiss the claimed need to test the safety of a new mascara than the development of a new treatment for stroke victims. However, many of the basic factors determining toxicity, carcinogenicity, and inflammation are well enough understood that in vitro studies (and, in a few cases, computer modeling) can be substituted. The Johns Hopkins Institute for Alternatives maintains current data on such substitutes. Even when these alternatives are not conclusive — for example, because they fail to fail to detect effects at an organic rather than a cellular level — they can be useful for initial screening. The most infamous tests, including the Draize test and the LD-50, are gradually being replaced by alternatives that use no living animals, or fewer animals in a less invasive way. The Draize test involves inserting the substance to be tested for irritancy into the eye of a rabbit who has been immobilized in a "stock," and then observing changes in the eye over a period of days. Various in vitro tests, or the use of chicken eggs, have often served as reliable alternatives. The use of computers for more sophisticated statistical analysis has allowed researchers to replace the crude LD-50 test, in which increasing doses of a material were given to colonies of mice or other animals until one found the level at which 50 percent of the animals died, with other tests that required fewer animals and did not always use death as an endpoint. Despite the availability of these advances and alternatives, animal testing remains an area in which many advances in animal welfare are still possible.

Continuing Issues

Of course the fundamental continuing issue is when, if at all, animals should be used in research. While many of the relevant arguments have been discussed in the previous section on theory, it will be useful to see how they apply specifically to animal research. Other continuing issues focus more specifically on current regulations and the general contemporary research environment: how well do they provide appropriate protection for the animals (72)?

Another continuing issue reflects society's (including many scientists') demand for a further increase in our moral sensitivity in animal research. The most obvious trend in social pressures is reflected in the growing popularity of animal protection groups, ranging from the radical People for the Ethical Treatment of Animals (PETA) through the more moderate organizations such as the American Society for the Prevention of Cruelty to Animals (ASPCA) or Working for Animals Used in Research, Drugs, and Surgeon (WARDS) whose aim is not total abolition of animal research. More precisely, the moderate groups hold that the use of animals ought to be abandoned when and where it is possible to do so (would any researcher disagree?), but they are more likely to agree with the scientific community about the

fact that the range of productive alternatives is today quite limited, thus accepting that animal research must continue for the foreseeable future. In terms of the three R's, moderates tend to see "refine" and "reduce" as more effective immediate options than "replace" — although they are likely to encourage further research on the development of alternatives. In terms of our categories of views about the moral status of animals, radical groups are generally liberationists while moderate groups at least implicitly adopt a differential perspective. The existence of this range of views within what is sometimes called the protectionist movement again illustrates the danger of the "animal rights/animal welfare" dichotomy discussed earlier. If the scientific community insists on viewing all protectionist groups as radical "animal rights people," opportunities for fruitful dialogue and identification of common ground can be missed.

When fruitful dialogue is possible, and common ground is identified, another trend that is just beginning may blossom. This refers to the increased efforts among scientists and regulators to address ethical issues explicitly and directly, with attention to the general moral principles which underlie our decisions about research on animals can and should be conducted (73,74). The American Association for Laboratory Animals (AALAS), the American Veterinary Medical Association (AVMA), the Animal Behavior Society, and the Scientists Center for Animal Welfare are but a few of the professional societies that have included sessions on ethics at their national conferences. This is in marked contrast to earlier, more polarized efforts in which scientists — sometimes with little background in ethical theory — took it upon themselves to demonstrate exactly what was wrong with animal rights, objections to research on animals, or Regan or Singer. By contrast, the efforts mentioned above represent a collaborative effort to formulate and evaluate the various ethical theories and principles that shape (well or badly) specific choices and regulations about animal research. Such collaborative efforts can give us a more solid grounding from which to address some of the more vexing specific questions about the current state of animal research and its evaluation: Should dogs, cats, or primates be singled out for special protection (as they are today), and if so, why? Should rats, mice, birds, or agricultural animals be excluded from USDA regulations, and why or why not? To what extent must considerations of scientific merit be blended with ethical issues, including but not limited to, IACUC reviews and journal publication criteria? After all, one cannot do a cost–benefit analysis of the sort required by NIH and USDA without some consideration of potential benefits (75).

On the Horizon

Crystal balls are notoriously unreliable, but some future trends are already apparent. Two major social influences on the future use of animals in research are apparent, and they pull in opposite directions. As indicated in the previous section, increased public interest in the use of animals will almost certainly demand that the scientific community continue searching for alternatives to whole animal models. At the same time an increased interest in maintaining good health, especially in an aging population, drives demands for more research on diseases and aging. This apparent inconsistency provides scientists with the opportunity to drive home an essential message: Given the current state of biomedical research, studies will require the use of animals, but such studies can and will be done with ever-increasing sensitivity to the welfare of the animals used.

The largest unknown quantity is the rapidly increasing power and impact of biotechnology and genetic engineering on animal research. Currently the most obvious effect is the ability to develop animal models to aid in the study of human diseases such as cystic fibrosis, Lesch-Nyham's disease, and some forms of cancer (OncoMouse) (76). This raises two sets of issues: The first is simply a new variant of an old problem, and the other poses new problems for researchers, IACUCs, and regulators. The old issue is that of deliberately producing animals with health problems that can sometimes be chronic, debilitating, and painful. While genetic engineering has made it possible for us to produce animals with some new diseases, old-fashioned selective breeding has long been used to produce animals with other equally serious problems. The new issues are that genetic engineering can result in animals that (*1*) have unpredicted, perhaps unpredictable, health and welfare challenges, (*2*) are often considered much more valuable than "standard" animals of the same species, and (*3*) heighten public fears about scientists "playing God" or tampering too much with the "natural" order of things (77).

The variation of the old problem — caring properly for animals that have, or are expected to develop, severe health problems — is related to genetic engineering only insofar as biotechnology affords the possibility of producing animals with diseases hitherto unknown in a given species. This poses important ethical problems for both researchers and IACUCs. The first is when it is justifiable to produce such animals at all, rather than using alternative forms of research such as epidemiological studies of naturally occurring incidents of the disease. Second, one must consider whether the effect on the animals' welfare will require special care — such as different housing or diet, more frequent monitoring by caretakers or veterinarians, or a prescribed regimen of analgesics. While standard procedures are designed to protect an adequate level of welfare for normal animals, they can be inappropriate for animals specifically bred to develop a chronic health problem. This poses a special burden on IACUCs and responsible veterinarians, who must be informed enough about special needs to ensure that those needs are met. Finally, one must consider the problem of identifying an ethically acceptable endpoint. Using death as an endpoint for a procedure, or even waiting until an animal becomes moribund before terminating the procedure and euthanizing the animal, always requires careful scrutiny and detailed justification, but when the research involves chronic, severe, or terminal health problems the issue of determining the appropriate endpoint, and who is responsible for making the necessary assessments, must be fully addressed at the outset.

As noted earlier, the moral questions just described are not unique to genetically engineered animals, even though

biotechnology increases the ability to create animals who are likely to develop a targeted health problem. Other difficulties are magnified even more by advances in biotechnology. What follows are a few examples of problems that have been anecdotally identified; there is a definite need for more carefully controlled studies in this area.

One often-noted difference between genetically engineered animals and those produced through traditional methods of selective breeding is that genetically engineered animals often tend to develop unanticipated phenotypic effects, not obviously related to the desired effect. Some of these effects, ranging from severe joint problems in the "Beltsville pigs" to lack of maternal instinct in mice bred as models for Lesch-Nyham's disease, have a direct impact on animal welfare and ethical concerns about whether such animals should be produced. Second, genetically engineered animals are viewed as more valuable than standard laboratory animals, if only from the standpoint of production costs; this can have both a positive and a negative impact on ethical deliberations. On the positive side, more valuable animals are likely to get more intensive care. On the negative side, researchers with a considerable investment in genetically engineered animals may be more reluctant to terminate a study or euthanize an animal on the basis of welfare concerns that do not coincide with research goals. Once again, this poses ethical challenges for both the researcher and the IACUC that must evaluate and monitor the protocol.

The final concern about genetic engineering, usually expressed as "playing God," is unfortunately too often not very well articulated, which makes the ethical concerns hard to assess. However, it is often likely to focus on research that involves introducing human genetic material or patterns into animals. When such research is specifically targeted — such as aimed at developing pigs whose organs can be used for transplantation into humans, or goats whose milk contains hormones useful for treating human diseases — such concerns seem less apparent. Whether or not "playing God" represents a legitimate ethical objection to genetic engineering, it is surely an area in which the research community must improve communication with the public.

BIBLIOGRAPHY

1. R.J. Hoage, ed., *Perception of Animals in American Culture*, Smithsonian Institute Press, Washington, DC, 1989.
2. G. Carson, *Men, Beasts, and Gods*, Scribner's, New York, 1972.
3. T. Regan and P. Singer, eds., *Animal Rights and Human Obligations*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1989.
4. R. White, in T. Regan and P. Singer, eds., *Animal Rights and Human Obligations*, Prentice-Hall, Englewood Cliffs, NJ, 1976, pp. 163–169.
5. R. White, *Hastings Center Rep.* **20**, 43 (1990).
6. P. Carruthers, *The Animals Issue*, Cambridge University Press, New York, 1992.
7. P. Singer, *Animal Liberation*, 2nd ed., Avon, New York, 1990.
8. T. Regan, *The Case for Animal Rights*, University of California Press, Berkeley, 1983.
9. B. Rollin, *Animal Rights and Human Morality*, Prometheus Books, Buffalo, NY, 1981.
10. S. Curtis, in T. Regan and P. Singer, eds., *Animal Rights and Human Obligation*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1989, pp. 169–175.
11. S. Sapontzis, *Morals, Reason, and Animals*, Temple University Press, Philadelphia, PA, 1987.
12. E. Pluhar, *Beyond Prejudice: The Moral Significance of Human and Nonhuman Animals*, Duke University Press, Durham, NC, 1995.
13. D. DeGrazia, *Taking Animals Seriously: Mental Life and Moral Status*, Cambridge University Press, New York, 1996.
14. P. Singer, *Practical Ethics*, Cambridge University Press, New York, 1979.
15. P. Singer, *Monist* **70**(1), 3–14 (1987).
16. R.G. Frey and W. Paton, in T. Regan and P. Singer, eds., *Animal Rights and Human Obligation*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1989, pp. 223–236.
17. M.A. Fox, *The Case for Animal Experimentation*, University of California Press, Berkeley, 1986.
18. I. Kant, reprinted in T. Regan and P. Singer, eds., *Animal Rights and Human Obligations*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1989, pp. 23–24.
19. J.M. Jasper and D. Nelkin, *The Animal Rights Crusade: The Growth of a Moral Protest*, Free Press, New York, 1992.
20. S.M. Russell and C.S. Nicoll, *Proc. Society for Experimental Biology and Medicine*, 1996, pp. 109–138.
21. R. Ryder, *Victims of Science*, David-Poynter, London, 1975.
22. M. Wreen, *Ethics and Animals* **5**, 47–60 (1984).
23. R.G. Frey, *Between the Species* **3**, 191–201 (1987).
24. J.U. Dennis, *J. Am. Vet. Med. Assoc.* **210**(5), 612–618 (1997).
25. J. Stone, *Canadian J. Philos.* **17**, 815–830 (1987).
26. W. Paton, *Man and Mouse*, 2nd ed., Oxford University Press, Oxford, UK, 1993.
27. C. Cohen, *N. Engl. J. Med.* **315**, 865–870 (1986).
28. J. Nelson, *Between the Species* **2**, 116–134 (1986).
29. J. Rachels, *Created from Animals*, Oxford University Press, Oxford, UK, 1990.
30. J. Bentham, *The Principles of Morals and Legislation*, ch. 17, sec. 1.
31. B. Rollin, *The Frankenstein Syndrome*, Cambridge University Press, New York, 1995.
32. B. Rollin, *Farm Animal Welfare*, Iowa State University Press, Ames, IA, 1995.
33. L.-M. Russow, *Between the Species* **8**, 224–229 (1993).
34. L.-M. Russow, *The Question of Animal Rights*, forthcoming.
35. R. Sorabji, *Animal Minds and Human Morals: The Origins of the Western Debate*, Cornell University Press, Ithaca, NY, 1993.
36. J. Rawls, *A Theory of Justice*, Harvard University Press, Cambridge, MA, 1971.
37. J. Narveson, *Canadian J. Philos.* **7**, 161–178 (1977).
38. J. Narveson, *Monist* **70**, 31–49 (1987).
39. M. Midgeley, *Animals and Why They Matter*, University of Georgia Press, Athens, GA, 1983.
40. A. Baier, in H. Miller and W. Williams, eds., *Ethics and Animals*, Humana, Clifton, NJ, 1983, pp. 61–77.
41. L. Birke, *Feminism, Animals, and Science*, Open University Press, Bristol, PA, 1994.

42. J. Tannenbaum, *Veterinary Ethics*, 2nd ed., Mosby, St. Louis, MO, 1995.
43. G. Varner, *Hastings Center Rep.* **24**, 24–28 (1994).
44. J. Feinberg, in J. Feinberg, ed., *Rights, Justice, and the Grounds of Liberty*, Princeton University Press, Princeton, 1980, pp. 185–206.
45. J. Rachels, in T. Regan and P. Singer, eds., *Animal Rights and Human Obligation*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1989, pp. 122–131.
46. F.K. Goodwin, *Contemporary Topics in Laboratory Animal Science* **31**, 6–11 (1992).
47. F.K. Goodwin and A.R. Morrison, *Scientist* **17**, 12 (1993).
48. A. Rowan, *Of Mice, Models, and Men*, SUNY Press, Albany, NY, 1984.
49. J. Tannenbaum and A. Rowan, *Hastings Center Rep.* **25**, 32–43 (1985).
50. B. Orlans, *In the Name of Science*, Oxford University Press, New York, 1993.
51. J. Smith and K. Boyd, eds., *Lives in the Balance: The Ethics of Using Animals in Biomedical Research*, Oxford University Press, New York, 1991.
52. H. LaFollette and N. Shanks, *Brute Science*, Routledge, New York, 1996.
53. D. Blum, *The Monkey Wars*, Oxford University Press, New York, 1994.
54. P. Cavalieri and P. Singer, *The Great Ape Project*, St. Martin's, New York, 1993.
55. C. Magel, *A Bibliography of Animal Rights and Related Matters*, University Presses of America, Latham, MD, 1981.
56. J.L. Nelson, *Am. Philos. Q.* **22**, 13–24 (1985).
57. R. Baird and S. Rosenblum, eds., *Animal Experimentation: The Moral Issues*, Prometheus Books, Buffalo, NY, 1991.
58. G. Langley, ed., *Animal Experimentation*, Chapman and Hall, New York, 1989.
59. H. Miller and W. Williams, eds., *Ethics and Animals*, Humana Press, Clifton, NJ, 1983.
60. National Research Council, *Guide for the Care and Use of Laboratory Animals*, National Academy Press, Washington, DC, 1996.
61. M. Phillips and J. Sechzer, *Animal Research and Ethical Conflict*, Springer Verlag, New York, 1989.
62. K. Bayne et al., *Current Issues and New Frontiers in Animal Research*, Scientists Center for Animal Welfare, Greenbelt, MD, 1995.
63. Available at: *http://www.uel.ac.uk/research/ebra/info.html*
64. Available at: *http://www.labanimal.com/col/ebra3.html*
65. D.E. Blackman, P.N. Humphreys, and P. Todd, *Animal Welfare and the Law*, Cambridge University Press, Cambridge, UK, 1994.
66. W.S. Russell and R.L. Burch, *The Principles of Humane Experimental Technique*, Methane & Co, London, 1959.
67. M.D. Mann, D.A. Crouse, and E.D. Prentice, *Lab. Animal Med.* **41**, 6–14 (1991).
68. D. Pratt, *Alternatives to Pain in Experiments on Animals*, Argus Archives, Washington, DC, 1980.
69. D.M. Foster and R.C. Boston, *ILAR J.* **38**, 58–89 (1997).
70. M. Mukerjee, *Sci. Am.*, 86–93, February (1997).
71. A. Smith et al., *ILAR J.* **38**, 82–88 (1997).
72. L.-M. Russow, *ILAR J.* **40**, 15–21 (1999).
73. *Chimpanzees in Research: Strategies for Their Ethical Care, Management, and Use*, National Academy Press, Washington, DC, 1997.
74. NASA Principles on the Use of Animals; *NASA Policy Directive for Animal Care and USE*, forthcoming, 3/98. Also available at: *http://www.nih.gov/grants/oprr/dearcolleage.htm*
75. B. Orlans et al., *The Human Use of Animals*, Oxford University Press, New York, 1998.
76. S. Donnelley, C. McCarthy, and R. Singleton, *Hastings Center Rep.* **24**(special supp.), S1–S32 (1994).
77. L.R. Batra and W. Klassen, *Public Perceptions of Biotechnology*, Agricultural Research Institute, Bethesda, MD, 1987.

See other entries ANIMAL, MEDICAL BIOTECHNOLOGY, LEGAL, LAWS AND REGULATIONS GOVERNING ANIMALS AS SOURCES OF HUMAN ORGANS; ANIMAL, MEDICAL BIOTECHNOLOGY, POLICY, WOULD TRANSGENIC ANIMALS SOLVE THE ORGAN SHORTAGE PROBLEM?; see also RESEARCH ON ANIMALS entries; TRANSGENIC ANIMALS: AN OVERVIEW.

RESEARCH ON ANIMALS, ETHICS, PRINCIPLES GOVERNING RESEARCH ON ANIMALS

JOHN KLEINIG
City University of New York
New York, New York

OUTLINE

INTRODUCTION

Research using live animals goes back at least as far as Galen, though it was not until the rise of medicoscientific experimentation in the seventeenth and eighteenth centuries that the use of live animals in research became systematic (1). In the past half century the increasing professionalization of research and practice and the growth of product-safety regulation have led to their prevalent use in education and product testing.

From its very beginning the practice of live animal research (or, now inaccurately, vivisection) has attracted controversy, even among those who have defended it. At one end of a wide spectrum are those who have wished to apply to animals the whole range of moral entitlements and protections that govern experimentation with humans. At the other end are those who would ascribe to an animal's cries no more moral significance than one would to the noise made by a squeaky door. Informing this vast range are metaphysical, epistemological, and moral

positions that are too complex and controversial to be resolved in the space of a single article.

Most who give the matter consideration would now acknowledge that nonhuman animals make some moral claims on us, even though there are serious disagreements over the nature and extent of those claims (2–4). Should we eat animal flesh? Does the domestication, transgenic breeding, and cloning of animals interfere with their natural integrity? Is hunting morally acceptable? But nowhere are the moral questions more troubling than in the domain of scientific research, experimentation, and testing. For here important values often seem to be in tension, and no simple formula for their resolution appears to be available (5–8). Whereas, in the case of humans, we can at least constrain our endeavors by appealing to the informed consent of research subjects, this is not a serious option in the case of nonhuman animals.

Furthermore, producing ethical guidelines for animal experimentation is not like producing guidelines for the withholding or withdrawal of lifesaving treatment from human beings. Complex though the latter is, there is sufficient congruence of both ends and means to allow the formulation of a reasonably explicit and manageable structure of benchmarks and stipulations for the guidance of those with whom such decisions will lie (9). In the case of animal experimentation, however, such is the plurality of ends and means, not to mention the diversity of subjects, that any general guidelines are likely to be unhelpfully vague or very limited in their application. Participant observation of animals in the wild, genetic experimentation with *Drosophila melanogaster*, and toxicity testing using laboratory-bred rats differ so much in their character and in the specific ethical questions they raise, that no single set of guidelines is likely to be fully responsive to the ethical issues that should be addressed. This is not to gainsay the value of the guidelines that have been produced in recent years (10–12), but it identifies one source of the dissatisfaction they have engendered. At the very least it mandates a plurality of guidelines.

MEANS AND ENDS

The alternative to comprehensive and exhaustive guidelines need not be an absence of moral moorings or of a structure of moral questions that would allow the ordered assessment of animal experimentation and its associated practices. The inadequacies of guidelines need not bespeak the absence of guidance. Indeed, animal experimentation of whatever kind has the form of a means–end relation, and there are established procedures for assessing such relationships. True, there are some very basic questions concerning the moral status of animals on which we are culturally confused and for which there exists no theoretical moral consensus. And these, along with more general theoretical controversies in ethical theory, will continue to bedevil our deliberations. Nevertheless, an ordered structure for identifying and responding to these issues is available and may be articulated as a framework for both primary ethical deliberation and the secondary development of specialized codified guidelines.

To be ethically acceptable, practices that involve a means-end relation must address and perform satisfactorily in relation to each of the following five questions: (*1*) Is the end morally acceptable? (*2*) Are the means appropriate to the end? (*3*) Are the means likely to realize the end? (*4*) Are the means proportionate to the end? and (*5*) Will the means undermine other good ends? Question 1 recognizes that ends, no less than means, need to be scrutinized. Question 2 acknowledges that what may be employed as means may be inappropriate to the end. Question 3 assumes appropriateness, but focuses on the probability that the end will be established. Question 4 allows that the means may establish the end, but it addresses the costs of achieving it. Question 5 concerns itself with external costs — those wider social costs that an institution or practice may have. Although — as will be clear from the discussion that follows — this menu of questions does not represent a simple checklist, in which the questions can be separately considered and okayed, it nevertheless provides a broadly comprehensive and ordered framework for the organizing of moral deliberation. The commentary to follow, though sketchy, will indicate how these questions can function like a main software menu to systematize ethical reflection on animal experimentation. Each menu item then needs to be accompanied by a further set of pull-down menus. We can order even if not simplify the complexity.

Is the End Morally Acceptable?

Research and experimentation are teleologically oriented activities. A subject is investigated or manipulated in the light of some end. Frequently that end is appealed to as a justifying consideration. Although it is commonly — and properly — asserted that "the end does not justify the means," it does not follow that the end is morally irrelevant to the means. The aphoristic prejudice against *any* justificatory appeal to ends tends to reflect the great significance we attach to autonomy or consent in dealings between human beings, and the general presumption against treating them instrumentally or paternalistically. Even so, that should not obscure the fact that an assessment of ends, and human ends in particular, constitutes a relevant and often important determining factor in the appraisal of human practices. Indeed, as the flip side of our concern that humans not be used merely as means, a focus on ends reflects the importance that we ascribe to them as expressions of human purposive activity.

Animal welfare committees, unlike human investigational review boards, will be guided for the most part by considerations of beneficence, largely unmediated by considerations of justice and autonomy, and in their case the evaluation of ends will assume greater significance. To assert this is not to deny that research animals may have ends of their own, or that their ends may not be genuinely competitive with human ends (see question 4), but it acknowledges the considerable justificatory weight that attaches to human ends as human (though not necessarily anthropocentric) ends. If we undervalue such ends, we erode the significance of the very enterprise that allows our questioning of those ends to carry weight.

Even so, we must make important distinctions. Animal experimentation may be directed to a variety of potentially legitimizing ends — the welfare of particular animals or animal populations or species, the welfare of individual human beings or some wider social good, the communication and advancement of human knowledge, organizational profit, some personal benefit to the researcher, and so on. These ends are not exclusive, and they may be given a different value and priority by different researchers, even by researchers engaged in the same series of procedures. And, of course, *we* (the politically potent community) may value some more highly than others. Even if human ends, as human ends, possess an intrinsic value, not all human ends are equally valuable, and some we may think unworthy of human advancement.

Although disagreements about the relative value of ends can make consensus difficult to achieve, we are not left wholly at the mercy or to the vagaries of individual preference. Relevant differentia can be articulated and brought to bear on judgments of priority. For example — though these do not constitute decisive (i.e., lexically ordered) considerations — experimental procedures directed to welfare will generally have better standing than those carried out solely to satisfy our curiosity or to expand or communicate our knowledge; public goods will generally take precedence over private benefits; procedures designed to benefit the subject of those procedures will generally have a stronger claim to our recognition than those designed to benefit others; and human welfare will generally take priority over animal welfare. This is because there is value to benevolence, a communal dimension to value, a special dignity to human life, and an integrity to animal lives that should weigh significantly in our decision making. But these considerations may exist in tension, and it is not possible to read off priorities in a mechanical fashion. Judgments that seek to accommodate them are singular without being arbitrary, and the problem of formalizing them is not peculiar to animal experimentation but reflective of more general problems in the appraisal of human conduct.

Some of the difficulties in making judgments about the relative importance of ends are linked to complexities in the ends themselves. The advancement of human knowledge, for example, may comprehend the satisfaction of curiosity, the exercise and development of our human powers of understanding, an increase in our grasp of the universe and of ourselves within it, and what is sometimes termed "basic research," which, though not directed to some specific application, usually anticipates some later — albeit unspecifiable — usefulness. And where some instrumental benefiting of ourselves is sought, it may be the alleviation or cure of some disease–mild or serious, rare or common, self-induced or unwittingly contracted — some positively enhancing, recreational, labor-saving, or aesthetic end, or some preventive or protective social goal. The possibilities are legion and jointly pursuable, and animal researchers will need to give some detailed account of them. Guidelines that differentiate and categorize ends may assist in this task, even though judgments about their relative importance will require the more sensitive deliberative scrutiny that a review committee may be able to provide. Even then, there is a serious practical problem posed by the fact that the interests of animals can be represented only by proxies who may not be sufficiently sensitive to them.

The fact that different researchers in a single project may have different priorities and may even be in pursuit of different ends highlights one of the difficulties involved in the evaluation of practices independently of their practitioners, and in the development of social policy. Otherwise, justifiable projects may sometimes be compromised by the unworthy, questionable, or only moderately worthy ends of those who engage in them. Dissertation research, the testing of commercially redundant substances whose marketing is designed only to give a company increased profitability, and the development of biological weaponry can represent the compromise or perversion of otherwise justifiable research agendae. Such ends may not justify the moral costs they involve. Unfortunately, we are not usually able to peer into the hidden motives of researchers to determine whether their private motives match their publicly asserted intentions. And social policies can achieve little more than a monitoring of formally stated ends.

As noted earlier, ends, though relevant to the justification of means, are not usually sufficient to justify them. The appropriateness, efficacy, proportionality, and character of the means also need to be taken into account. Means have a "life" of their own which needs to be considered, not just as they are associated with particular ends.

Are the Means Appropriate to the End?

However worthy ends may be, unless the means used to further them are appropriate to their realization, they will fail to provide the justificatory support expected of them. This is particularly important for the use of animals in scientific research, which imposes exacting demands on experimenters if their results are to be valid and reliable.

Experimental design is often viewed only in relation to scientific validity, and the determination of scientific validity is frequently thought to be independent of moral considerations. That, however, is an oversimplification. For one thing, unless an experimental procedure is suited to the realization of scientifically valid results, it will be inappropriate to the ends to which it is supposedly directed, and any justificatory value those ends might have possessed will be forgone. Moral costs will remain uncompensated. Moreover, and more fundamentally, decisions about the level of significance that will be required of experiments involving animals need to take into account the importance of the ends to which the experiments are directed, and the moral costs in suffering or other deprivations that will be caused. The decision to require a p value of .01 rather than, say, .05, has implications for sample size, the level of control that is exercised over variables, and so on, decisions that inevitably confront the researcher with the costs that will be involved in his or her inquiry. In other words, what appears to be only a matter of scientific validity will also involve issues of moral acceptability.

One of the common complaints about experimental procedures involving animals has been that researchers treat their animal subjects in ways that compromise the validity of their results. Rough handling, poor housing, and generally inadequate control over extraneous variables may jeopardize experimental integrity and give rise to misleading or worthless results. Animal life is wasted, needless suffering may be caused, and scarce resources are squandered. The training of those who are to handle experimental animals must encompass not only issues of technique but also sensitivity to the ecology of animal lives. Much of the current concern over experimental conditions was generated as a result of a 1984 raid by the Animal Liberation Front on the Regional Head Injury Center of the University of Pennsylvania, and the subsequent circulation of stolen videotapes by People for the Ethical Treatment of Animals as *Unnecessary Fuss*. Here, worthwhile ends were needlessly compromised by the careless and callous attitudes of researchers and their assistants (13).

But validity is not affected only by careless treatment of the animal subjects. Sample size, species selection, and other elements of research design may also have an important bearing on the credence to be given and conclusions to be drawn from experimental results. The debate over using pound as against purpose-bred animals turned in part on the extent to which the use of one rather than the other would introduce uncontroled variables. And where animal research is intended to have implications for other species or for human beings, there needs to be some assurance that the experimental subjects are similar in relevant respects.

This latter concern is particularly intractable. The evolutionary theory that may seem to allow for some continuity between humans and animal species, such that research results gained from one can be applied to the other, can also cut the other way and be used to suggest that there now exist fundamental discontinuities and that the validity of using animal research for human welfare ends is problematic. Although I think these difficulties have been exaggerated, they are not without force (7). Moreover, how high the probability of transferability needs to be may depend in part on other considerations, such as the importance of the end. The use of simian immunodeficiency virus (SIV) to research AIDS was fairly speculative, but in view of the seriousness of the AIDS problem, there was more to be said for it than would have been the case had the problem been less pressing.

One of the continuing moral dilemmas of human-oriented animal research is the need to affirm two propositions: (*1*) research animals are sufficiently like humans to allow reasonable inferences to be drawn from data involving one for the other; and (*2*) research animals are sufficiently unlike humans to justify (morally) our using them for experimental purposes. The two propositions need not be in tension, though the possibility that they are in any particular case must always be considered.

But even the most rigorously designed and monitored experiment is likely to have little to be said for it if there is a very low probability that the data it yields will advance the end it is ultimately intended to serve.

Are the Means Likely to Realize the End?

Though it is true that a number of significant scientific breakthroughs have been the outcome of happenstance and guesswork more than careful planning, such serendipitous occurrences cannot be appealed to as a substitute for the requirement that researchers make some case for the *likelihood* that their investigation will advance the ends to which their work appeals.

Likelihoods of course are always somewhat speculative, and will vary, and so will increments in knowledge and the advancement of particular ends. There is no measurable probability or simple likelihood of success to which all experiments should be required to conform. The degree of probability that might be expected of a particular project will depend on a variety of factors, such as the importance and/or urgency of the ends, the costs in animal life and suffering, the scarcity of resources, and the availability of alternatives.

Trade-offs between these factors are not easy to craft and require a well-rounded sensitivity. Although moral decision making is not a matter of numbers, one of the arguments for requiring committee approval of protocols involving animals is that the diverse interests involved may not otherwise be adequately represented. In theory at least, representative animal care committees (Institutional Animal Care and Use Committees) may provide an environment for the articulation and rational balancing of interests.

The reasonably expected likelihood of an end's furtherance will also depend on the level of research already undertaken. Some experimental ground has already been so thoroughly explored that the likelihood of new discoveries is very slight; other territory may offer only theoretical possibilities, with (as yet) relatively little empirical data. In some cases the more ambitious protocol may be preferable to the conservative one.

Are the Means Proportionate to the End?

Where good ends are sought, and enter into the justification of means, the costs incurred by those means have to be entered against the goodness of the ends. An experiment directed to and likely to advance a good end may nevertheless be unacceptable because its costs are disproportionate to the goodness of the end. Ends are not privileged with respect to means.

In this context, "costs" include not only straightforwardly economic or utilitarian costs, but what we can more broadly term "moral costs" — including the loss or abridgement of certain values, in particular, the value attaching to a life that is allowed to flourish free of burdensome constraints. The tasks of determining and arbitrating between these costs can raise the most intractable problems for judging animal experimentation.

Experiments involving animals may intrude on them in various ways: They may be killed, but even if not, their "lives" may be disrupted or constrained, and pain/suffering may be caused them. Although animals will obviously differ in the sophistication of their lives, we should not disregard the fact that such lives as they have possess an integrity and intrinsic value not entirely disanalogous to

the organic integrity and value that is possessed by human bodily life. A functioning organism cannot be reduced to a normatively neutral complex of chemical interactions, but is distinguished by an organic *telos* whose invasion should generate justificatory questions. In assessing such intrusions, not only should the immediate experimental protocol be taken into account, but also the costs incurred by housing and care, aftercare, and the destruction of natural environments or social bonds.

Finding an adequate moral language in which to cast these discussions is not easy. That many — albeit not all — (experimental) animals can experience *pain* is not generally disputed, though the extent to which those animals can *suffer* is somewhat more problematic. We might differentiate pain from distress, discomfort, anxiety, and fear, recognizing that each makes different assumptions about the capacities of its subject, and that some animals will be capable of experiencing some but not other kinds of such suffering. Furthermore, whether such pain and/or suffering should count morally for the same as human pain and suffering is even more problematic. There is little doubt that part of what is problematic about causing pain and suffering to humans is the way in which they tend to subvert autonomy and its fruits, and this is not likely to be an issue in (most) animal experimentation. Whether or not it is an issue is reflected in (thought not simply resolved by) debates about whether animals possess "interests" and/or "rights," and whether we have duties *to* them or simply *regarding* them.

The point of these debates is not usually to establish whether we owe animals any moral consideration, but to decide what kind of consideration is appropriate, and — remembering that we are not involved in a simple consequentialist calculus — to determine how weighty it is with respect to the various human interests that might be realized in experimentation. Even if it is argued that human interests, by virtue of their origination in deliberative activity, have morally relevant features that animal interests lack, it does not follow that every human experimental interest will take priority over any animal interest. One of the common complaints about the Draize test (in which potential irritants are tested on the eye tissues of animals) has been that the cost in animal pain/suffering often far exceeds the worth of adding a new product to an already well-supplied market. Human interests, no less than animal welfare interests, must be scrutinized and ranked.

Problematic though these judgments are, we should not assume too quickly that they are impossible to make. Insofar as the evaluation of alternatives is an enterprise undertaken by practical decision makers, it is ultimately up to *us* as decision makers, as rational beings and normative agents, to determine how much weight we will accord to environmental and/or social destruction as against physical pain, and how important these are with respect to the advancement of knowledge or human welfare. Contrary to those who see something fatally anthropocentric in such judgments, the perspective of those who have to take responsibility for what they do is the only appropriate perspective to take.

Although judgments of proportionality are in some ways too complex to allow of simple codification and commensurability, some assistance to researchers can probably be provided by setting out in a roughly ascending order different degrees of intrusiveness (taking into account not only the kind of intrusion caused, but also its intensity, duration, relievability, and so forth), cross-referencing it with different levels of animal complexity. Human ends might be similarly ranked, taking into account such factors as urgency, whether long- or short-term, those likely to be benefited, whether the ends to be served can be served in other ways, and so on. Some have complained that many of the human medical problems to whose alleviation much animal experimentation is directed are the result of voluntarily adopted lifestyles, and that lifestyle changes would not be an unreasonable expectation (albeit not, perhaps, a sufficient reason for refusing to engage in animal experimentation): The costs to animals, along with others, might be incorporated in a motivational package addressed to the problem of "unhealthy" lifestyles.

In one of its associated expressions, the proportionality requirement mandates the use of the least costly alternative consistent with the ends being realized. Using 200 experimental animals when 100 would do, using experimental animals when computer modeling or tissue cultures would do, causing pain and stress when the use of anesthesia would not compromise the outcome, using untrained animal handlers when experienced handlers would cause less animal distress, and so on, all represent abuses of the proportionality requirement, since there is a less costly way of achieving the same end. The so-called "three Rs" of animal experimentation — replacement, reduction, and refinement (14) — express this dimension of the proportionality requirement.

Some judgments here may be very difficult to make. It may, for example, be hard to determine whether or not the multiple use of single animals is to be preferred to the use of fewer procedures on more animals, or how the "costs" of using pound animals are to be assessed against those of breeding animals specially for experimental purposes. Do numbers count, as well as the amount of suffering? And how do we factor in the production of transgenic animals for experimental purposes? If an animal is bred to be disease prone, does this constitute a violation of species integrity, or does the new animal now have a natural end that makes experimental procedures (e.g., the testing of anticarcinogenic agents) more acceptable? Does the patenting of such animals provide some control over their use, or does it take us too far in the direction of an unacceptable commodification of animals, in which we come to see them as *no more than* commercially exchangeable tools?

Will the Means Undermine Other Good Ends?

Even if the ends to which an experimental procedure is directed are eminently worthwhile, and the procedure is appropriate to those ends and likely to advance them, and even if the direct costs of the procedure are proportionate to the ends sought, there may be other dimensions of its implementation that need to be taken into account and ranged against it. Sometimes, in pursuing one end, we

may undermine or jeopardize other ends to which we are independently committed.

Although it would need to be supported by data rather than merely conjectured, the claim that some kinds of animal experimentation tend to brutalize or barbarize researchers suggests how the pursuit of some worthwhile ends — say, human health — might undercut others — say, civilized sensibilities (15,16). One of the deeper anxieties that fueled nineteenth-century opposition to animal experimentation was concern at the dominance of science and its arrogant oligarchy of expertise, the depersonalization of social decision making, a growing detachment from the world of nature of which we are a part (17). It has been followed in the twentieth century by criticisms of human selfishness and profiteering, our concern with self-indulgence and self-advancement without regard to the costs for other living things (18).

A better documented example of the undercutting of other ends might be the effect of primate research on the persistence of an endangered species. The integrity of a species and the value of species diversity may be compromised if primates are used — or used without regard for their survival — for research into cancer or AIDS. Of course, as we noted earlier, ends themselves may be amenable to ranking, and it would not follow merely from the fact that a means of pursuing a good end would subvert another that it should be eschewed. Nevertheless, our consideration of costs should not be limited to those directly associated with the experiment at hand.

Serious though the foregoing claims are, like all claims they have not gone unchallenged. Researchers may just as easily see their activity as one of responsible stewardship rather than one of arrogant domination: the knowledge gained in animal research is seen as serving the good of human and animal well being, a task that falls to humans because of their unique endowments (18).

The resolution of such conflicts is unlikely to be a simple one. Like most human activities, from sport to road building, animal research is likely to be attended with larger social costs, whatever its benefits, and it will behove us to address them as they arise and seek to ensure that they do not fall victim to the political sloganizing that has characterized much of the current debate.

INSTITUTIONALIZATION OF JUDGMENT

In view of the complex nature of the ethical questions confronting the use of live animals in research, a two-part process for evaluating such research can be proposed. The first will consist in the development of formal guidelines for experimentation and research, guidelines that are sufficiently specific to the ends and subjects of the research to avoid the charge of vagueness. Guidelines might, for example, be developed that will be specific to field research (19–22), to biomedical experimentation (10–12), to product testing (23), and to education (24,25), as ends that will tend to generate different requirements and different questions. A further subdivision might have regard to the animals involved, taking into account levels of consciousness, social ecology, replaceability, and so on. Such guidelines will then address questions of intrusiveness with regard to these various "structural" factors.

It is most likely that such guidelines will be seen as the responsibility of national, international, or professional bodies — bodies with a wide enough representation and jurisdiction to ensure not only that a broad spectrum of opinion has been canvased but also that the resultant guidelines will possess the public standing that will allow them to be used as a meaningful standard in holding researchers accountable, whether by assessing eligibility for funding or by informing legal standards of proper use and handling.

Beyond that, however, there will need to be an informed and sensitive application of these guidelines to specific research protocols, and this might be best achieved through the activity of an Institutional Animal Care and Use Committee (IACUC), in which the various interests at stake in animal research may be represented. The task of an IACUC will not be to apply the guidelines in a formulaic or algorithmic manner but to make a judgment, in the light of the guidelines, about the ethical acceptability of a proposed research study or whether existing research facilities meet acceptable standards for animal care. This is important: Bureaucratic guidelines are almost always too crude for the purposes for which they are drawn up (18).

The role of IACUCs is disputed. Some believe that they should provide no more than scientific assessments of research protocols, and even those who believe that their mandate should extend to ethical questions often wish to limit that questioning to means — ends being seen as given — or to questions of general institutional policy. It is argued that the use of IACUCs to monitor specific protocols and even day-today institutional practice allows judgments to made by the inexperienced, diverts valuable resources from research, overburdens the monitoring system, and restricts academic freedom (26). Although these objections are not decisive, they warn that where the power of an IACUC is considerable, there is a corresponding responsibility on the part of institutions to ensure that their memberships are wisely constituted and adequately resourced (13).

Here too, some national guidelines might be appropriate to ensure that the representation of IACUCs does not too easily fall prey to the political winds that often affect even scientific research. Some effort should be made to ensure that the concerns of animal advocates are represented as well as researchers, and that wider public concerns about both scientific research and animal welfare are allowed voice.

Neither the provision of guidelines nor the approval of an IACUC will guarantee that good decisions are made, but they probably represent the best formal steps that fallible and contending humans can take to reach acceptable solutions.

CONCLUSION

The structure of moral deliberation proposed in this article has the merit of providing a framework of questions that forces to raise the basic issues that need to be confronted by

researchers wishing to engage in animal experimentation. In its broad outlines, the structure is both comprehensive and rich. It is comprehensive in that provides for all the morally significant questions to be asked. It is rich in that it allows these questions to be pursued at different levels of generality.

It does not, however, provide a simple procedure for cranking out answers to questions that demand sensitivity and judgment more than formulae. There is great diversity in the ends and subjects of animal experimentation, and researchers in each kind of experimental situation will need to determine that situation's moral ecology before they will be able to grapple with the difficult judgments that will often have to be made. Nor does the general structure provided come with ready made "pull-down" sub-menus. There is little doubt that those sub-menus provide the sites for some of the most difficult and intractable problems.

Nevertheless, this article proposes that within each domain of research — basic research, field research, medical research, product testing, and so on — researchers construct a series of fairly specific questions they will need to ask themselves, based on the general questions canvassed in this article. These questions could then take into account the more particular ends being pursued, the kinds of animals likely to be involved, the costs likely to be encountered by that research, and so forth. If the responses to these questions are then considered by an animal welfare committee or IACUC whose membership is collectively capable of appraising the social value and scientific merit of the proposed experiments, as well as representing the various animal and human interests involved, this might come as close as what can expect to come to a balanced judgment on an issue that will continue to challenge the quality of our moral perception.

BIBLIOGRAPHY

1. N.A. Rupke, ed., *Vivisection in Historical Perspective*, Croom Helm, London, 1987.
2. P. Singer, *Animal Liberation, N.Y. Review*, New York, 1975.
3. T. Regan, *The Case for Animal Rights*, University of California Press, Berkeley, 1983.
4. P. Carruthers, *The Animals Issue*, Cambridge University Press, Cambridge, UK, 1992.
5. B. Rollin, *The Unheeded Cry: Animal Consciousness, Animal Pain, and Science*, Prometheus Books, Buffalo, NY, 1989.
6. A. Rowan, F. Loew, and J. Weir, *The Animal Research Controversy*, Tufts University School of Veterinary Medicine, Medford, MA, 1995.
7. H. LaFollette and N. Shanks, *Brute Science: Dilemmas of Animal Experimentation*, Routledge, London, 1996.
8. F.B. Orlans et al., *The Human Use of Animals: Case Studies in Ethical Choice*, Oxford University Press, Oxford, UK, 1998.
9. Hastings Center, *Guidelines for the Termination of Life-Sustaining Treatment and the Care of the Dying*, Hastings Center, Briarcliff Manor, Westchester, NY, 1987.
10. Council for International Organizations of Medical Sciences (CIOMS), *International Guiding Principles for Biomedical Research Involving Animals*, CIOMS, Geneva, 1985.
11. Australian Government Printing Service, *Australian Code of Practice for the Care and Use of Animals for Scientific Purposes*, AGPS, Canberra, 1990.
12. Institute of Laboratory Animal Resources (ILAR), National Research Council, *Guide for the Care and Use of Laboratory Animals*, National Academy Press, Washington, DC, 1996.
13. F.B. Orlans, *In the Name of Science: Issues in Responsible Animal Experimentation*, Oxford University Press, New York, 1993.
14. W.M.S. Russell and R.L. Burch, *The Principles of Humane Experimental Techniques*, Methuen, London, 1959.
15. T. Aquinas, *Summa Contra Gentiles*, Bk III, Ch. cxiii, University of Notre Dame Press, Notre Dame, IN, 1975.
16. I. Kant, *Lectures on Ethics* (L. Infield, trans.), Century, NY, 1930.
17. R.D. French, *Vivisection and Anti-vivisection in Nineteenth Century Victorian England*, Princeton University Press, Princeton, NJ, 1975.
18. J.McA. Groves, *Hearts and Minds: The Controversy over Laboratory Animals*, Temple University Press, Philadelphia, PA, 1997.
19. American Society of Ichthyologists and Herpetologists, American Fisheries Society, and the American Institute of Fisheries Research Biologists, Guidelines for Use of Fishes in Field Research, *Copeia, Suppl.*, pp. 1–27 (1987).
20. American Society of Ichthyologists and Herpetologists, The Herpetologists' League, and the Society for the Study of Amphibians and Reptiles, Guidelines for Use of Live Amphibians and Reptiles in Field Research, *J. Herpetol.* **4**(Suppl.), 1–14 (1987).
21. Acceptable Field Methods in Mammalogy: Preliminary Guidelines Approved by the American Society of Mammalogists, *J. Mammal* **68**(4, Suppl.), 1–18 (1987).
22. American Ornithologists' Union, Report of Committee on Use of Wild Birds in Research, *Auk* **105**(1, Suppl.), 1A–41A (1988).
23. Foundation for Biomedical Research, *The Use of Animals in Product Safety Testing*, Found. Biomed. Res., Washington, DC, 1988.
24. Schools Animals Care and Ethics Committee, *Animals in Schools: Animals Welfare Guidelines for Teachers*, NSW Department of School Education, New South Wales, 1991.
25. National Science Teachers Association, *Guidelines for Responsible Use of Animals in the Classroom, NSTA Rep.*, p. 6, NSTA, Arlington, VA, 1991–1992.
26. N.H. Steneck, *Ethics and Behav.* **7**, 173–84 (1997).

See other entries ANIMAL, MEDICAL BIOTECHNOLOGY, LEGAL, LAWS AND REGULATIONS GOVERNING ANIMALS AS SOURCES OF HUMAN ORGANS; ANIMAL, MEDICAL BIOTECHNOLOGY, POLICY, WOULD TRANSGENIC ANIMALS SOLVE THE ORGAN SHORTAGE PROBLEM?; see also RESEARCH ON ANIMALS entries; TRANSGENIC ANIMALS: AN OVERVIEW.

RESEARCH ON ANIMALS, LAW, LEGISLATIVE, AND WELFARE ISSUES IN THE USE OF ANIMALS FOR GENETIC ENGINEERING AND XENOTRANSPLANTATION

F. BARBARA ORLANS
Kennedy Institute of Ethics, Georgetown University
Washington, District of Columbia

OUTLINE

INTRODUCTION

A new industry has emerged in recent years that uses laboratory animals in ways that have never before been tried. Animals are used as experimental subjects of studies involving genetic manipulation in which there has been a deliberate modification of the genome — the material responsible for inherited characteristics — to produce genetically modified animals or sources of organs and tissues for transplantation into humans (xenotransplantation). Mice are the most-used species in transgenic studies; pigs are the preferred species for xenotransplantation. As a result of these novel technologies, new animal welfare issues have arisen. There is public concern about the sheer increase in the numbers of animals used, the potential increase in animal suffering, and new forms of exploitation of animals. Existing legislation is not able to give these subject animals the quality of consideration and degree of protection to which they are entitled. U.S. federal laws, last amended in 1985, fall behind the current need for protecting the welfare of these animals. For instance, areas of concern are that mice are not included under the Animal Welfare Act, adequate limits are not placed on the invasion of the integrity of an animal, pigs are confined in limited space and barren environments, and more training of laboratory personnel is needed in clinical monitoring of animal pain and distress.

ANIMAL PROTECTIVE LEGISLATION

Two U.S. federal laws govern the use of laboratory animals in biomedical research. The first, now called the Animal Welfare Act (AWA), was passed in 1966. It required registration of animal research facilities with the United States Department of Agriculture (USDA), federal inspections, and the humane treatment and care of certain species of animals (dogs, cats, nonhuman primates, rabbits, hamsters, and guinea pigs). In 1970 Congress changed this wording so that additional warm-blooded species could be included as determined by the Secretary of Agriculture. Any future changes in what species were included would therefore be up to USDA to announce in its rule making. But no action was taken and numerous other animals were left unprotected. Strengthening amendments to the AWA made in 1970, 1976, and 1985 required the use of pain-relieving drugs, the establishment of Institutional Animal Care and Use Committees (IACUCs) to oversee compliance with the regulations, and promotion of psychological well-being of primates (1).

A second mechanism of control emerged in the 1960s covering the practices of grantees of the National Institutes of Health (NIH), which is part of the Public Health Service (PHS). NIH had provided federal grants for animal experiments since 1946. In 1963, in an effort to forestall the increasing efforts to establish federal legislation, NIH published for the first time voluntary guidelines called the *Guide to the Care and Use of Laboratory Animals* (commonly called the NIH Guide). Under the Health Research Extension Act of 1985 (P.L. 99-158), the NIH Guide is no longer a "guide" but law, and it now covers all federal agencies (e.g., the Department of Defense) and not only PHS. It is now called the Public Health Service Guide for the Care and Use of Laboratory Animals (2). In the 1960s these publications dealt only with husbandry standards — minimal caging size, sanitation, nutrition, and the like. Over time the scope of these publications has broadened to include provisions on experimental procedures as well.

Thus became established the two primary mechanisms for maintaining standards that continue to this day in the United States — AWA and its amendments administered by the Animal and Plant Health Inspection Service (APHIS) of USDA, and the policy of NIH/PHS administered by the Office for Laboratory Animal Welfare, (previously the Office for the Protection from Research Risks) of NIH. These laws and subsequent rule making govern the current conduct of animal research. The two oversight mechanisms cover different constituencies, although there is overlap in their purview. Recent efforts have been undertaken to make the provisions of each compatible with the other.

Funding for enforcement of AWA by USDA has always been a problem. USDA is required to inspect research facilities, dog and cat dealers, and zoos, since these are all covered under AWA. When AWA was first passed in 1966, an appropriation of $300,000 was barely achieved. Annual appropriations rose slowly for several years, but from 1992 to 1999 the appropriation has remained static at $9.2 million. (This compares with a congressional appropriation of over $17 billion to NIH for biomedical research in fiscal year 2000. NIH grants comprise an important national source of funding for animal research.) A shortage of USDA personnel has also been a problem. In Fiscal year 1997, for instance, a staff of about 73 animal care inspectors conducted almost 16,000 inspections to ensure compliance with AWA regulations (3). Currently a consortium of professional scientific and animal advocacy organizations is pressing for an annual increase in funding of at least 3 to 4 million dollars.

In 1998 there were 1227 animal research facilities registered with USDA under the AWA (4). This number

compares with some 970 institutions (in February 2000) that must comply with the PHS policy. Some overlap between the two groups exists. Still outside the provisions of any national policies are the academic and commercial institutions that either do not receive federal funding or use species of animals which are exempted. How many such exempt institutions there are is unknown, but the number probably runs to several thousands and includes privately funded facilities that conduct genetic manipulations on mice and rats.

Animal research facilities now have greater responsibilities than previously. The current legislation mandates that each research facility using animals establish an oversight IACUC with members appointed by the chief executive officer of the facility. Each committee is composed of no fewer than three members: one a veterinarian, and another not affiliated with the institution. In practice, animal researchers both chair the committee and dominate its membership. The most recent inquiry by the NIH office that administers the PHS policy found that only two of the approximately 1000 IACUCs have chairpersons who are not animal researchers.

Since 1985, protocol review by IACUCs has been mandatory. For instance, the PHS policy requires IACUCs to review relevant sections of PHS grant applications to ensure that (*1*) procedures with animals will avoid or minimize discomfort, distress, and pain to the animals, consistent with sound research design; (*2*) appropriate sedation, analgesia, or anesthesia is used; (*3*) animals that would otherwise experience severe or chronic pain or distress that cannot be relieved will be painlessly killed; (*4*) laboratory personnel are appropriately qualified and trained in the procedure(s) they are using; (*5*) methods of euthanasia are consistent with those prescribed by the American Veterinary Medical Association; (*6*) procedures involving animals are designed and performed with due consideration of their relevance to human or animal health, the advancement of knowledge, or the good of society; (*7*) the animals selected are of appropriate species and quality and the minimum number required to obtain valid results; and (*8*) methods such as mathematical modeling, computer simulation, and in vitro biological systems are "considered." In accomplishing these tasks, each IACUC approves, disapproves, or modifies the proposed animal experiment. The effectiveness of these committees is variable.

The AWA 1985 amendments also added other provisions: Training must be provided to laboratory animal personnel in the humane care and use of animals, the environment in which nonhuman primates are housed must promote the animals' psychological well-being, and dogs must be given exercise. The Secretary of Agriculture was required to issue standards governing these provisions.

The first of these new provisions fared reasonably well in the 1991 USDA rule making implementing the AWA amendments. As a result institutions for the first time began to provide training for their personnel in animal handling, anesthesia, and euthanasia, and what the concept of "Three R" alternative means — to replace animal experiments with nonanimal methods where feasible, to reduce the numbers of animals used, and to refine procedures to minimize or eliminate animal pain and distress, concepts that are all included in the laws. In response, the climate changed appreciably. IACUCs began to be more alert to the qualifications of research investigators to conduct traumatic procedures and to insist on training of the laboratory personnel who were not familiar with the techniques involved. The role of the veterinarian in providing this on-site training became progressively important. The legal requirement that the experiments not be duplicative led to greater use of the computerized library resources at the increasingly influential Animal Welfare Information Center at USDA and at the National Library of Medicine.

Psychological Well-Being

The other two new AWA provisions fared less well. The congressional requirements to promote the psychological well-being of primates and to provide exercise for dogs proved highly controversial. Congress had left unclear exactly how it wanted USDA to write the rules. Researchers protested the inclusion of these requirements in the law, arguing that "well-being" was unmeasurable, exercise for dogs was unnecessary, and any changes would be too costly.

There was considerable delay in USDA's promulgating rules governing primate psychological well-being. What appeared finally was permissive vis-à-vis the biomedical community. Instead of setting specific standards (as wanted by the humane community), the regulations allow each laboratory to determine how it will improve treatment of research animals, and a great deal of discretion is allowed (as wanted by the research community).

Former Congressman John Melcher, a veterinarian and the person responsible for adding the amendments regarding primate well-being and exercise of dogs, wrote of his regret about the 1991 rule making in the *Washington Post* (5). He said: "Imagine a small cube of a cage three feet on a side and three feet high. Within this cube lives a primate — often a baboon or a rhesus monkey — that could weigh as much as 55 pounds. Baboons usually stand on all four feet, but in such a space they cannot walk anywhere. They cannot stand upright or stretch their arms in such a cage. Yet this is a common caging for the animals used for scientific research. . . . The USDA has failed [in their new regulations]." The USDA rules went into effect in 1994, nine years after the passage of the law.

In July 1999, many years later, USDA reported that research facilities do "not necessarily understand how to develop an environmental enhancement plan that would adequately promote psychological enrichment," and therefore additional policies have been proposed (6). As of February 2000, public comments on these proposals were being assessed before issuance of additional policies.

Despite all these problems, without the 1985 law, funding for research projects to explore environmental enrichment would not likely have been forthcoming. Importantly, NIH started funding projects designed to test the beneficial effects of primate housing that allowed the expression of normal behaviors so that the animals would not be bored and come to express stereotypical obsessive behavior, such as constant rocking

or bar chewing — signs of psychological trauma. In fact the expression of stereotypic behaviors may indicate a disordered nervous system, bringing into question the validity of data derived from them (7).

Research has established that enrichment schemes are beneficial to the animals (8–10). Examples include group housing, more space, addition of climbing apparatus and manipulative devices such as chew toys and mirrors, and feeding enrichments. As a result of these efforts there has been a reduction (although not elimination) in the proportion of laboratory chimpanzees, baboons, and other primates that are psychologically damaged and a concomitant improvement in the quality of science.

Expanding Coverage to all Species Used

More than 30 years after the enactment of AWA, several widely used species of laboratory animal are still not covered, notably mice, rats and birds, despite the fact that according to commonly accepted estimates, they comprise 80 to 90 percent of all animals used. Farm animals were not included until 1990.

The animal welfare community has long fought to have *all* species of animals used in experiments included in AWA. No act of Congress is needed, only that USDA amend its rule making. (The Public Health Service policy *does* include all species inasmuch as all vertebrate animals are covered, so at PHS grantee institutions this is not an issue.) With regard to AWA, initially USDA had enough on its hands just to get the law into operation. But as time went on, the exclusions of certain species became a glaring problem.

Over the years individuals and groups concerned about the welfare of animals have exerted pressure to drop exclusion of agricultural farm animals, mice, rats, and birds. USDA has resisted on the ground of financial cost. However, in 1990 USDA finally ruled that horses, sheep, goats, cows, and pigs when used for *biomedical or other nonagricultural research* are covered by AWA. (Excluded still are farm animals used in genetic engineering research to increase productivity for food and fiber purposes and also to produce various biologics and pharmaceuticals.) As of 1990, pigs used for xenotransplantation and other biomedical research must be maintained and cared for in compliance with AWA standards and the facilities are subject to USDA inspection. (Pigs used in such research in PHS-funded facilities have always had to comply with PHS standards.) Between 1990 and 1998, the use of farm animals for biomedical research has more than doubled (66,702 to 157,620 per year) (4). In particular, the use of pigs has grown significantly.

Mice, Rats, and Birds. USDA is the target of complaints about failure to expand coverage of AWA to other species. This is because Congress had given USDA discretion about which species to include in addition to those initially mandated in 1966. After years of trying persuasion, in 1990 the Animal Legal Defense Fund brought suit against USDA to amend the regulations to include mice, rats, and birds. USDA objected to these inclusions because of lack of money and resources. The animal rights group won its case. In a judgment issued January 8, 1992, in the U.S. Distric Court, Judge Charles R. Richey ruled that USDA's exclusions were "arbitrary and capricious" and that USDA must issue new rules to include these species (11). USDA appealed the court's decision.

Pressures continued to mount. In 1998 the Alternatives Research and Development Foundation (a branch of the American Anti-Vivisection Society) filed a petition with USDA requesting that the agency amend its definition of "animal" to include mice, rats, and birds. The *Federal Register* announced receipt of the petition as well as the agency's response (12). Again USDA's arguments reflected its previous opposition to inclusion on the basis of lack of resources for implementation. Also USDA stated its belief that the majority of these animals were already being afforded certain protections, asserting that 90 percent of these animals are provided oversight by PHS assurance, voluntary accreditation, or both. The agency stated that most biomedical research in the United States is performed in laboratories funded at least in part by PHS. In addition 600 facilities in the United States are accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC) and therefore voluntarily comply with their standards (which, like those of the PHS, cover mice, rats, and birds).

If mice, rats, and birds were included under AWA, the impact would fall primarily on two groups of animal users: (*1*) two- and four-year liberal arts and community colleges that, in general, use birds, mice, and rats for student education in preference to any other species; and (*2*) commercial genetic engineering companies.

There are an unknown number of facilities that conduct research on transgenic animals. Such facilities neither fall under the PHS policies nor need to be accredited, since the accreditation system is voluntary.

Mice are the species most frequently used to decode human ills. They are genetically modified to have human diseases such as diabetes, cancer, multiple sclerosis, arthritis, and a host of other ailments. In some experiments, multiple pathologies and serious animal welfare problems such as chronic pain and weak legs resulting in inability to stand up have been reported. Since commercial genetic engineering facilities are outside the law, they are not inspected by federal officials, nor do they have to have IACUCs for protocol review. They are also not required to use approved euthanasia methods. Although some companies are doubtless maintaining acceptable standards of animal care and use, there is no public assurance that they are, and there is no overt public accountability. With the added factor of secrecy in this highly competitive enterprise, public concerns arise. Indeed, major ethical concerns arise from a dangerous combination of factors, including lack of federal oversight in an industry based on the use of procedures which can cause considerable suffering in order to model severe human disorders, and the pursuit of profit.

Arguments for and against Inclusion of Additional Species. The vast majority of animal advocacy organizations have voiced support for including mice, rats, and birds. Among the arguments presented is that expanded

coverage is a matter of justice in that animals of similar moral worth should be treated equally (13). On this view, there should be a difference in some morally relevant characteristic if animals are to be treated differently. Mice and rats are so similar to three species currently regulated under AWA (hamsters, guinea pigs, and gerbils) that it is arbitrary, against common sense, and unjust to exclude them from legal protection. Indeed, all five species, mice, rats, hamsters, guinea pigs, and gerbils are similar in many ways. All are commercially purpose-bred for research and widely used; physiologically and anatomically they are, to a large extent, commensurable. Furthermore, and importantly from an ethical standpoint, the burdens they bear as subjects of biomedical experiments are of the same order: They have similar sensibilities in their perception of pain, and all are likely to be killed before the end of their normal life span. It is reasonable to assume that these species have an interest in not being subjected to pain or suffering and not having their lives prematurely foreshortened. Inasmuch as humans cannot be treated differently, unless there is some morally relevant basis on which to do so, dissimilar legal protection cannot be justified for animals that have similar relevant characteristics.

Several organizations have voiced opposition to expansion of coverage to all species used. These include the National Association for Biomedical Research, which states that expansion of AWA to include mice, rats, and birds is "a luxury we can do without" (14). The Federation of American Societies for Experimental Biology argues that expansion "would represent redundant regulation" (15). Other comments underline the fact that this would represent an unnecessary burden and hamper enforcement of existing regulation. However, some scientifically based organizations such as the Association for Accreditation and Assessment of Laboratory Animal Care (the accrediting agency) and the Scientists Center for Animal Welfare have voiced approval of including rats, mice, and birds. As of February 2000, USDA was analyzing the public comments they have received in response to their announcement in the Federal Register to determine what future action to take.

NATURE OF PUBLIC CONCERNS

Public response to new biotechnologies has been both enthusiastic and cautious. Genetically modified animals provide an extremely powerful tool for the development of disease models, since the mechanisms of gene regulation will become better understood. In addition the use of genetically modified mice as models of human diseases closely mimics the human disease and may even, in time, replace the need to use more acutely sentient animals (nonhuman primates) as models. But several reservations have also been expressed. Five issues of public concern are discussed below: the overall increase in (*1*) numbers of animals used, (*2*) the sum of animal suffering, (*3*) the invasions into the integrity of the animal, (*4*) the standards of housing and care of pigs, and (*5*) the monitoring of animal pain and distress. These concerns arise despite acknowledgment that the end result of these novel experiments is likely to be of significant benefit to humans.

Increase in Animal Use

When the U.S. Congress passed AWA in 1966, it recognized that keeping proper records is essential to ensure public accountability. It ordered that the numbers of laboratory animals used be counted and publicly reported each year, and this tabulation has been performed since 1973. But since mice, rats, and birds are not included under the definition of "animal," they are not counted. Data on the most-used species are missing from the statistics. This lack of information is detrimental to animal advocates who wish to track trends in animal use as part of their endeavor to reduce use and to target areas for reduction in animal pain and distress. It hampers commercial estimates of the future need for laboratory animals. It also prevents the public from participating in an informed debate.

Other sources of information show that use of mice is increasing in the United States. A 1999 article reported that Harvard Medical School will probably double its use of mice over the next five years — to about one million mice annually (16). Harvard is no exception. In 1991 NIH reported the use of 294,000 mice in intramural research; this number had increased to 648,000 in 1997. At both institutions, the increased use of mice is attributed to an increased number of experiments involving genetic manipulation.

Even more telling are statistics from the United Kingdom, since these are national data. In UK data both the numbers of animal procedures are reported and their purpose. Of the 1998 total of 2,659,662 animal procedures performed, genetically manipulated animals comprised 447,612 or approximately 20 percent of the total, and their use had doubled since the preceding year (17). In 1998 mice were used in 96.6 percent of the procedures comprising genetic modification. Other species used, in descending order were rat, pig, sheep, and other species.

The data in Table 1 provide some notion of the extent of animal experimentation worldwide. There are over 28 million animals counted in official statistics. This is an underestimate because many countries that use animals for experimentation do not count the numbers used. Not available, for example, are data from South America, Eastern Europe, the Middle East, Russia, Africa, and Asia. Regulations governing the humane use of laboratory animals exist in all geographic areas represented in Table 1 — some regulations being more and others less rigorous than those of the United States. In several countries where animal experimentation is conducted, no legislation exists (18).

In all probability the United States is the largest user of laboratory animals worldwide. In fiscal year 1998, a total of 1,213,814 animals were officially reported to have been used in research in the United States (19). When this figure is adjusted for uncounted species, the total comes to over 12 million animals per year in the United States alone (see Table 1). Over the period from 1973 to 1998, the total numbers of animals counted has fluctuated between 1.7 to 1.2 million animals per year, indicating a decline over the last six years.

Table 1. Number of Laboratory Animals Used in Research, by Country

United States (1998)	12,138,000[a]
United Kingdom (1998)	2,660,000
France (1997)	2,609,000
Canada (1997)	1,472,000
Belgium (1996)	1,516,000
Germany (1998)	1,532,000
Australia (New South Wales, South Australia, Victoria, Tasmania, and Western Australia) excludes fish (1996/7)	1,141,000
Italy (1996)	1,094,000
Netherlands (1997)	713,000
Norway (1997)	630,000
Spain (1996)	507,000
Switzerland (1998)	492,000
Denmark (1997)	380,000
New Zealand (1998)	309,000
Sweden (1997)	267,000
Austria (1996)	205,000
Finland (1998)	195,000
Ireland (1998)	69,000
Portugal (1996)	50,000
Hong Kong (1998)	27,000
Greece (1996)	19,000
Total	28,025,000

Note: Numbers represent official statistics of all countries and regions for which information could be found. Numbers are given to the nearest thousand and figures in parenthesis indicate year of count. The data presented are not necessarily comparable from country to country because of differences in animal species included. For instance, fish comprise 91 percent of the animals used in Norway but are not counted in the United States. Also some countries count experimental procedures and others individual animals.

[a]The United States counts only about 10 percent of all animals used in experimentation, since the most used species — rats, mice, and birds — are not protected under the relevant legislation and are exempt from counting. The official count for 1998 was 1,213,814, and this figure has been multiplied by 10 to allow for uncounted species and to achieve approximate comparability with data from other countries.

Increased Animal Suffering

In 1985 scientists at the USDA Beltsville Research Center called in the media to see the first ever genetically modified animals — creatures who became known as "the Beltsville pigs." In an attempt to produce faster-growing animals, these pigs had been genetically modified to express very large quantities of human or bovine growth hormone. The experimental purpose was to bring potentially greater profits to the food industry The public reaction to the pictures was of shock and criticism because of the obvious animal suffering. Some animals had damaged vision or deformed skulls, and others were unable to walk properly. The long-term deleterious effects for these animals were demonstrated two generations later and included gastric ulcers, arthritis, cardiomegaly, and nephritis (20).

Animals are now increasingly used to model human diseases; these studies can involve severe animal suffering. Among the painful diseases that have already been produced by genetic manipulation of mice are cancer, cystic fibrosis, Huntington's disease, and a rare but severe neurological condition called Lesch-Nyhan's syndrome that causes the sufferer to self-mutilate (21). Because the technology is in its infancy, the outcome for the animals is still somewhat hit or miss. Multiple pathologies are frequent: legless animals have been born, and abnormalities in genital organs, liver, kidney, joints, and vision also appear. Until techniques are worked out, survival levels are poor and considerable wastage of animals can occur. Unexpected, uncontrolled, and even undetected animal suffering often result. Particular attention needs to be paid to the ethical justification in terms of likely benefit to human health compared with the likely suffering of the animals.

In some countries statistical data may become available to measure this increased sum of animal suffering. A refinement in national statistics not found in the United States is that data are presented according to the "severity band" or "invasiveness" of the procedure, either minor, moderate, or severe. As of February 2000, six countries mandate that investigators rank their proposed procedures according to the degree of pain and distress (22). Two countries, Canada (23) and the United Kingdom (24), have also begun categorizing the level of adverse state resulting from specific genetic procedures. The Canadian guidelines on genetically modified animals stipulate that proposals to create novel transgenics initially should be assigned CCAC category of invasiveness level D (moderate or severe distress or discomfort) at least until the phenotype has been evaluated (25). In theory, statistical data on genetic modification procedures could be developed that report both on the numbers of animals used and the severity of the procedures used, ranked as minor, moderate, or severe. The rankings of severity would be made by the investigators in concurrence with IACUC. Over years, this would provide information on, for instance, the reduction in severity of effects from genetic modification procedures as the techniques become refined.

In the United States, major reforms would be needed to equal the data already available from other countries. U.S. official statistical data are deficient in not including mice, rats, or birds at all, in not specifying the experiments' purpose, and in not ranking the procedures' severity. Public concerns are fanned by the lack of animal data and lack of disclosure.

Integrity of the Animal

A particularly troublesome issue is interference with the integrity of an animal. The 1997 Experiments on Animals Act of the Netherlands requires that biomedical experiments on animals must be conducted "in recognition of the intrinsic value of animal life" (26). This is the first law in the world with such a statement. But what limits should be placed on preserving the integrity of a life form?

"Naturalness," "integrity," and "intrinsic/inherent value" are concepts that are open to differences in interpretation. For example, are laboratory animals simply tools to exploit, or do substantial alterations to the genome violate species-specific life — their "telos"? What constitutes the pigness of a pig, that is, the telos of an animal? Some critics believe that a pig should not be altered to the point that it ceases to be recognizably a pig. They question whether the biotechnology industry is attempting to

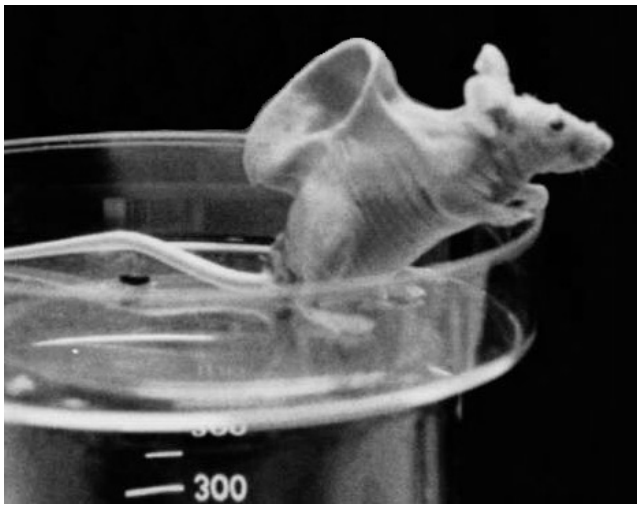

Figure 1. This is an actual photo of a genetically engineered mouse with a human ear on its back. It appeared in the *New York Times*, A11, October 11, 1999, in a full page advertisement protesting this invasion of the boundary between lifeforms. Credit Associated Press.

"capture [control of] the evolutionary process" and distort life forms in ethically unacceptable ways. "Who appointed the biotech industry as the Gods of the 21st Century?" was the heading of a 1998 protest against biotechnological manipulations of living beings. This advertisement, published in the *New York Times*, showed a photograph of a mouse that had been genetically manipulated to cause a human ear to grow on its back, see Figure 1. Many people worldwide were shocked at the sight of this photograph. The sponsors of the protest were a coalition of 19 organizations including the International Center for Technology Assessment, the Council for Responsible Genetics, and the Humane Society of the United States. Amazing mixtures of genes have been tried. For instance, jellyfish genes have been installed in monkey embryos, as well as injected into mouse sperm, and eggs and pregnancies created (27). (The goal of these studies was to test a technique that might eventually be used to create monkeys with added human genes.) The question is: Is there a boundary between life forms that should not be crossed?.

It is true that humans already share many genes with other animals — for instance, about 98 percent of the genes of chimpanzees and humans are identical, so only about 2 percent represent uniquely human characteristics. Because of the similarities, placing a single gene in another species may not, in itself, be so objectionable. But if future developments allow the transfer of uniquely human characteristics to animals, then ethical concerns would be increased. Mench provides a useful discussion of the issues of gene transfer (28).

Standards of Pig Housing and Care

There is considerable public support for the belief that laboratory animals should be provided with humane housing appropriate to their needs. Freedom of movement and the opportunity to express natural behaviors are viewed as basic animal needs. Pigs are intelligent animals with a range of bodily movements and natural behaviors which, if thwarted, result in manifestations of psychological damage such as obsessive, stereotypic behaviors. In this respect, pigs are like baboons and other primates. As discussed above, the living quarters for primates have been significantly improved over recent years. Will lessons learned about the importance of primate environmental enrichment flow over to other species such as pigs?

Among researchers and ethicists, a consensus is emerging that pigs are the preferred species for the routine supply of organs for xenotransplantation. According to the UK Nuffield Council on Bioethics, such use is "ethically acceptable" (29). The rationale for using pigs rather than nonhuman primates is that pigs have less highly developed mental capacities because they are less closely related to humans and that there is less chance of transmitting diseases from pigs to humans than from primates to humans. The use of pigs as a source of tissue for human use introduces the necessity of genetic modification — the creation of transgenic animals bearing human genes — in order to reduce the risk of hyperacute rejection. Further use of immunosuppressive treatments of the patients, to mitigate the danger of rejections, means that all tissues for transplantation must therefore be produced under sterile conditions in order to reduce the transmission of infectious diseases from animals to humans.

Initial production of transgenic pigs begins with impregnation by artificial insemination and removal of fertilized eggs to be microinjected with the required human gene. The pigs produced through this breeding procedures are used to stock the expansion herd through early weaning procedures. The source animal herd is established as a qualified pathogen-free herd. The pigs are born via hysterectomy or hysterotomy, taken from the sow, and reared in groups in isolated environment. The piglets are kept in isolators for 14 days, having no contact with the sow or the sow's milk (30).

Pathogen-free pig housing conditions are restrictive and closely controlled; they can therefore be stressful. Such housing conditions can mean, in some laboratory facilities, completely barren enclosures. The walls may be stainless steel or sometimes they are like tiled bathrooms. Flooring may be slatted fiberglass, or some other sterilizable matting. The flooring may not be ideal for pigs to walk on but is chosen for its hygienic qualities. Social deprivation is another problem because these pigs are individually housed and may be out of visual contact with other animals. There may be plenty of physical room, but if there is nothing *in* the room, the environment deprives the animal of normal social and play behaviors. The intelligent and social nature of pigs makes such deprived housing stressful.

Are such barren environments essential, or could enrichment alternatives be introduced? An alternative is to maintain the pigs in groups that are treated as a microbiological unit. Pregnant sows show a clear preference for a bedded surface rather than an unbedded surface, and sterilized straw is available (31) or possibly irradiated straw, as the pigs particularly enjoy having straw to root in. On the market are toys suitable for pigs housed in pathogen-free environments, such as teflon

balls and other products (32). In cases where pathogen-free environments are not essential, an even greater range of environmental complexity is possible, such as inclusion of wood shavings to provide soft surfaces, scratching posts, chains, footballs, concrete blocks (33), and showers. Often these enrichments are omitted because they are more trouble and take extra time to maintain. Enrichments are being used in some laboratories, but encouragement is needed to make such enrichments standard practice.

A high quality animal welfare system for maintaining pigs, the Nuertinger System (34) developed in Germany, is used by some laboratories conducting xenotranplantation research, such as Imutran Ltd. in Cambridge, UK (35). Neither gestation nor farrowing crates (discussed below) are used. The Nuertinger System has a number of animal-welfare-friendly features: It comprises a warm insulated bed and a cooler area for loafing, feeding and drinking, which gives the pigs a choice of environment and temperature, allowing them to choose where they can rest comfortably; see Figure 2.

Gestation and Farrowing Crates. Another issue concerning the breeding of pigs, not specific to xenotransplantation, is the use of gestation crates (also called gestation stalls) and farrowing crates — controversial practices developed in intensive food production units but now becoming increasingly used in laboratory settings. Both systems involve significant restraint of the sows, so that the animals can only stand up and lie down; they cannot turn around or walk. The animals cannot express their normal behaviors and become severely stressed, especially in gestation crates because the period of confinement is longer. Use of both gestation and farrowing crates is standard practice in the U.S. pork industry. Public protest against the use of gestation crates in particular, and to a lesser extent of farrowing crates, has been raised both in the United States and Europe. In the United Kingdom, confinement of sows during gestation and pregnancy in gestation crates has been prohibited since January 1, 1999 (36).

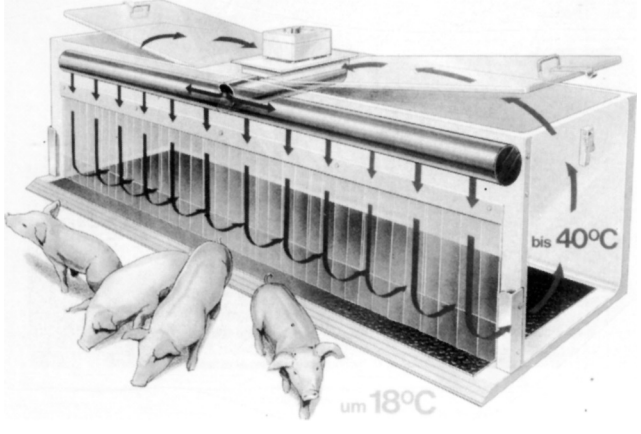

Figure 2. The Neurtinger System for housing pigs. The animals are group housed and can move freely through a plastic curtain between two areas that are maintained at different temperatures, up to 40 °C (104 °F) in one area, and about 18 °C (64 °F) in the other. Among the various environmental variables that can affect the welfare of pigs, temperature is the most important. Often piglets prefer to lie in a warm bed with their heads in the cool fresh air. For suckling, the air can be maintained at a cooler temperature. A controlled ventilation system regulates the temperatures. Credit: HAKA.

Gestation crates are used in food production units because they are economical of space and therefore inexpensive, and because they involve minimal human labor. In the agricultural industry, where profit is a dominating factor, these crates have come into almost universal use. The dimensions and use of gestation crates, as given below, are taken from the *Ag Guide* (37). They have overall dimensions of 5 by 7 feet, with a confining area that is 22 to 24 inches wide that constrains the sow and prevents her from turning around. Swine are large animals and can weight up to 600 pounds. Pregnant females are held in this confinement for four months — the whole period of gestation — and the animals are severely stressed. The animals lack freedom of movement and social interaction with conspecifics: They are unable to root, and they are deprived of expressing their strong maternal instincts of nest building. These deprived housing conditions are beyond the pregnant animals' abilities to cope, and as a result, they frequently exhibit stereotypic behaviors such as bar chewing, vacuum chewing, or head waving. Despite these welfare problems, gestation crates are used to breed animals in some American laboratories. Furthermore pigs are obtained for laboratory use from farms where confinement gestation crates are standard practice.

In laboratories there have been few pressures to avoid using gestation crates despite the fact that humane standards traditionally have been more rigorous for laboratory animals. The one-time clear distinction between farm and laboratory practices has become blurred. Traditionally there have been higher animal welfare standards in laboratories than in farm situations, but this tradition is being broken with the use of gestation crates. The rationale for using gestation crates in laboratory settings is weakened by the fact that the numbers of pigs maintained is relatively small, and there is no pressure to use minimal space to bolster profits as in food production units.

Farrowing crates represent another intensive farming practice that is now in use in laboratories. They too involve restraint of sows and restriction of normal movement but are far less objectionable than gestation crates because they are used for shorter periods, from several days to a few weeks, depending on the period of suckling. The conventional farrowing system, as described and sanctioned in the *Ag Guide*, is rectangular and measures 5×7 feet (1.5×2.1 m). But the sow resides in a crate within that area that is typically 2×7 feet (0.6×2.1 m). This width of 24 inches (0.6 m) restrains the sow so that she can, again, do no more than stand up and lie down. The main objective of farrowing crates is to slow the sow as she lies down, so that the piglets can escape to the sides and avoid being crushed. Piglet death is a serious issue, and a balance should be found between allowing the sow postural adjustments and freedom of movement against the crushing deaths of the babies.

John Webster, professor of animal husbandry at the University of Bristol, England, objects to farrowing crates on the ground that they restrict the opportunity for social contact between sow and piglets, and that the sow that is so confined during the period prior to farrowing is unable to satisfy her powerful motivation to build a nest (38). "The farrowing crate is certainly not designed to assist farrowing. The sow is uncomfortable, frustrated and compelled to drop her piglets on the same spot she drops her faeces," he states. The 1996 UK Advisory Group on the Ethics of Xenotranplantation considers the use of farrowing crates "undesirable" and states that "in theory it may be possible to avoid their use in the future" (29, p. 80).

Some refinements on traditional farrowing crates are already in use. For instance, the ellipsoid crate (39) and the Ottawa crate (40) (both developed in Canada) involve lesser degrees of constraint than conventional farrowing crates. Both systems allow the sow to turn around in addition to other movements and also permit easier visual and tactile contact between the dams with their piglets. They do not increase pig crushing rates. The ellipsoid and Ottawa crates are in use in Canadian farms and are suitable for laboratory use. Additional research is currently being conducted, particularly in the United Kingdom and Europe, to find alternative farrowing systems that are more welfare-friendly.

In the laboratory the choice of minipigs (rather than farm pigs) avoids the use of both gestation and farrowing crates. Miniature swine are preferred for this and other reasons. They are bred specifically for research from small wild species who have a mature body weight ranging from 132 to 198 pounds (60–90 kg) compared with domestic farm pigs who typically weigh in excess of 550 pounds (250 kg) at maturity. Recently minipigs have gained popularity for research studies including genetic modification and xenotranplantation.

Monitoring Animal Pain and Distress

Our understanding and recognition of animal pain and distress has advanced significantly in recent decades, and a considerable literature exists (41–45). Currently what is new is the use of scoring systems to keep track of the health status of laboratory animals during intensive periods when adverse effects are developing. With novel experimental techniques to model human diseases, it is especially important to maintain constant clinical monitoring of animal pain and distress. Researchers need to know when to intervene to relieve suffering by the use of drugs, by intense supportive care equivalent to that given to a human in the same state, or other actions.

In the last few years, several useful programs for intensive clinical monitoring of laboratory animals have been proposed. They require that the health status of the animal be evaluated at regular intervals and graded either on a numerical scale or a simple + or − (presence or absence) of physiological conditions or behavior, which together provide an assessment of severity of adverse conditions (46,47).

One such monitoring system in use at NIH includes assessment of the severity of a neurological impairment of mice on a grade from 0 to 5. Grade 0 is normal and grade 5 is moribund. Grades 1 to 4 show increasing signs of incapacity include clumsiness, incontinence, flaccid tail, abnormal plantar response, mild parapareses, trouble initiating movement, inability to move one or both hind legs, noticeable gait disturbance, moderate quadriparesis, and quadriparalysis (48). Qualified staff make frequent clinical assessments of the animals. At each grade, specific interventions have been established such as administration of fluids, dietary changes, expression of bladder, and provision of supplemental heat. The point at which early euthanizing of the animals should occur is specified.

It has long been recognized that a limit should be placed on the suffering of all animals used in experimentation. One way to achieve this is to set early, humane endpoints of experiments. The experimental design should establish the earliest point at which adequate scientific data have been collected and the experiment can be stopped, thus minimizing animal pain, distress, or lasting harm (49–51). In cancer studies, for instance, the endpoint should not be the death of an animal but the earliest point at which adequate scientific data are obtained. New initiatives are needed to foster use of these humane experimental designs (52).

Currently there are a few veterinary surgeons, animal ethologists, and others who are able to make such intensive clinical assessments and determinations of criteria for early endpoint on genetically engineered animals. But there are not enough persons so qualified; additional training programs are needed. However, evaluation of pain and distress for genetically modified mice may be difficult in some laboratory facilities, particularly those using micro-isolator cages and where there are large numbers of knockouts that are being subjected to a variety of breeding strategies to determine the effect of genetic deletions. Knockouts refer to animals who have had one or more genes removed or "knocked out." Defects in transgenic animals can be subtle but still affect welfare. Examples are mismothering, aggression, and spatial disorientation. These require extra special monitoring.

CONCLUSION

It was in 1985 that two important events took place — public awareness of the Beltsville pigs, and the latest amendment to AWA The Beltsville pigs demonstrated not only the scientific potential of genetically modifying animals but also the new welfare problems involved. With no amendment to the law in 15 years, U.S. laws lag behind existing needs because they were implemented before the full welfare implications of genetic modification were recognized. Scientific advances in genetic research challenge ethical norms, and the implications should be carefully considered before such work is approved. There is an

expectation in the science community and society generally that genetic modifications should occur within a framework of legislative controls that minimizes the impact on the animals involved. Some national guidance and mandates are needed to help deal with these issues. Animal welfare concerns raised in this article are also being grappled with in other countries: policies are being prepared to address societal issues arising from genetically modified animal experimentation. One such proposal directed toward the European Union has just been published (53). It includes a specific cluster of questions around the issues of justification, scientific relevance, animal suffering, and wider social, economic, and environmental impacts of animal studies involving genetic modification. This proposal may help point the way toward a reconsideration of national policies and a development of fresh initiatives.

ACKNOWLEDGMENT

I wish to sincerely thank the following colleagues for their helpful comments during the preparation of this manuscript: Clément Gauthier, Gilly Griffin, John Gluck, Joy Mench, David Morton, and LeRoy Walters. I also acknowledge the able library assistance of D'Anna Jensen.

BIBLIOGRAPHY

1. Food Security Act of 1985 P.L. 99-198, amendments to AWA Improved Standards for Laboratory Animals Act.
2. National Institutes of Health, *Public Health Service Guide for the Care and Use of Laboratory Animals*, No. 86-23 Rev., NIH, Bethesda, MD, 1985.
3. *Fed. Regist.* **64**(18), 4357 (January 29, 1999).
4. United States Department of Agriculture, *Animal Welfare Report Fiscal Year 1998*, APHIS, Washington, DC, 1998.
5. J. Melcher, *Washington Post*, September 8, C5 (1991).
6. R. DeHaven, *Lab Animal* **29**(1), 44–46 (2000).
7. M.H. Lewis et al., *Brain Res.* **513**, 67–73 (1990).
8. J. Bielitzki et al., *Lab. Animal Sci.* **40**(4), 428–431 (1990).
9. K. Bayne et al., *Am. J. Primatol.* **23**, 23–35 (1991).
10. K.J. Field, J. Denny, and G. Kubicz, *Lab. Primate Newslett.* **36**(2), 5–7 (1992).
11. *Madigan*, 781 F. Suppl. at 805–806.
12. *Fed. Regist.* **64**(18), 4356–4367 (1999).
13. F.B. Orlans, *Lab Animal* **28**(8), 13 (1999).
14. *Foundation Biomedi. Res. News* **16**(1), 1–2 (1999).
15. *FASEB Newslett.* **11** (June 1999).
16. J.F. Lauerman, *Harvard Mag.* **52** (January/February 1999).
17. *Statistics of Scientific Procedures on Living Animals: Great Britain 1998*, Home Office, London, July 1999.
18. F.B. Orlans, in H. Kuhse and P. Singer, eds., *Companion to Bioethics*, Blackwell, London, 1998, pp. 399–410.
19. USDA, APHIS, *Animal Welfare Rep.*, No. 36 (1998).
20. V.G. Pursel et al., *Science* **244**, 1281–1288 (1989).
21. R. Weiss, *Washington Post*, June 7, A1, A12 (1999).
22. F.B. Orlans, *ATLA* (in press 2000).
23. CCAC Guidelines, *Transgenic Animals*, Canadian Council on Animal Care, Ottawa, Canada, 1997.
24. *Supplementary Guidance to Applicants for Project Licences: Projects to Generate and/or Maintain Genetically Modified Animals*, Home Office, London, 1999.
25. *Categories of Invasiveness in Animal Experiments*, Canadian Council on Animal Care, Ottawa, Canada, 1991.
26. Experiments on Animals Act, 1997, Section 1a, Veterinary Public Health Inspectorate, Rijswijk, The Netherlands.
27. G. Kolata, *N.Y. Times*, December 23, pp. A1, A20 (1999).
28. J.A. Mench, in J.D. Murray et al., eds., *Transgenic Animals in Agriculture*. CAB International, New York, 1999, pp. 251–268.
29. *Animal-to-Human Transplants*, Nuffield Council on Bioethics, London, 1996, pp. 49, 51–52.
30. Advisory Group on the Ethics of Xenotransplantation, *Animal Tissue into Humans*, Report, Her Majesty's Stationary Office, Norwich, UK, 1996, pp. 78–80.
31. Available from: David Nunn Ltd., Shrewsbury, Shropshire, UK.
32. Available from: Bioserve, Frenchtown, NJ; Primate Products, Redwood City, CA; and KLASS San Jose, CA; *www.nal.usda.gov/awic/pubs/enrich/supplier.htm*
33. *Animal Welfare Information Newsletter* **7**(3–4), 1–8 (1996–1997).
34. Available from: HAKA, Josef Haeufele GmbH & Co. KG, Robert-Bosch-Str. 6-9, D-89155 Dellmensingen, Germany, Tel: +49/7305/96100; Fax: +49/7305/9610-40.
35. *Animal Welfare: Xenotransplantation*, Imutran Ltd., Cambridge, UK, 1999.
36. Welfare of Livestock Regulations, 1994 (S.I. 1994/2126)-Schedule 3, Home Office, London, UK.
37. *Guide for the Care and Use of Agricultural Animals in Agricultural Research and Teaching*, 1988, rev. ed., Federation of Animal Science Societies, Champaign, IL, 1999.
38. J. Webster, *Animal Welfare: A Cool Eye towards Eden*, Blackwell Science, Oxford, UK, 1994.
39. Z. Lou and J.F. Hurnik, *J. Animal Sci.* **72**(10), 2610–2616 (1994).
40. D. Fraser, P.A. Phillips, and B.K. Thompson, *Livestock Prod. Sci.* **20**, 249–256 (1988).
41. D.B. Morton and P.H.M. Griffiths, *Veterinary Rec.* **116**, 431–436 (1985).
42. D.B. Morton, *ATLA* **18**, 29–39 (1990).
43. FELASA Working Group on Pain and Distress, *Lab. Animals* **28**, 97–112 (1994).
44. P. Flecknell, *AWIC Newslett.* **8**(3–4), 8–14 (1998).
45. M.H. Ullman-Culleré and C.J. Folz, *Lab. Animal Sci.* **49**(3), 319–323 (1999).
46. D.B. Morton, in F.F.M. van Zutphen and M. Balls, eds., *Animal Alternatives, Welfare and Ethics*, Elsevier Science, Amsterdam, The Netherlands, 1997, pp. 235–241.
47. D.B. Morton, in *Proc. Joint ANZCAAT/NAEAC Conference on Ethical Approaches to Animal-Based Science*, Auckland, New Zealand, September 19–20, 1997, pp. 131–148.
48. V.A. Hampshire, J.A. Davis, and C.A. McNickle, *Lab Animal* **29**(5), 40–45 (2000).

49. D. Mellor and D.B. Morton, *ATLA* (in press 2000).
50. G. Griffin and H.B.W.M. Koeter, in *Proc. Third World Congress on Alternatives and Animal Use in the Life Sciences*, ATLA (in press 2000).
51. *Choosing an Appropriate Endpoint in Experiments Using Animals for Research, Teaching and Testing*. Canadian Council on Animal Care, Ottawa, Canada, 1998.
52. C.F.M. Hendriksen and D.B. Morton, in *Proc. International Conference*, November 22–25, 1998, Zeist, The Netherlands Royal Society of Medicine, London, 1998.
53. V.C. Delpire, T.B. Mepham, and M. Balls, *ATLA* **27**, 869–881 (1999).

See other entries ANIMAL, MEDICAL BIOTECHNOLOGY, LEGAL, LAWS AND REGULATIONS GOVERNING ANIMALS AS SOURCES OF HUMAN ORGANS; ANIMAL, MEDICAL BIOTECHNOLOGY, POLICY, WOULD TRANSGENIC ANIMALS SOLVE THE ORGAN SHORTAGE PROBLEM?; FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE; see also RESEARCH ON ANIMALS entries; TRANSGENIC ANIMALS: AN OVERVIEW.

S

SCIENTIFIC RESEARCH, ETHICS, SCIENTIFIC MISCONDUCT

KENNETH J. RYAN
Harvard Medical School
Boston, Massachusetts

OUTLINE

INTRODUCTION

Scientists have traditionally tried to shield their work from ethical scrutiny under the guise that they are involved in a value-free pursuit of the truth, but in the second half of the twentieth century this position became untenable. Ethical concerns were expressed by the public as well as some scientists about such issues as the dangers of atomic warfare, nuclear radiation, human subjects biomedical and behavioral research, recombinant DNA, assisted reproductive technologies, and global warming (1–3). In view of the rapid and massive intrusion of scientific discoveries and their application into human affairs for both good and ill, it was becoming difficult to maintain that science was essentially value free. As Stephen Toulmin suggested "it is meanwhile becoming clear that the professional organization and priorities of scientific work can no longer be concerned solely with considerations of intellectual content and merit, as contrasted with the ethical acceptability and social value, either of the research process itself, or its practical consequences" (1). In addition, by the latter half of the twentieth century a good deal of scientific research was being supported by public funds especially in the United States. This introduced political as well as social concerns with how, and for what purposes, scientists pursued their research. With such ethical concerns already in place, it is not difficult to understand the intense dismay that greeted at first, sporadic and then a flood of reports beginning in the 1970s of scientific research papers that were fraudulent, particularly when they involved clinical research and matters related to human well-being. The inquiries into this scientific misconduct raised a series of questions about (*1*) the extent of the problem, (*2*) the definitions that should be applied to misconduct in contrast to scientific error or chance, (*3*) the plight of an accused scientist or the whistleblower, (*4*) the nature of scientific discovery and the role of ethics in the education and practice of scientists, (*5*) the role of the government and public in the oversight process, and (*6*) the roles of the academic community, universities, industry, research laboratories, and the individual scientist in dealing with the problem of scientific misconduct (4).

SCIENTIFIC MISCONDUCT

Incidence and Types of Scientific Misconduct

The incidence of fraud or misconduct in science has in the past generally been considered by scientists, science historians, and sociologists to be rare and inconsequential. The scientist has traditionally been regarded as the seeker of true knowledge. Advances in science and technology have resulted in space exploration, modern forms of transportation and communications, and benefits to humankind in agriculture and medicine. The traditional view is that if there is fraud in an important area of science, it will be uncovered quickly by failed attempts to repeat it by other scientists or by inside information from the laboratory concerned. On the other hand, if the published research work is in an unimportant area of science, or is of a trivial nature, little damage is supposedly done to science or the public except for the cluttering of the literature. This line of reasoning has become suspect. The "romantic ideal" that scientists seek the truth and that errors or frauds are uncovered by a continual process of repeated studies has been challenged. While the ideal model may prevail over the long term for primary or critical data, there are many exceptions. A Noble laureate and director of the National Institutes of Health, Harold Varmus, challenged the idea that science is self-correcting in 1994 at a National Academy of Science convocation on scientific misconduct (5). Kiang suggested that errors (or fraud) in the archival record often are never corrected, simply ignored and eventually forgotten (6). The question of the extent of science fraud is, however, of long-standing. Robert Merton, a sociologist of science wrote in 1957 that he believed fraud was rare in science. For an earlier period, he quoted Darwin who "knew of only three falsified statements" in all of science. From the same period, Merton described Charles Babbage's 1830 "inventory of fraud in science" which included cooking and trimming of research data (7). In the 1990s several reports suggested that scientific misconduct was increasing and more widespread than previously believed and that such fraud in science was potentially damaging to both science and the public interest (8). The American Association

for the Advancement of Science (AAAS) surveyed its membership and reported in 1992 that 27 percent of those replying had personal knowledge of an average of 2.5 cases of *suspected* fabrication, falsification or theft of research in the prior ten years (9). Thirty-seven percent felt the incidence of misconduct was rising. In 1993 Judith Swazey and colleagues reported on a survey that involved 2000 graduate students and 2000 faculty in chemistry, civil engineering, microbiology, and sociology. While the study could not report on the frequency of misconduct, it did provide insight into the rates of exposure of students and faculty to various forms of misconduct (10). The authors reached the conclusion that although science misconduct is not rampant, it certainly is not rare. Between 6 and 9 percent of students and faculty reported knowledge of faculty who falsified or plagiarized data while one-third of faculty reported incidents of student plagiarism. Review of audits by the Food and Drug Administration of drug research conducted between 1977 and 1988 revealed problems in 12 percent of studies before 1985 and 7 percent thereafter. The problems uncovered included failure to perform studies for which results were given or changing of data (11).

That fraud is not always uncovered quickly and is not always innocuous was demonstrated in the criminal conviction of the psychologist Stephen Breuning in 1988 some years after he reported on the clinical effects of drugs on hyperactive retarded children. Much of his research was simply never performed, although his "work" inappropriately influenced the care of the mentally retarded (12). In the industrial arena there was the case involving the officers of Industrial Bio-Test Laboratories, Inc. who were convicted of fraud in the reporting of toxicity data on which drug companies and pesticide manufacturers relied for the effects of drugs and pesticides on laboratory animals. This ultimately affected the review and approval process of pesticides and drugs by the Environmental Protection Agency and the Food and Drug Administration (12).

When oversight of research misconduct was formalized by the National Science Foundation (NSF) and the National Institutes of Health (NIH) in the 1990s, they reported that some 20 to 40 alleged cases a year of various forms of misconduct reached the investigational stage for each agency (13). Whether one is alarmed by the numbers may depend on whether one takes solace in the small percentage of federal grantees actually involved in misconduct or whether one is concerned with a potential corrosive effect of even a small percentage of cases on the education of research scientists and the trust of the public. The actual numbers of misconduct cases also ultimately depend on what is included as research or scientific misconduct.

The types of research misconduct that have been chronicled in books and articles have varied widely and given rise to the confusion about the frequency and seriousness of the problem. One distinction that has been made is between science in general and research specifically, as when the Public Health Service changed the titles of its oversight committee from the Office of Scientific Integrity to the Office of Research Integrity in 1992 (14). The government makes the obvious distinction in its oversight role between research supported by federal funding and that supported from privates sources, as well as research performed to comply with regulatory agencies. The private sector is obviously responsible for research not funded or regulated by the government and the definition of misconduct could be more rigorous (or less) than federal standards. Other distinctions that have been suggested are between basic and applied research and research involving human subjects. Human Subject Research has generally been under separate government oversight over matters of informed consent, risks to subjects, and equity as distinct from concerns about fraud (15). The main emphasis in government definitions with respect to fraud has been on the scientific record with concern for the information published, formally presented at scientific meetings, or offered in progress reports to grants. In this case examples of misconduct have included reporting on work never performed or manipulating research data to obtain desired results as in adjusting points on a curve, omitting points that fall off the curve, or omitting data that make the conclusions less appealing or untenable. Misconduct might also include the use of inappropriate statistics to achieve significance that is not otherwise demonstrable. While sloppiness or gross errors in research might not warrant government sanctions, they might limit academic advancement or job security (13).

The other major source of misconduct involving the scientific record is in the acts of plagiary, which involve misappropriating the words or ideas of another scientist without giving credit for the source and representing them as one's own. The other forms of misconduct fall into such categories as sexual harassment of colleagues or students in the research environment, secreting data or notebooks that are not exclusively ones own, sabotaging or destroying another scientist's equipment or experimental results. Specific misconducts that might or might not be lumped with protecting the research record include failing to follow federal regulations on the protection of human subjects, treatment of animals, environmental protections especially with radioactivity, or appropriate use of grant funds. The Office of Inspector General of the NSF indicated in 1991 that of the cases of alleged misconduct: 20 were about plagiarism, 9 were for fabrication and falsification, and 8 were in the other "serious deviations" category (13).

The choice of what to include in one's definitions or categories obviously depends on anticipated use. If the list is for educational purposes in academic institutions, inclusiveness might be the goal. If the objective is to check compliance with federal funding or regulatory agencies, then definition of the misconduct is expected to be more restrictive and precise.

Defining Misconduct

Formal definitions for research misconduct have originated largely through government oversight of funded or regulated research. The first definition was published in 1986 in the NIH Guide to Grants and Contracts, as part of an interim policy before final regulations were formulated. Prior U.S. congressional hearings on scientific misconduct resulted in a provision of the Health Research

Extension Act of 1985 which required that research institutions establish procedures to investigate scientific fraud, and that the director of the National Institutes of Health develop an administrative method to respond and deal with it. The first definition of misconduct in science was thus established as "serious deviation such as fabrication, falsification, or plagiarism, from accepted practices in carrying out research or reporting the results of research; or material failure to comply with Federal requirements affecting specific aspects of the conduct of research, for example, the protection of human subjects and the welfare of laboratory animals" (16). The scientific community had an opportunity for critical commentary in response to the publication of proposed policies and procedures of the Public Health Service (PHS) in the Federal Register in 1991 and prior to that to advance the final Notice of Proposed Rulemaking in 1988. A similar procedure was followed by the National Science Foundation (NSF) in 1987 to receive public commentary and establish a definition covering grants under its jurisdiction. The final NSF definition developed in 1991 closely followed that of the PHS: "Misconduct in science and engineering is fabrication, falsification, plagiarism, or other serious deviation from accepted practices in proposing, carrying out, or reporting results from activities funded by NSF: and retaliation of any kind against a person who reported or provided information about suspected or alleged misconduct and who has not acted in bad faith." The second portion of this NSF definition covering retaliation was criticized by outside commentators as not being part of scientific misconduct, but the view of the NSF was that retaliation was a serious deviation from accepted practices and would work against scientific integrity. This clause has remained. The PHS has not had retaliation in its definitions but was later called upon by Congress to develop procedures for protecting whistleblowers. It was also noted that the Whistleblower Protection Act of 1989 protects only federal employees and not whistleblowers who are employees of the grantee institutions. The PHS definition of misconduct was later changed in response to some of the commentary it received. The PHS definition established in 1989 is: "Misconduct or misconduct in science means fabrication, falsification, plagiarism, or other practices that seriously deviate from those that are commonly accepted within the scientific community for proposing, conducting, or reporting research. It does not include honest error or honest differences in interpretations or judgments of data." The final sentence was included in response to commentary from the scientific community. On the other hand there has been no dropping of the clause "practices that seriously deviate from those that are commonly accepted" which has been repeatedly criticized vigorously by scientists for both the PHS and NSF definitions as too vague and a potential basis for unfair treatment of scientists (16).

The U.S. Congress as part of the 1993 NIH Revitalization Act created a Commission on Research Integrity (CRI) with a mandate to make recommendations on a definition for research misconduct. Research misconduct had been a continuing subject for congressional hearings in the 1980s and 1990s, and there was apparent dissatisfaction by the Congress over the way the scientific community, academic institutions and federal agencies were dealing with the problem. The Commission was a 12-member public advisory body composed of scientists, lawyers, sociologists, and ethicists that was asked to make recommendations to the Secretary of Health and Human Services on the definition of research misconduct, an assurance process for compliance of institutions with federal regulations, administrative processes for dealing with misconduct, and, finally, recommendations to protect whistleblowers.

The CRI report did make recommendations to change the definition of research misconduct in an attempt to "provide vital guidance for personal and ethical judgments and decisions concerning the professional behavior of scientists and to provide a legal framework for formal proceedings" (14). From the public record and testimony, the CRI developed a sense of the types of research misconduct which were most common and needed addressing in a definition. The Commission felt that a fundamental principle for scientists was to be truthful and fair in the conduct of research and dissemination of their findings. The recommended definition of the CRI was that "research misconduct is significant misbehavior that improperly appropriates the intellectual property or contributions of others, that intentionally risks corrupting the scientific record or compromising the integrity of scientific practices. Such behaviors are unethical and unacceptable in proposing, conducting, or reporting research or in reviewing the proposals or research reports of others." The commission gave as examples of research misconduct to substitute for fabrication, falsification and plagiarism; the categories of misappropriation, interference, and misrepresentation were explained as follows:

> *Misappropriation* is intentional or reckless plagiary. This means presentation of the documented words or ideas of another without attribution appropriate to the medium of presentation, or to the use of any information in breach of any duty of confidentiality associated with the review of a manuscript or grant application.
>
> *Interference* is intentional and unauthorized taking, sequestering, or material damage to research related property of another. This includes, without limitation, the apparatus, reagents, biological materials, writings, data, hardware, software, or any other substance or device used or produced in the conduct of research.
>
> *Misrepresentation* is an intention to deceive or reckless disregard for the truth. This means stating or presenting a material or significant falsehood or omitting a fact so that what was stated or presented as a whole stated or presented a material or significant falsehood (14).

This somewhat complex and legalistic definition was not well received by the scientific community, but it engendered a broad debate on research integrity. The Commission's definition was in fact a more explicit rendering of fabrication, falsification, and plagiary with a substitution of interference for the "serious deviation from accepted practices" clause that many found objectionable (14,16). The CRI definition did introduce the concept that omission of critical data should be identified as misconduct

and made the concept of intent explicit, something that has been discussed by Dresser (17). The Commission further defined two other forms of professional misconduct that included obstruction of research misconduct investigations, which still is in the NSF definition, and noncompliance with federal regulations, which had been in an early PHS definition.

In line with a recommendation of the Commission, a federal interagency task force was appointed to develop a common definition of research misconduct for all government departments. By the end of 1999 they could not agree on a common definition; the secretary had not recommended the CRI definition, and the NSF and PHS definitions from 1991 remained in force.

Plight of the Whistleblower, Plight of the Scientist

The plight of the whistleblower has been championed by the distinguished scientist and journal editor John Edsall who believes that whistleblowers are necessary for the maintenance of honest science. He described cases occurring over a 30-year period in which major harassment and difficulties were faced by the whistleblower, often more severe than the person accused of the misconduct (18). A book entitled *The Whistleblowers* reported on a 6-year study of 64 individuals (19,20). A detailed report on the consequences of whistleblowing for the whistleblower and the exonerated accused was also prepared by the Research Triangle Institute for the Office of Research Integrity in 1995 (21). Sixty-nine percent of whistleblowers and 60 percent of those exonerated of misconduct suffered negative consequences, usually more severe early in the process of investigation and lessening with time. The converse of this is that significant numbers did not suffer adverse outcomes. The public sympathy engendered for the whistleblower from congressional hearings in the 1990s contributed to the inclusion of a mandate to address their protection by the CRI (14). At that time there was much publicity about Margot O'Toole, a junior scientist and whistleblower who had problems with Thereza Imanishi-Kari and David Baltimore at the Massachusetts Institute of Technology. Tom Devine likened whistleblowing to professional suicide. He chronicled the tales of 20 witnesses who appeared before the CRI. They complained of censorship, loss of job, academic expulsion, retaliatory investigations, denial of access to their data and laboratories, as well as threat of deportation and physical harm (22). To quote Tom Devine: "Everyone pays lip service to the ideal that science is the search for the truth, and scientific integrity is a concern for all. But whistleblowers actually live those values." In response to the CRI mandate, a Whistleblower's Bill of Rights was proposed as part of the CRI report (14,22). Even though the National Academy of Science advises beginning researchers that they have an "unmistakable obligation to act" if they suspect someone is violating the ethical standards of science, there is much cynicism. Often the whistleblower must wait years to be vindicated or must resort to the courts to get attention and redress. The scientific community felt that the Whistleblower's Bill of Rights provided too little concern for the accused scientist (4). There is always the possibility that the whistleblower is wrong or is not acting in good faith.

The plight of the scientist has been typified by two cases settled in 1997 in which after many years of turmoil, and adverse effects on their careers and reputations, the accused scientists were vindicated. The case of Thereza Imanishi-Kari, who was a collaborator of David Baltimore, was settled by an appeal board of the PHS some 10 years after the first charges of misconduct were made. No misconduct was ultimately found although there were criticisms of the research record keeping by Imanishi-Kari (4). All this was chronicled in a *New Yorker* article entitled: "The Assault on David Baltimore" (23) and a book, *The Baltimore Case*, both by Daniel Kevles (24). Baltimore, who initially suffered for the defense of a scientific colleague, became a hero in the Kevles book, and Margot O'Toole, the whistleblower, was recast by Kevles in a less sympathetic light. It is not certain that Kevles gave an evenhanded recounting of the controversy, and at the very least the maintenance of research records was inadequate and contributed to the problem. It is likely that Baltimore, Imanishi Kari, and O'Toole were all victims of a bad system for defining and dealing with misconduct by the scientific community, academic institutions, and the government (4). The other case illustrating the plight of the accused scientist was that of Bernard Fisher who ultimately gained vindication by going to court to receive an apology from the government for its inept oversight and an apology and financial settlement from his institution, the University of Pittsburgh. Fisher was the director of an interinstitutional NIH-funded clinical cancer trial. It was reported that a physician from one of the participating hospitals entered ineligible subjects into the study, and Fisher was caught up in the question of how and when to reveal this information, although it had little effect on the conclusions drawn from the study (25). The CRI also heard testimony from many aggrieved scientists, even professors who claimed to have been unjustly accused of misconduct and had their careers destroyed. This is reviewed in the report noted above on those accused but exonerated of scientific misconduct (21).

ETHICS AND SCIENCE

History, Sociology, and Philosophy of Science: Looking for Causes of Misconduct

The history, sociology, and philosophy of science are so interconnected as to make distinction between these disciplines difficult, but each has contributed to the literature about the values held by the scientist in conducting research, in elaborating scientific theories, in relating to other scientists, and in relating to society at large. Ultimately these values held by scientists contribute to the character of science, and they are a promising place to look for the factors that motivate scientists and the means by which progress in science occurs or is hindered. The so-called normative structure of science with its virtual absence of fraud due to rigorous policing by science itself enunciated in the 1940s by Merton may be an unreal perception of how science works today or for that matter how it operated previously (26).

As noted earlier, the ideal of an incorruptible science based upon timely self-correction is questioned in the real world. Harriet Zuckerman has considered the causes of misconduct under three headings: individual psychopathology, anomie, and alienation (26). The cause of misconduct had been attributed by many scientists simply to individual psychopathology, but this has been criticized as "too convenient and self-serving" to explain what is going on. It is not only the individual guilty of misconduct but also the laboratory environment in which he or she works that affects behavior. Anomie is a theoretical sociological construct for deviance based on a high value being placed on a goal for which the means are not readily available, inducing people to choose dubious routes to try to achieve the highly desired end. Another theoretical cause of misconduct is alienation, which occurs when there is a disconnect between the daily laboratory work and the ultimate goals of the research. Also, when there are gross differences in reward and recognition among the laboratory team, alienation, personal animosity, and misconduct are more possible (26). It reached a point in one prominent laboratory in Rockefeller University where attempts at poisoning were reported (27). The rise in misconduct is also believed due to the change in the way scientific research is organized and funded. Research in the last half of the twentieth century has become more costly, competitive, complex, specialized, and collaborative at the same time that being successful in research creates the opportunity for academic advancement, rewards, and recognition (13). Research may be attracting scientists largely because of external and secondary rewards of fame and fortune, which can corrupt, rather than the incorruptible primary pursuits of asking questions and seeking knowledge for its own sake.

Another interesting sociological perspective on scientists was offered by Bernard Barber in his 1961 *Science* article, "Resistance by Scientists to Scientific Discovery" (28). Although it is generally accepted that there may have been outside religious, political, and ideological impediments to progress in science over the years, the idea that science has cultural and social forces within itself that resist progress is a novel concept. Barber noted that new ideas may be resisted because they clash with existing substantive concepts like the notion of the irreducibility of the atom which clashed with the discoveries of electrolytic dissociation, the discovery of X rays, and the theory of the electronic composition of the atom, all of which were resisted when first proposed. Methodological concepts based on the senses or models of mathematics clashed with the ideas of analyzing colors with prisms, radioactive measuring, electromagnetic theory, or experimentation and provided resistance to their introduction. Religious beliefs of scientists were the basis for resistance to Darwin's theories. Low professional standing can impede acceptance of work, as occurred with resistance to Mendel's theory of inheritance. Such social forces working within science may give rise to suspicions of misconduct when none exist merely because it is conceptually difficult to accept the results of a groundbreaking study. Rivalries based on specialization, societies, schools, or seniority have also impeded acceptance of scientific work, but Barber ends on an optimistic note that scientists even with their human faults are in his view more objective and open-minded than society in general.

A major innovation in the philosophy of science was introduced in 1962 by Thomas Kuhn in his *The Structure of Scientific Revolutions*, which introduced the view that discovery and progress in science are not completely rational and that science is not a steady progression toward an "objective" truth (29). Taking into account the work of Bernard Barber and his own studies in the history of science, Kuhn created a new vocabulary and definitions to cover "normal science," paradigms, incommensurability between paradigms, paradigm shifts, and scientific revolutions. Normal science works within a paradigm with shared rules and standards and is closest to the traditional view of how science works. Normal science works on a historical record, but it does not bring about striking new discovery. When anomalies occur and questions arise that cannot be answered within normal science, there may be a paradigm shift or revolution with different rules and standards and a radical break with the past. The switch from the vision of Ptolemy to Copernicus, from Newton to Einstein, and from creationism to Darwin have all been given as examples of scientific revolutions. Again, the workings of science may raise suspicions of misconduct within the rubric of practices that seriously deviate from those commonly accepted within the scientific community unless there is a broad understanding of the history and sociology of science, of how research is actually conducted in specific fields, and of how theories are developed and tested.

History of Scientific Misconduct

There have over the years been many reports of suspected or proven fraud or misconduct in science, but it is best to divide the cases temporally based on whether they occurred in remote or recent times. This is a selected list of cases and is not meant to be exhaustive (30–34).

Science Misconduct in Ancient Times. Science misconduct as ancient history is probably more speculative than factual, and there are usually accusers and defenders for almost every case involving the legendary scientists of the past. The contributions of most of these scientists are little diminished by these accounts, and there is little evidence of any intent to deceive in cases before the twentieth century. For details, reference should be made to listed sources (30–34).

Claudius Ptolemy of the second century A.D. is accused by both French and American astronomers of not making measurements claimed but extrapolating them from an earlier Greek astronomer or of deriving the data from theoretical projections rather than from personal observation. Historians dispute this suggesting instead that his observations were adjusted in keeping with the standards and methods then in use.

Galileo Galilei of the seventeenth century has been accused of not having performed experiments as described but of creating data to conform to his theories.

Isaac Newton, the great physicist of the late seventeenth and early eighteenth centuries, is believed to have used "fudge factors" after the fact to adjust data to

meet theoretical predictions. While some have termed this fraud, others have characterized this as making approximations to test if a theory is in fact feasible.

Gregory Mendel published his work in 1865 that is the basis of modern genetics. It is claimed that the reported observations of the frequency of inherited traits and the expected values are too good to be true. This has been attributed to the manipulation of data by an obliging assistant, experimenter bias, or a more innocent difficulty in sorting the categories.

Robert Millikan published papers in 1910 and 1913 which won him the Nobel Prize in 1923 for determining the electric charge of the electron. Although Millikan claimed to have published all his data points, review of his notebooks reveals that he was selective and left out one-third of his observations. A scientific rival Felix Ehrenhaft of Vienna who lost out in the recognition of the Nobel Prize found fractional charges rather than the exact multiples described by Millikan. With the newer knowledge of subelectronic particles, it is possible that Millikan was wrong not only in his conclusion but also in his selection of data points to publish (31).

Sir Cyril Burt was a famous British psychologist who published studies in the 1950s on identical twins raised together and apart and concluded that IQ is largely inherited. It is claimed that many of the studies were never done and that the conclusions were designed to support his belief that intelligence is determined more by genes than environment (32).

Science Misconduct in Recent Times. Science and research misconduct described in the last 40 years of the twentieth century is based on more substantial evidence than cases dug up from ancient history. In more recent times investigations of misconduct are usually based on laboratory records and often on confessions.

Harold Bates worked in the laboratories of Mel Simpson, Professor at Yale and then with Professor and Noble laureate Fritz Lipmann at Rockefeller University, and published papers with them in the *Journal of Biological Chemistry* in 1960. A paper on the biosynthesis of cytochrome C coauthored with Simpson and a paper on the biosynthesis of glutathione coauthored with Lipmann had to be retracted because the work could not be verified (18).

George Webster, an established investigator of the American Heart Association, joined the Enzyme Institute of the University of Wisconsin and in 1965 published a paper in the *Journal of Biological Chemistry*. His work was challenged by Efraim Racker whose own work was at odds with Webster's. Webster announced that he did not have the original data to back up his work, although the paper has never been formally retracted (18).

William Summerlin was a dermatologist who joined the famous immunologist Robert Good at the University of Minnesota and moved with him to the Sloan Kettering Institute in 1973. Summerlin claimed from his research studies that he could treat mouse skin and human corneal tissue by culture outside the body to abolish immunological rejection when the treated tissue was subsequently transplanted to other unrelated animals. He claimed to be able to treat skin tissue from black mice, which would keep it from being rejected when transplanted to a different strain of white mice. In 1974 it was discovered by a technician that Summerlin had used a black pen to darken the spots on white animals and the credibility of all his work collapsed. The one mouse with a successful transplant was shown to be a hybrid, which would be expected to have a graft survive from the donor strain. The claim of successful transplantation of human cornea to rabbit eyes was also shown to be a deception. After this case there was a good deal of soul-searching about the pressures put on young scientists in large, highly publicized laboratories and the fact that their work is too readily "accepted" by their supervisors who have expectations of the kind of result they are seeking.

John Long was research pathologist at the Massachusetts General Hospital in the laboratory of Paul Zamecnik and rose to the rank of associate professor. He claimed to have a human tumor cell line from patients with Hodgkin's disease. In 1979, when collaborators asked for primary data on some joint experiments, Long was found to have altered data. He later admitting falsifying the results and it was also discovered that his so-called human Hodgkin's tumor cells were derived from a lymphoid cell line of the owl monkey, possibly by contamination. Long resigned his laboratory position and returned to clinical practice.

Vijay Soman was a scientist from India who started as a postdoctoral fellow in the laboratory of Philip Felig, professor of endocrinology at Yale and worked his way up to the rank of associate professor. In 1980 he and Felig published a paper in the *New England Journal of Medicine* that was shown to be partially plagiarized from a manuscript Felig was asked to review and rejected. Felig had given the paper to Soman to read because it was in their area of research. Soman not only plagiarized exact wording from the rejected manuscript; he made up most of the data he submitted. Review of 14 prior publications by Soman with Felig revealed that only 2 had supporting data and 12 of the questionable papers had to be retracted. Soman returned to India, and Felig who had gone on to be chairman and professor of medicine at Columbia, lost his position and returned to Yale.

Mark Spector was a promising graduate student in the laboratory of Efraim Racker at Cornell University when in 1981 he reported an exciting sequence of events, a cascade hypothesis for the process of transformation of normal animal cells into tumor cells. A collaborator became suspicious when only Spector could repeat certain key gel experiments crucial to the hypothesis. It was found that Spector's work was largely a fabrication. After this, Spector's credentials were checked, and it was found that he had a prior record of forgery and previous research work could not be repeated. The irony of this case is that it was Efraim Racker who exposed the fabrication of George Webster some 20 years earlier as noted above, only to be deceived himself in his own laboratory (18,30).

John Darsee was a promising young cardiologist who worked in the laboratory of Eugene Braunwald of Harvard and the Brigham & Women's Hospital. In 15 months of

work at Harvard, Darsee contributed 5 papers coauthored with Braunwald and many abstracts to national meetings. When asked by the laboratory head to submit raw data on some dog experiments, Darsee was observed by other lab workers in the act of fabricating the data in a contrived experiment. Rather than being an isolated event as first believed, it turned out that Darsee had a long trail of fraudulent research stretching back to his medial school and college days. Eight of 10 publications that Darsee had released before coming to Harvard had to be withdrawn or corrected. Darsee left Harvard and obtained a clinical post.

A junior colleague of Dr. Francis S. Collins, head of the Human Genome Project, was accused in 1996 of fabricating data in five research papers on leukemia, which had to be withdrawn. The colleague was identified by the *New York Times* as Amitov Hajra, a graduate student at the University of Michigan who worked in Dr. Collins's laboratory at the NIH. The fraud was uncovered when the reviewer of a subsequent paper submitted for publication by Collins and his student in the journal *Oncogene* questioned the data and suggested intentional deception. Ironically, what was obviously suspicious to an outside reviewer was missed by Collins and others in the laboratory (4,35).

A scientist from Immunex, a biotechnology company, was accused of plagiarizing the gene structure for an interleukin molecule that was obtained by reviewing an article for the journal *Nature* authored by scientists from a competing company Cistron Biotechnology. The paper was rejected, but it is claimed that the reviewer then patented the gene structure presented in the article. The controversy resulted in a suit alleging misappropriation of data during peer review. The case was settled out of court, but at least illustrates the occurrence of alleged misconduct in the biotechnology industry as well as in academic laboratories (36,37).

Plagiarism is a frequently reported form of misconduct. A case often cited is that of Elias Alsabti, an Iraqi who worked in both England and the United States and simply copied previously published works and sent them to obscure journals under his name or took progress reports or grant applications of others as bases for articles. In all he had some 60 fraudulent publications between 1977 and 1980. When found out, he simply went to new positions to ply his deceptions. The checking of references and his past as he went from position to position was obviously seriously deficient.

In 1987 Shervert Frazier, a professor of psychiatry at Harvard, admitted to plagiarizing works of others for review articles between 1960 and 1975. He subsequently resigned his academic post but continued to practice psychiatry at McLean Hospital (33).

The foregoing list of misconduct cases from distinguished laboratories is merely a selection of the many that could be described involving clinical and basic research. While the emphasis is on biological research, misconduct has been a problem in all fields of science, engineering, and technology. It is interesting that in many cases the misconduct was carried out by a junior or mid-level scientist who was inadequately supervised in a high-powered high-profile laboratory. Many of the transgressors were considered brilliant, hard-working, technically skilled, and almost too good to be true. They worked harder and longer and published more than is usually possible or reasonable. These stories are reminiscent of the adage for consumers about deceptive advertising — that attractive offers that seem too good to be true probably are too good to be true.

GOVERNMENTAL ROLES IN SCIENCE AND MISCONDUCT

During World War II scientists were involved in many research projects sponsored by governments in pursuit of advantages in the war effort. Out of such work came the manufacture of Penicillin, the invention of radar, advances in transportation computers, and communications and the use of nuclear energy for the atom bomb. After the war, government science advisor and MIT professor Vannevar Bush in his report to President Truman, "Science: The Endless Frontier" outlined the benefits that would flow from continued governmental support for basic research (38). As a consequence a new system was established that sited the conduct of research at universities, created project grants that could be awarded to scientists for research in their own laboratories at universities, and developed advisory and review committees of private scientists to serve the government on a part-time basis, which gave rise to the peer review system for approving and giving priority to grant applications. David Guston characterized this general arrangement as the "social contract for science" (39). Initially scientists were hesitant to accept government funds lest onerous restrictions and oversight be imposed on their academic and scientific freedoms. The "contract" provided for federal support, but with the responsibility for oversight left to the traditional mechanisms for academic governance at the private universities and institutions. There was considerable faith in the ability of scientists to regulate themselves and to ensure the integrity of the scientific process. Although cases of misconduct in science as described earlier were being reported during the 1960s and 1970s, there seems to have been little interest in the Congress to get involved, trusting instead the mechanisms ordinarily used by scientists and universities to discipline their members. Ironically, the reports of incompetent university and federal investigations of misconduct triggered the interest of congress, rather than the occurrence of the misconduct itself (4,40). The first congressional hearings for oversight on scientific misconduct occurred in 1981, chaired by Albert Gore in the House of Representatives and by Orrin Hatch in the Senate. Gore's committee looked into the Darsee affair and heard testimony from Drs. Philip Felig and John Long about their reactions to their experiences, described in the cases stated above. Philip Handler, president of the National Academy of Sciences, testified before the Gore committee that the problem of science misconduct had been grossly exaggerated by the press and that scientists should be allowed to take care of it by themselves. Senator Hatch was concerned with the institutional responses to the cases and with waste and fraud. He was particularly incensed that a

cancer scientist Marc Strauss from Boston University received new funding from the National Cancer Institute while he was under investigation for prior fraudulent practices (34,40). The Gore hearings did lead to legislation on misconduct which was included in the 1985 NIH reauthorization bill. The provision in the bill dealt with the responsibility of universities to develop mechanisms for dealing with misconduct. Congressional hearings were held again in 1988 when the late Ted Weiss from the House Committee on Government Operations looked into abuse of whistleblowers and delays in dealing with allegations of scientific misconduct at the institutional level. Meanwhile John Dingell, chairman of the House Committee on Energy and Commerce Subcommittee on Oversight and Investigation, started exploring the Thereza Imanishi-Kari, David Baltimore case described earlier. These latter hearings would go on for several more years until 1993 and involve Representative Dingell and David Baltimore in acrimonious exchanges on the congressional committee's aggressive investigation (4). In response to the congressional hearings and wide spread coverage in the press about the plight of whistleblowers and the cases of scientific misconduct, the PHS in 1989 set up an Office of Scientific Integrity (OSI) within the NIH to receive reports from institutions and conduct investigations on alleged misconduct. An Office of Scientific Integrity Review (OSIR) was set up within the Office of the Assistant Secretary for Health as an oversight function for the OSI (16). Dissatisfaction with the way this arrangement dealt with the Iminishi-Kari, David Baltimore case and with one involving Robert Gallo of the NIH brought criticisms of these offices. Congressional, scientist, and press criticisms of the OSI and OSIR were that they had inconsistent policies and vague rules, that they were biased against defendants, and used illegal procedures. The most important issue was they were not bringing the high-profile cases to a satisfactory conclusion as far as Dingell's committee was concerned (4,40). In 1992 the Department of Health and Human Services in response to the concerns merged the OSI and OSIR into one Office of Research Integrity (ORI) and placed it in the Office of the Assistant Secretary of Health (13,14). In addition an appeals process for those found guilty was made available with a trial-like hearing before a Research Integrity Adjudication Panel. These panels were appointed by the DHHS Appeals Board and included a scientist. In 1995 the ORI was moved to the Office of the Secretary DHHS (14). Concerns about the process of dealing with alleged misconduct cases continued to be expressed by the scientific community and the congress. In 1993, as part of the NIH Revitalization Act, the Congress as noted earlier (14) authorized a Commission on Research Integrity (CRI). The mandate was to consider a new definition of research misconduct (described earlier) to recommend an assurance process for institutions to comply with PHS regulations on misconduct, to recommend government and institutional policies to deal administratively with investigations, and to make recommendations on the protection of whistleblowers (14). From 1987 the National Science Foundation (NSF) handled cases of misconduct out of their Office of Inspector General, which conducted any investigation. An adjudication process separate from the investigation was available with due process rights. The NSF drew much less criticism than the PHS for the way misconduct cases were investigated. It is not clear whether the more favorable reaction by scientists to the way misconduct cases were handled by NSF was due to the types of cases chosen or the process used. For both the PHS and NSF, investigations were carried out by the scientist's own institution whenever possible as long as they were handled competently (14,40). An interesting phenomenon was the activities of two NIH scientists, Ned Feder and Walter Stewart (41,42). For about 10 years the pair devoted their time, without NIH authorization, to checking out accusations of fraud and plagiarism in science research at the NIH and elsewhere, and urging action by anyone who would listen to them. They championed the side of the whistleblower, Margot O'Toole in the Thereza Imanishi-Kari case and probably increased congressional interest in the problem. They also looked at a neglected aspect of the Darsee case, the responsibility of his many coauthors. They reviewed 109 papers published by Darsee with 47 coauthors, even papers in which fabrication was not alleged. Stewart and Feder reported what they thought were many cases of errors, republication of information, use of common controls for several studies, and lapses from standards that they felt reflected poorly on the coauthors, journal editors, and reviewers. Their paper took four years to be published because of criticisms and even threats of libel by Darsee's coauthors. The article was finally published in the journal *Nature* with an editorial and with a critique of the paper by Dr. Braunwald, Dr. Darsee's mentor at Harvard (43,44). In the four years required to have their paper published, Stewart and Feder made much of the idea that they were being "censored," and had a sympathetic ear in the U.S. Congress. The pair also developed a software package that proved useful to substantiate or refute allegations of plagiarism, the so-called plagiarism machine. Since they were working without authorization while on the government payroll and their zealousness offended many scientists, they were finally moved to new positions at NIH in 1993 and told to restrict work on misconduct investigations to their own time (42).

Office of Research Integrity (ORI)

The role of ORI, which was formed in 1992, is to manage PHS research integrity activities. One function is to investigate misconduct in the NIH intramural programs, but the major activity is to oversee extramural misconduct investigations conducted by grantee institutions. ORI also develops model policies and procedures for handling allegations of misconduct which institutions can adapt for their use. Other responsibilities include evaluation of institutional policies and procedures, and investigation of whistleblower retaliation complaints. The ORI, despite much criticism from scientists and whistleblowers, remains as of 1999 the major federal watchdog for PHS grants (14).

ACADEMIC AND INSTITUTIONAL ROLES IN SCIENTIFIC MISCONDUCT

As the reports of misconduct cases increased and persisted over several years, universities and professional organizations began to study the problem and develop policies and procedures for dealing with alleged cases of misconduct (45,46). Since 1992 the NIH had required that all institutions receiving NIH training grant awards provide educational programs in research integrity for the trainees. As educational materials were developed, common themes emerged for normative rules of behavior. Guidelines for laboratory research practices were developed and shared by many academic institutions. The common features covered the areas where problems had arisen. These included laboratory procedures for recording, storing, and safeguarding research data used for the preparation of scientific papers and progress reports. Also covered were authorship practices such as whose name goes on a paper and in what order; who can be legitimate authors, the responsibility of coauthorship, the problems with honorary authorship, and, finally, mentoring responsibilities for junior research investigators when research is carried out in a training environment (45,46). Unacceptable behaviors were identified in the form of lists with definitions, with some derivations from the PHS and NSF definitions of research misconduct outlined previously. These included falsification and fabrication of data, plagiarism, dishonesty in publications, deliberate violation of regulations, failure to report misconduct, and failure to respect property of others. As factors that might influence behavior were considered, attempts were made to relieve pressures on faculty in the area of promotion and consideration for tenure. For example, some universities began to reduce the number of publications that may be considered in a promotion file for each rank considered (46). With the transfer of basic science findings into biotechnology and the opportunity for scientists to receive stock from joint ventures with industry, academic institutions developed rules about conflicts of interest and disclosure. It became apparent that if scientific integrity were to be considered an important value within an institution, the leaders of the institution would have to pay more than lip service to the concept (47). In addition the PHS and NSF requirement that institutions deal effectively with allegations of misconduct and protection of whistleblowers led to administrative changes within individual universities. In general, a specific official was identified within institutions to receive all allegations and to start the process of response in motion (46).

The Scientific Research Society Sigma Xi prepared a booklet for science students in 1984, with a third printing in 1991, entitled *Honor in Science* (48). This covers misconduct in science and whistleblowing. It starts with a section on "why honesty matters." The Committee on Science, Engineering, and Policy of the National Academy complex updated an educational booklet in 1995, *On Being a Scientist*, as a guide for teaching about scientific integrity (49).

A Panel on Scientific Responsibility and the Conduct of Research was formed under the sponsorship of the Committee on Science, Engineering, and Public Policy of the National Academy complex and published in 1993, in two volumes, a comprehensive report entitled: *Responsible Science, Ensuring the Integrity of the Research Process* (13,46). One critical review of the work noted three shortcomings: (*1*) "It insists on an unworkably narrow definition of misconduct." (The definition was fabrication, falsification, and plagiarism; see the discussion above.) (*2*) "It acknowledges studies of science as a social endeavor without taking to heart their arguments and implications." (*3*) As a consequence of the preceding weaknesses, the report's analysis of putative causes of misconduct and its proposed remedies are inadequate to the challenges now confronting science" (50). The argument has been also made by John Bailar, who works in epidemiology and biostatistics and for many years was statistical consultant to the *New England Journal of Medicine*. He feels that falsification, fabrication, and plagiarism are less a threat to the integrity of science than the day-to-day handling and reporting of data and the use of statistical methods (51). Some professional societies have developed codes of ethics or guidelines for responsible research. The American Society of Biochemistry and Molecular Biology developed a code of ethics in 1998 and seemed to learn a lesson in the process that trust can be expected only by acting responsibly (52). The Society for Neuroscience developed guidelines for scientific communications (53).

Another group that is responsible for publication and authorship practices consists of the editors of scientific publications. They have changed policies to discourage honorary and irresponsible authorship (54). A suggestion has been made by Drummond Rennie that experimental auditing of published manuscripts be undertaken after advance warning to scientists, with an eye to establishing what the real frequency of science fraud is. The information would not be used to investigate or prosecute cases. Although this process was never implemented, there have been calls for much more aggressive audits such as those conducted by the FDA (11) or suggested by the CRI (14). Since 1989 the *Journal of the American Medical Association* requires authors to sign a statement that they will produce data upon which the manuscript is based for examination by the editors or their assignees (55), and many journals now require disclosure of any conflicts of interest.

When all else has failed in bringing cases of misconduct to some satisfactory resolution, whistleblowers have occasionally gone to court. There is a tendency now to use the False Claims Act, which results in triple damages if successfully prosecuted. In one such case Dr. Condie in 1983 felt that a colleague had falsified or fabricated data in published papers and grant applications and brought this to the attention of officials at the University of Utah and University of California, San Diego. The ORI was also involved, and neither they nor the universities found misconduct. In 1989 Condie brought the action to court and asked the Justice Department to take over the case which they did. The universities were ultimately found negligent in 1994 and ordered to pay a total of $1.6 million dollars of which Condie would share 15 percent. After the court findings, the ORI then reached a settlement with the

investigator to be excluded from federal grants for 3 years and to publish retractions or corrections of the disputed scientific papers (56). The 11-year time frame certainly indicates that the wheels of justice turn slowly when it comes to dealing with misconduct cases.

The tension between the scientific community and the government that supports it remains, but learning how to deal with scientific misconduct has been recognized by scientists as a serious challenge. In the meantime U.S. government policies are still in the process of evolution (4,57,58), and a common definition for misconduct is still being developed.

No one believes that misconduct can be completely prevented, but there is hope that it can be discouraged, and that if it occurs, it can be detected early and its impact mitigated. The FDA experience suggests that an audit system will reduce misconduct and have the greatest impact on the most serious and flagrant violations (11). Quality control has been used in industry to maintain adequate standards in general, and some forms of this process could be adapted to the research setting to deal with both inadvertent error and misconduct.

BIBLIOGRAPHY

1. S. Toulmin, *Hastings Cent. Rep.* **9**(3), 27–34 (1979).
2. L.R. Graham, *Hastings Cent. Rep.* **9**(3), 35–40 (1979).
3. F. Dyson, *N.Y. Rev. Books* **44**(6), 46–49 (1997).
4. S. Glazer, *CQ Res.* **7**(1), 1–24 (1997).
5. B. Alberts and K. Shine, *Science* **266**, 1660–1661 (1994).
6. N.Y. Kiang, *Sci. Eng. Ethics* **1**(4), 347–356 (1995).
7. R.K. Merton, *The Sociology of Science*, University of Chicago Press, Chicago, IL, 1973.
8. P. Woolf, *Hastings Cent. Rep.* **11**(5), 9–14 (1991).
9. W.J. Broad, *N.Y. Times*, March 27, p. A16 (1992).
10. J.P. Swazey, M.S. Anderson, and K.S. Louis, *Am. Sci.* **81**, 542–553 (1993).
11. M.F. Shapiro, in S. Lock and F. Wells, eds., *Fraud and Misconduct in Medical Research, BMJ* Publishing Group, London, 1993, pp. 128–141.
12. S.M. Kuzma, *Univ. Mich. J. Law Reform* **25**(2), 357–421 (1992).
13. Committee on Science, Engineering and Public Policy (US), *Panel on Scientific Responsibility and the Conduct of Research, Responsible Science*, vol. 1, National Academy Press, Washington, DC, 1992.
14. Report of the Commission on Research Integrity, *Integrity and Misconduct in Research*, U.S. Government Printing Office, Washington, DC, 1996.
15. R.J. Levine, *Ethics and Regulation of Clinical Research*, 2nd ed., Urban & Schwarzenberg, Baltimore, MD, 1986.
16. A.R. Price, *J. Higher Educ.* **65**(3), 286–297 (1994).
17. R. Dresser, *J. Am. Med. Assoc.* **269**(7), 895–897 (1993).
18. J.T. Edsall, *Sci. Eng. Ethics* **1**(4), 329–340 (1995).
19. M.P. Glazer and P.M. Glazer, *The Whistleblowers*, Basic Books, New York, 1990.
20. F. Hoke, *Scientist* **9**(10), 1, 15 (1995).
21. J.S. Lubalin and J.L. Matheson, *Sci. Eng. Ethics* **5**, 229–250 (1999).
22. T. Devine, *Scientist* **9**(10), 11–12 (1995).
23. D.J. Kevles, *The New Yorker*, May 27, pp. 94–109 (1996).
24. D.J. Kevles, *The Baltimore Case*, Norton, New York, 1998.
25. P. Peck, University of Pittsburgh and NCI Settle with Dr. Fisher, *OB.GYN. News*, October 1, P11 (1997).
26. E.J. Hackett, *J. Higher Educ.* **65**(3), 242–260 (1994).
27. W.M. Carley, *Wall Street J.*, July 26, pp. A1, A6 (1994).
28. B. Barber, *Science* **134**, 596–602 (1961).
29. T.S. Kuhn, *The Structure of Scientific Revolutions*, 2nd ed., University of Chicago Press, Chicago, IL, 1970.
30. W. Broad and N. Wade, *Betrayers of the Truth*, Simon & Schuster, New York, 1982.
31. A. Kohn, *False Prophets*, Blackwell, Oxford, UK, 1989.
32. D.J. Miller and M. Hersen, eds., *Research Fraud in the Behavioral and Biomedical Sciences*, Wiley, New York, 1992.
33. M.C. LaFolette, *Stealing into Print*, University of California Press, Berkeley, 1992.
34. S. Lock, in S. Lock and F. Wells, eds., *Fraud and Misconduct in Medical Research*, BMJ Publishing, London, 1993, pp. 5–41.
35. L.K. Altman, *N.Y. Times*, October 30, p. A12 (1996).
36. E. Marshall, *Science* **270**, 1912–1914 (1995).
37. E. Marshall, *Science* **273**, 1162–1164 (1996).
38. D.K. Price, *Daedalus* **107**(2), 75–92 (1978).
39. D.H. Guston, *Centen. Rev.* **38**(2), 215–248 (1994).
40. M.C. LaFollette, *J. Higher Educ.* **65**(3), 261–285 (1994).
41. P. Gwynne, *Scientist* **2**(13), 1, 8 (1988).
42. F. Hoke, *Scientist* **9**(3), 3–15 (1995).
43. W.W. Stewart and N. Feder, *Nature (London)* **325**, 207–214 (1987).
44. E. Braunwald, *Nature (London)* **325**, 215–216 (1987).
45. Institute of Medicine Report, *The Responsible Conduct of Research in the Health Science*, National Academic Press, Washington, DC, 1989.
46. Committee on Science, Engineering and Public Policy, *Panel on Scientific Responsibility and the Conduct of Research, Responsible Science*, vol. 2, National Academic Press, Washington, DC, 1992.
47. C.K. Gunsalus, *Acad. Med.* **68**(9), Suppl. S-33–S-38 (1993).
48. *Honor in Science*, Sigma Xi Press, Research Triangle Park, NC, 1991.
49. Committee on Science, Engineering and Public Policy, *On Being a Scientist, Responsible Conduct in Research*, National Academy Press, Washington, DC, 1995.
50. E.J. Hackett and S.M. Solomon, *BioScience* **43**(10), 717–719 (1993).
51. J.C. Bailar, III, *Chron. Higher Educ.*, April 21, pp. B1–B3 (1995).
52. F. Grinnell, *Sci. Eng. Ethics* **5**, 205–217 (1999).
53. M.J. Zigmond, *Sci. Eng. Ethics* **5**, 219–228 (1999).
54. D. Rennie, *J. Am. Med. Assoc.* **271**(6), 469–471 (1994).
55. D. Rennie, *J. Am. Med. Assoc.* **270**(4), 495–496 (1993).
56. P. Hilts, *N.Y. Times*, July 23, p. 9 (1994).
57. K. Ryan, *Prof. Ethics Rep.* **9**(2), 1, 6–7 (1996).
58. K. Ryan, *Chron. Higher Educ.*, July 19, pp. B1–B2 (1996).

See other entries SCIENTIFIC RESEARCH, ETHICS, VALUES IN SCIENCE; SCIENTIFIC RESEARCH, LAW, AND PENALTIES FOR SCIENTIFIC MISCONDUCT.

SCIENTIFIC RESEARCH, ETHICS, VALUES IN SCIENCE

RACHELLE D. HOLLANDER
National Science Foundation
Arlington, Virginia

OUTLINE

INTRODUCTION

This article reviews recent scholarship to demonstrate ways in which science is value laden and how these value dimensions have ethical importance. It identifies the values scientists generally accept as those that should govern their work, as well as other values that influence support for science, its conduct and outcomes. It discusses some implications of these values for ethical issues in biotechnology. The approach is to review relevant philosophical and social science literature that examines relationships among science, ethics, and values, and to identify and analyze some examples of ethical issues in biotechnology.

The article contains two basic themes. One focuses on values and science. It reviews literature in the philosophy and sociology of science to describe a controversy that occupies the attention of scholars in those fields, as well as that of some scientists. The philosophers and scientists want to answer a question like: Does science tell us something important and true about reality? The sociologists want to answer a question like: Does the scientific community have a distinctive set of norms or values?

This first theme considers such topics as: Is science value laden or value free? What is the relationship of intrinsic and extrinsic values to science? What are constitutive and contextual values in or in relationship to science? From philosophical and sociological perspectives, these are the issues that comprise an attempt to answer the questions: What is science? Is science a set of independent and abstract criteria? Or is it what scientists do, and where they do it, and what they do it with? Or all of the above (1)?

The second theme focuses on the ethical implications of values in science. It takes the position that the embedding of science in society means the conduct of science, and its outcomes always have ethical implications. Furthermore societal commitments to innovation mean that these ethical implications are momentous and complex. Thus people can, and should, make moral judgments about science in its undertaking and with respect to its influences and outcomes. These are judgments about the moral responsibilities of individuals and the social responsibilities of institutions, about the moral and social implications of scientific practices, and about the effects from these endeavors for humans and their communities and environments. Making well-considered judgments can be assisted by moral theories and conceptual analysis, by systematic research to understand phenomena, and by public deliberation that takes care to include a wide variety of views. The proportionality principle in ethics assigns greater responsibility to persons more likely to influence outcomes. It would follow that people in positions of influence in biotechnology have greater responsibility to promote activities that encourage these kinds of considerations and deliberations.

The contemporary world and the political states that comprise it have made a substantial commitment to the production of science and its integration into society. Research institutes and educational, industrial, and governmental organizations house scientific laboratories and research facilities. Thereby the conduct of science has societal impacts in and of itself. To do science requires social institutions and social commitments. The production of this encyclopedia and the attention of its readers testifies also to the societal influences of science. Many entries demonstrate and discuss the changes in individual lives, social institutions, and the environment that result from doing science. It is clear, then, that science in its conduct and impacts cannot be value free. But this is only the beginning of the story.

SCIENCE AND VALUES

Numerous terms are used in the debate over whether or not, and how, science incorporates values. Perhaps one of the earliest and simplest contrasts is between value-free and value-laden science. This contrast goes back to the philosophers of the Vienna Circle in 1924 to 1936 (2). From their perspective, meaning itself was to be limited to value-free statements that lent themselves to sensory proof, or some kind of derivation from sensory proof. Later philosophers could not reconcile this position with quarks and genes, or the logical difficulty of moving from data to evidence to theory, and the philosophers of the Vienna Circle gave up that project. However, the questions of what are the distinctive features of science and whether they lead to results with a special claim to value-free truth remain important to scientists and those studying science. They are of social importance too, since scientific claims and claims about science often underlie large social expenditures and influence political and social outcomes.

In examining this issue further, it is useful to consider the contrast between objectivity and subjectivity, and between constitutive and contextual values. Scientists are content to view science as having constitutive values. These values are also called epistemic, as well as internal or intrinsic to science. They are contrasted with external or extrinsic values. Constitutive values would include the value placed on observation and experiment, on prediction and explanation, on the building and testing of theory and models, and on the development and testing of methods

that can reduce the probability of error or self-deception in scientific work. Science remains self-contained in satisfying these values, or most scientists view it to be so. Scientists and others have thought their satisfaction provides objective methods leading to objective truth. For good reviews of recent debates in philosophy of science about the status of scientific claims to privileged standing in understanding the natural world, see Couvalis (3) or Klee (4). Both argue for the preservation of scientific claims to objective understanding of the natural world.

The strength of this perspective on scientific values can be seen in the wide acceptance of the work of sociologist Robert Merton. In 1942 Merton characterized science as satisfying four norms: communism (open sharing of findings), universalism (use of general criteria for judging knowledge claims), disinterestedness (not acting for personal advantage), and organized skepticism (withholding belief). Only after several decades, in the 1970s, was this view challenged. Ian Mitroff proposed that the opposite views operated as counternorms in science, while Aaron Cicourel and Michael Mulkay challenge the sociological foundations of norms as categories, by arguing that norms are comprised in particular acts of individual actors (1, pp. 398–400).

Philosopher Joseph Rouse provides a current critique of both realist views that science tells how things really are and relativist or constructivist views that all scientific beliefs can be explained sociologically. He believes that both of these accounts suffer from their presupposition that a global assessment of science as one kind of knowledge is possible. Instead of a global view, the study of the actual variety of scientific practices might provide better grounds for an assessment of their significance. Understanding science as historically situated responses to past scientific effort in anticipation of and contest over its future development allows people to evaluate particular scientific claims without the presumption of a particular view about the coherence of scientific knowledge or about the relationship of that knowledge to reality (5).

From time to time some scientists and policy makers indicate that science is value free, or objective. Often in doing so, they want to use such statements to provide support for a favored position or to undermine one they do not support. If constitutive, or epistemic, values and the Mertonian norms were the only values in science, perhaps it would be possible for them to do this. Scientists would (sooner or later) find the truth. The hard work involved in making social decisions using that truth would be "handed off" from scientists to policy makers or other social decision makers. Scientists discover what they can, and then societal values and priorities need to take over. This model implies a clean division of labors and categories, value-free science, and value-laden policy.

Before identifying some flaws in this position about the separability of science from values, this entry needs to point out one immediate limitation. No one believes that technologies are value free. Technologies, from abacuses to zylophones, are devised and selected specifically to serve human purposes or values. If science is used in the development of technologies, that development is not value free. Furthermore technologies often make science possible, so at least in this sense, science is not value free. This encyclopedia is about biotechnology, indicating a kind of science that is not separable in entirety from technology. Thus, even if there is a kind of science that can fulfill the conditions for a value-free science, it does not seem to be the kind under consideration here. Discussing the ways in which science, thought to be value-free, can and does incorporate values is important both for an adequate understanding of what science is, and of its ethical implications. If a strong case can be made that science does incorporate nonconstitutive values, the case for their presence in the science (or sciences) of biotechnology is strengthened.

This entry uses notions of constitutive and contextual values found in the work of Helen Longino (6). Contextual values are values independent of the scientific goals or constitutive values identified above. Contextual values are social, cultural, economic, political, moral, and personal. These values can influence science, while it remains what all might agree is good science according to those constitutive values. Consider, for instance, a social or political commitment to fund materials science and engineering to understand properties of corrosion and fracture; consider the commitment to protection of human or animal subjects. These social values influence whether and what science gets done, but they do not necessarily result in bad science or science whose epistemic value is affected negatively. There are many examples that show that contextual values affect constitutive values, even when such science remains good science according to the constitutive values identified above. Reports of data falsification leading to retractions of published papers and findings provide obvious examples where contextual values produce bad science. A notorious example in the 1970s involved coloring mice to indicate successful skin grafts; the miscreant blamed personal stress and exhaustion (7).

Longino identifies five ways in which contextual values affect science. They affect scientific practices, the questions that get attention, and the description of data. Contextual values also affect the background assumptions and the global assumptions with which science operates (6, p. 86).

A good example of the deep interpenetration of science and contextual values comes in Nelly Oudshoorn's socio-historical analysis of the development of understanding about female sex hormones, which led to the invention of birth control pills (8). Her analysis shows how society influences questions that get attention, scientific practices, and the data that get collected.

This story about the making of sex hormones arose with new developments in the chemical life sciences at the turn of the twentieth century. The actors in the drama were physiologists working in laboratories, gynecologists treating women, and the pharmaceutical industry. The materials these actors needed influenced what would count as knowledge. The episode produced and transformed gender bias in science.

Before the 1920s the three groups began to interact over hormonal products. In the 1920s they ran up against problems in getting the amounts of gonadal material required to do the work they wanted to do. The

material was expensive and difficult to find, especially for the laboratory scientists. Because of this need, cooperation began to intensify between the scientists and pharmaceutical companies that supplied hormones from animal gonads. Gynecologists relied on these groups for quality hormonal products — whose benefits for their patients were, by the way, quite unclear.

In 1926 two German scientists happened on the long-sought source — human urine. Urine from pregnant women was particularly good. To gain access, pharmaceutical companies and laboratory scientists relied on the gynecologists' new, inexpensive source. But this was a prolific source of female sex hormone, not male. Urine from men would certainly be a suitable source for male sex hormone, but there was no institutional context for its collection. There were no clinics specializing in men's reproductive systems in the 1920s.

Though laboratory scientists had been interested in the role of both male and female sex hormones in growth and development of the body and in sexual differentiation, the ease of access to certain kinds of research materials lent support to, or privileged, the development of certain knowledge claims rather than others. This material source, along with the institutional context that focused on women's reproductive disorders, saw men disappear as a focus for research. Not until the late 1960s was the study of male reproductive disorders institutionalized as a medical specialty (8, p. 80).

This example demonstrates how contextual factors — ranging from the values different groups placed on social desiderata, such as potency or birth control, to the possibilities created by access to particular material such as a kind of urine — influence the very *constituents* of science. What counts as evidence, and for what — what can be used for the purposes of constitutive values — is intimately wrapped up in the social circumstances and priorities of times and places. This in turn influences what diseases or disorders get so labeled and treated, and when they do. This result has unavoidable ethical implications.

Background and global assumptions in science have implications for the interpretation of data in questionable ways that carry both policy and ethical implications. Longino provides the interesting example of research on human origins. This research uses one of two organizing principles: man-the-hunter and woman-the-gatherer. Each uses a story of the gendered development of tool use to promote a view of the favored sex as the initiator of activities from which defining human traits of intelligence and cooperation evolve. Each story promotes the view that men and women make tools for different purposes. However, while the remnants used to construct these stories provide evidence sufficient to conclude humans shaped them, they do not provide evidence sufficient to conclude if men or women used them, or how (6, pp. 106–108).

These examples demonstrate that contextual values affect science at very fundamental levels, but that these effects might not overthrow claims of scientific objectivity — in the sense of trying to work toward scientific understanding of natural phenomena, no matter how partial or influenced by nonepistemic motives and understandings. Recognizing this kind of partiality places a severe limitation on the adequacy of science or scientific answers at any particular time, for non-scientific purposes. The social responsibility for recognizing this limitation, and figuring out how to respond, falls both to scientists and non-scientists, particularly those in influential policy positions.

Another kind of partiality is less esoteric. Scientists and their academic sponsors, with commercial interests in development of products that follow from their discoveries, can be found using press conferences to promote their latest results. The tragicomic episode announcing the discovery of cold fusion provides a recent example demonstrating the limits of Mertonian norms in influencing scientific behavior (7, pp. 11–12). If the public is somewhat skeptical in its reception of such activities, the skepticism can be regarded as healthy prudence. Scientists who wish to honor the constitutive ideals of replication or peer review would also be likely to withhold approval. Here too, the social responsibility for recognizing this limitation, and figuring out how to respond, falls both to scientists and nonscientists. The proportionality principle would require more from scientists and others in positions to be influential than would be expected or required from those not so placed.

SCIENCE, ETHICS, AND THE STATE

The previous examples demonstrate that and how scientific endeavors require societal support. Science occurs in organizations and institutions and plays an important role in promoting particular social goals and interests. Public and private sector organizations sponsor scientific research with the expectation that their support will lead to public and private gain. These results bring with them questions about the nature and extent of gain; these questions have both utilitarian and distributional components.

Jurgen Schmandt and James Everett Katz identify three ways in which contemporary societies value science — as a product, as evidence, and as method. As a product, science is promoted and controlled in the interests of innovation. As evidence, science is used and interpreted for policy purposes. And scientific methods — analysis, experiment, empirical techniques — are valued in themselves (9). They are valued at least in part because they can serve for purposes of innovation and evidence.

The use of science in these ways means that, inevitably, people in democratic societies that encourage public involvement will call it to account. They will ask whether the use is justified, and consider the ethical implications of the choices that are made. As human interventions become pervasive (ozone holes), mammoth (Three Gorges Dam), and more sophisticated (recombinant DNA technology), so do the scale and requirements of human accountability, including scientific accountability. Two kinds of ethical issues for biotechnology are worth examining in this regard. One concerns ethics and risks; the other, the relationship of science to political consent (10).

Ethics and Risks

As a force for innovation, science brings inevitable risks. When human beings engaged in scientific activities or the use of science-based technologies create risks, those affected will ask whether the activities have been undertaken responsibly. Below are some examples.

Differential Impacts. Equity issues that arise in the management of radioactive or hazardous wastes provide a useful example of the kinds of problems resulting from commitment to science for innovation. Placement of these wastes raises questions of differential impact for locus, labor, and legacy (11). Locus raises the ethical isues of locating facilities that may harm those nearby while benefiting those far away. Legacy raises a similar question for those removed in time. These are issues of distributive justice; they ask about the fairness of the distribution of benefits and costs. Similarly for labor, although here the questions are complicated by issues of consent. Good justification of any state decision requires attention to these ethical issues.

Science and Differential Impacts. Looking more closely at the issue of evidence in this waste disposal example illustrates a particularly important quality in the value dimensions in science and how they raise ethical questions. To determine whether or not people are harmed, a scientific test is used. The decision to use a test has moral dimensions. It is a choice made by human beings, and it may help or harm them or their surroundings. Once the choice is made, decisions about which test to use also involve moral dimensions. Whichever tests are used, they will result in false negatives or false positives, at least to some degree. The choice in either direction will be a moral choice; thus a science-based choice involving constitutive values contains contextual value implications. If the test selected will result in false positives — that is, one that will tell us that some persons are affected who are not — these persons may be subjected to a variety of harms ranging from risky treatments to stigmatization or job loss. If the test selected will result in false negatives — that is, one indicating that affected persons are not affected — persons may suffer unattended illness and premature death. Independent of purposeful wrongdoing, application of science to questions of individual and societal well-being will involve value choices of ethical importance (12).

The previous examples show how science contains values dimensions with ethical implications. They show that these dimensions are not avoidable. Governments as well as other social institutions need to take these implications into account in order to behave responsibly. The sociopolitical context in which biotechnology is developing requires persons and institutions promoting it to pay attention to the contextual and constitutive values that affect biotechnology, as well as the ethical implications that are part of and follow from its promotion.

Value Conflicts. The case of deliberate release of genetically engineered organisms gives examples of direct relevance to this encyclopedia. Using Longino's categories, Soemini Kasanmoentalib provides a list of examples in which contextual values influence the scientific development and assessment of biotechnologies (13). Under practices, he points out that the multinational operation of companies doing biotechnology R&D allows them to select countries with less rigorous regulations for testing. In this case the commercial values of the company may coincide with the values some scientists place on being free to do their research, but they may conflict with those of the public and other scientists who are more concerned with the potential for harm. Here the conflict arises between the positive constitutive value scientists place on doing research and the negative contextual value placed on potential harm to vulnerable human populations. In another case in which constitutive and contextual values conflict, scientists at a university working with commercial support may find the value they place on open discussions and publication challenged by the need to patent or keep secret their findings. Here the value conflict can involve the constitutive value of open discussion and what a firm might argue is the utilitarian value that its product will provide. The firm might also point to the contextual value of promise keeping, if the scientists have signed an agreement. All of these kinds of cases, where values conflict, need careful ethical consideration.

Science and Short-Term/Long-Term Interests. Commercial interests are not the only interests in economic growth. Governmental desires to foster innovation may result in limiting the kinds of questions that are asked and data that are required before approvals to plant or market products of biotechnology are gained. The Kasanmoentalib article notes that criteria of what should count as risk or damage, or appropriate ecological end points by which environmental stress can be measured, are difficult to establish, and can be limited to the gross and near-term. Under specific assumptions the selection of a model on which to base the risk assessment can be less conservative, ignoring synergistic effects for which it is difficult to devise tests. Thus global assumptions favoring the reductive approaches to assessing risk found in molecular biology and genetics can be favored over those from ecology which incorporate more concern for synergistic, inclusive, and long-term effects.

Here, keep in mind that the decision to delay introducing a new genetically engineered organism may pose ethical risks also. For instance, a plant engineered to resist a pest may require less pesticide and provide more food to an impoverished area. The decision to hold off on its introduction may trade off short-term need for more crop against long-term concern for ecosystem health. How to identify and balance short-term and long-term interests is a difficult ethical question. The stories in June 1999 in the science journal *Nature* and many of the world's newspapers about the lethal effect of bioengineered corn pollen on Monarch butterflies show that these concerns are of more than theoretical importance.

Particular choices in these circumstances are no more scientific than their opposites. Given that this is so, the answers to the questions: What is risk? What is acceptable risk? What is acceptable evidence of risk?, are themselves

not just scientific. On the one hand, parties outside the sciences must help to establish standards for acceptable risk. On the other hand, careful scientific or what can better be called meta-scientific considerations of epistemic issues for science, of what scientists are entitled to say they know, are essential (12). Careful analysis of the justifications for the assumptions and findings of various sciences is essential, both to identify and show the limits of science, and to help make scientific progress possible, by challenging the conventional wisdom in particular fields. It is better not to call these choices scientific, since that can connote value-free or objective choices, in a way that they cannot be.

Science and Political Consent

Science and Political Values. Recent controversies over labeling of food products that contain elements that are biotechnologically engineered show another dimension of the intricacies of interactions between science, technology, values, and ethics. The U.S. government and some commercial biotechnology interests are arguing that there are no grounds for labeling bioengineered products as such, since they pose no additional safety risks because of the bioengineering. They are using a narrow definition of a safety risk here. Such risks would be those, for instance, created by incorporating an allergen or toxin into a plant. Engineering an herbicide or herbicide resistance into a plant would not constitute such a risk. Long-term risks to ecosystems are not being considered (13).

Another category of ethical concern needs to be mentioned here. As Paul Thompson points out, people want to know a great many things about food products besides and beyond the kinds of safety concerns identified above (10, p. 75). They want to know where the foods come from. They want to know the processes by which they have been made. It is one thing to say that government should ensure that health-related claims on a product are not false and misleading, and that science should be used to assist in making that determination. It is quite another thing to say that no information except for scientifically validated claims about safety should be allowed on food products. Saying that is to substitute science for the consent of the governed. For persons who wish to buy hot sauce from Louisiana, or free range chickens, or organic produce that is not bioengineered and is raised on farms that use no manufactured pesticides or herbicides, such rulings would abolish choices they currently have and value having. It would substitute scientific rule-making for democratic choice, indirectly silencing a political voice.

Science and Moral Theory. Moral theories, unlike scientific theories, provide judgments about what human beings should do, not what the world is. However, the previous discussion serves to illustrate how science involves values and moral implications. It illustrates how human judgments about the worth of science exist in a context in which questions central to moral discourse — about what promotes human well-being and what is fair — arise. It illustrates how human judgments called scientific may incorporate contextual as well as constitutive values that presuppose particular answers to the questions: What kinds of scientific research should be done? Who should own the results? How should the results be described in the open market? Who is benefited or harmed? How? And how much?

The previous discussion can also illustrate limits of moral theories in resolving issues. Some scientists and regulators in government agencies and industry spokespeople might insist on utilitarian grounds that no products be labeled not to contain bioengineered components. They might believe that the greatest good for the greatest number will be served by having people buy less expensive safe products than they might otherwise choose, because of unjustified fears, if they see an equivalent, more expensive product with the "no biotech" label. Insisting on the right to have such a label gives pride of place to, or privileges, the rights view. That view insists that persons have rights to make choices that others think they are making for less than satisfactory reasons. Is one of these views the morally right one? On what grounds would that decision be made?

One way to try to answer this question is to try to resolve the facts of the case, where the facts are the empirical claims that are being made. Is the utilitarian claim true? Is it true in all circumstances? There may be circumstances in which ungrounded fears will influence consumer purchases, but these may be relatively few. It is easy to imagine ways to overcome such fears. Further, consumers may begin to feel manipulated and distrustful of a system in which they believe information is being kept from them because of commercial interests. With this scenario, utilitarian theory itself may be better served in a marketplace which allows labeling that includes information in addition to scientific claims about safety. A world in which more rights and freedoms can be honored may be a better world, by utilitarian standards, even when it allows choices that are ostensibly less well grounded by some current scientific standards. While science can be enlisted to serve a particular moral point of view and, in this case, given a utilitarian cast, the claims underlying such an outcome needn't be accepted. If this response is accepted, a moral conflict can be settled by finding a creative middle way in which moral theories can be reconciled (14).

This creative middle way may allow the preservation, or perhaps even the transcendence of utilitarian and rights-based moral theories. Another moral point of view is worth mentioning, one that William Aiken called holistic (15). This view, also referred to as the interconnectedness approach, points out that utilitarian approaches that proceed by examining trade-offs, or costs and benefits, risks and benefits, can lose sight of the connections that mean that the natural world does not operate like a balance sheet. These connections mean that negatives cannot simply be traded off against positives; negatives may be necessary to the maintenance of a desirable whole. Life requires evolution, predation, and death. From this perspective, neither rights views nor trade-off views give due recognition to larger values that need consideration in the relationships between human beings and the larger environment or natural world. This approach demands attention to the values we wish to maintain in our social

practices as well as the kinds of long-term consequences that may be overlooked in standard moral theories. The satisfaction of these values for such things as communities and ways of life requires consideration of long-term consequences and the connections that structure and sustain wholes. The interconnections approach may not reduce to either utilitarian or rights-based approaches. However, it may provide a connection between ethical and ontological theories, between theories about what ought to be the case and what is the case. It may require and provide a context which is hospitable to raising new ethical issues and challenges.

Currently scientific approaches to risk in the formulation of environmental policy take what Thompson labels a purist, rather than a hybridization, view. The purist view separates risk into its components: risk to human health, animals, environment. While social consequences could be a component also, the current policy process in the United States does not recognize social consequences as a legitimate topic for discussion. For the public, however, risk is an amalgam or hybridization of at least all of these components (10, pp. 232–236). Additionally, the public view of risk contains a concern for the responsiveness of those with authority and power — be it scientific, political, financial, corporate, organizational — to this amalgam. Unless science and scientists recognize that the approaches they take to risk incorporate these constraints, which have moral dimensions, they are operating under false assumptions with respect to both the ethical and value implications of their work. If scientists and others with the kinds of power identified above can recognize and be responsive to these implications, it may be possible to reconcile scientific and social progress.

SUMMARY

The task of this entry is to show some of the ways in which science is value laden and how its doing and results have ethical implications. The first section shows that both intrinsic or epistemic and extrinsic or contextual values inhere in science. The discussion and examples show how contextual values affect the scientific search for explanations of natural phenomena and the outcomes of that search, and how the search and outcomes incorporate social priorities and biases. The discussion and examples point out that these outcomes have ethical implications. Thus, the material and institutional circumstances surrounding hormonal research in the 1920s led to an emphasis on the study of female rather than male reproduction. Scientific discoveries in this area have been of enormous social benefit. However, the partial understanding that science provided placed limits on its appropriate use as an underpinning for societal decisions. Less esoteric, current examples of scientific partiality — for instance, that arising from scientific promotion of research results for commercial purposes — also give rise to healthy caution. These limits on scientific understanding require careful attention from scientists and others in positions to influence social policy and programs.

The second section continues the discussion of how values enter science and discusses ethical implications in the pursuit of science as a social or national priority. This pursuit affects individuals, organizations, communities, and the environment, and brings with it inevitable ethical questions about the nature and distribution of benefits and harms. The section examines issues of ethics and risks from science-based innovations, and issues of science-based innovations and conflicts over consent. Science-based technologies include and create differential risks. Figuring out what these risks are, and what their ethical implications are, is a complex task. The findings deserve careful attention from those in positions to influence social policy and programs. The same is true with issues raised by the relationships between science and regulation, as the example of food labeling makes clear. The call for science-based labeling gives priority to particular moral and political views as well as a particular view of scientific truth. Once again, these views require careful identification and consideration from those in positions of authority, if they wish to be accountable and responsible for their actions.

BIBLIOGRAPHY

1. T. Gieryn, in S. Jasanoff, G. Markle, J. Peterson, and G. Markle, eds., *Handbook of Science and Technology Studies*, Sage Publ., Thousand Oaks, CA, 1995, pp. 393–443.
2. D.D. Runes, *Dictionary of Philosophy*, Philosophical Library, New York, 1960.
3. G. Couvalis, *The Philosophy of Science, Science and Objectivity*, Sage Publ., London, 1997.
4. R. Klee, *Introduction to the Philosophy of Science, Cutting Nature at its Seams*, Oxford University Press, New York, 1997.
5. J. Rouse, *Engaging Science: How to Understand Its Practices Philosophically*, Cornell University Press, Ithaca, NY, 1996, see "Introduction," and Chapter 6.
6. H.E. Longino, *Science as Social Knowledge*, Princeton University Press, Princeton, NJ, 1990, p. 4.
7. D.B. Resnik, *The Ethics of Science: An Introduction*, Routledge, London and New York, 1998, pp. 75–76.
8. N. Oudshoorn, *Beyond the Natural Body: An Archeology of Sex Hormones*, Routledge, London, 1994, Chapter 4, pp. 65–81.
9. J. Schmandt and J.E. Katz, *Sci. Technol. Hum. Values* **11**(1), 40–52 (1986).
10. P.B. Thompson, *Food Biotechnology in Ethical Perspective*, Blackie Academic & Professional, Chapman & Hall, London, 1997, pp. 57–79.
11. R.E. Kasperson, ed., *Equity Issues in Radioactive Waste Management*, Oelgeschlager, Gunn and Hain, Cambridge, MA, 1983.
12. R.D. Hollander, in D.G. Mayo and R.D. Hollander, eds., *Acceptable Evidence, Science and Values in Risk Management*, Oxford University Press, New York, 1991; see also C.F. Cranor, *Regulating Toxic Substances: A Philosophy of Science and the Law*, Oxford University Press, New York, 1993, pp. 40–48.
13. S. Kasanmoentalib, *J. Agric. Environ. Ethics* **9**(1), 42–60 (1996).

14. C.E. Harris, M. Pritchard, and M. Rabin, *Engineering Ethics: Concepts and Cases*, Wadsworth Press, CA, 1995, pp. 135–138.
15. W. Aiken, in K.A. Dahlberg, ed., *New Directions for Agriculture and Agricultural Research: Neglected Dimensions and Emerging Alternatives*, Rowman & Allanheld, Totowa, NJ, 1986, pp. 31–41.

See other entries SCIENTIFIC RESEARCH, ETHICS, SCIENTIFIC MISCONDUCT; SCIENTIFIC RESEARCH, LAW, AND PENALTIES FOR SCIENTIFIC MISCONDUCT.

SCIENTIFIC RESEARCH, LAW, AND PENALTIES FOR SCIENTIFIC MISCONDUCT

BARBARA MISHKIN
Hogan Hartson L.L.P.
Washington, D.C.

OUTLINE

INTRODUCTION

Individual scientists and their research institutions (academic, nonprofit, or industrial) risk a variety of penalties if they engage in scientific misconduct. Penalties may be imposed administratively by the federal agency (if any) that supported the research or by the Food and Drug Administration (FDA) which must approve each new drug, biologic, or device before it may be marketed. Additional penalties may follow criminal prosecution or civil claims pursued through the courts.

Potential penalties range from increased supervision of research to sizable fines and imprisonment. The scientists and their institution may also be excluded from further participation in government funded or federally regulated research for a period of years (or sometimes, permanently). In addition, if the discredited research was supported by a federal grant or contract, the funding agency may demand that the full amount be returned. The agency invariably insists that articles or reports found to be the result of misconduct be formally withdrawn, and the publicity surrounding findings of scientific misconduct can tarnish the reputation of the research institution and destroy the career of the scientists involved. When the research institution is a commercial entity, the misconduct findings and penalties also may affect sales of its products and the value of its stock or its ability to make an initial public offering.

ADMINISTRATIVE SANCTIONS

The two government agencies that provide most of the funding for biomedical research in the United States are the National Institutes of Health (NIH) and the National Science Foundation (NSF). Each has regulations requiring investigation of alleged research misconduct and appropriate action if the allegations are confirmed (1). (Table 2) The FDA has similar authority to impose sanctions for violations of rules governing the development and evaluation of new drugs and medical devices. The actions taken by the agencies supplement any disciplinary action imposed by the research institution.

The NIH and NSF developed their scientific misconduct procedures following a series of congressional hearings in the 1980s which criticized the responses of government agencies and the recipients of federal grants to allegations of fraud in science. The hearings and related press accounts publicized several incidents in which data were fabricated or falsified, and others in which papers submitted to a journal were found to have been plagiarized.

The federal definition of research misconduct has been controversial. Whatever the definition, confirmation that one or more scientists engaged in such misconduct will lead NIH or NSF to impose administrative actions and civil or criminal penalties appropriate to the nature and seriousness of the misconduct that occurred.

The Public Health Service and National Science Foundation

The Public Health Service, of which NIH is a part, and the NSF may take one or more of a range of administrative actions at the conclusion of a scientific misconduct investigation, unless an appeal is filed. In addition the research institution or the funding agency, or both, may impose "interim administrative actions" even before an investigation has been concluded, if necessary to protect human or animal subjects, prevent improper use of federal funds, or safeguard the public interest (2). Although described as administrative actions rather than penalties, the distinction may make little difference to the scientist or institution subjected to the action.

Interim Administrative Actions. The stated government purpose of interim administrative action is to ensure the

proper use of public funds, the protection of research subjects, and the fitness of the principal investigator to continue to direct the research project. Although research institutions are expected to take interim administrative actions as appropriate, the funding agency may take one or more actions in addition, to protect the federal interests. Administrative actions once taken are reviewed periodically and may be modified as necessary in light of new information. The possible interim actions that may be taken by a federal agency are as follows:

- Total or partial suspension of ongoing research
- Total or partial suspension of the accused researchers from eligibility to receive additional federal grants or contracts
- Prohibition or restriction of certain research activities (e.g., research involving human subjects or animals)
- Requirements for supervision and prior approvals to ensure/ compliance with federal law and to protect public health and safety
- Delaying the award of pending grants or contracts
- Revoking agency approval of key research personnel to direct or perform research activities

Administrative Sanctions for Scientific Misconduct (Table 1). If allegations of scientific misconduct are confirmed by an investigation, the funding agency may impose one or more of the following sanctions:

- Send letter of reprimand to the scientist's institution
- Require increased supervision of the scientist's research and publications
- Require that a supervisor certify the accuracy and integrity of information submitted to the agency in grant applications and progress reports
- Restrict the use of agency funds to certain activities
- Conduct special reviews of all grant applications from the guilty scientists or their institutions
- Prohibit the scientists from serving on agency advisory committees
- Suspend or terminate ongoing research support
- Debar the scientists or their institution from eligibility for federal research support for a given period of time

Table 1. Penalties for Scientific Misconduct: Federal Regulations

FDA regulations	Citation
Administrative actions for noncompliance	21 CFR, pt. 56, Subpart E
Civil money penalties	21 CFR, pt. 17
Disqualification of clinical investigators	21 CFR, §312.70
Disqualification of testing facilities	21 CFR, §§58.200-58.219
HHS regulations	
Government-wide debarment and suspension (nonprocurement)	45 CFR, pt. 76
Responsibility of PHS awardee and applicant institutions for dealing with and reporting possible misconduct in science	42 CFR, pt. 50, Subpart A
NSF regulations	
Government-wide debarment and suspension (nonprocurement)	45 CFR, pt. 620
Misconduct in science and engineering	45 CFR, pt. 689

The U.S. Food and Drug Administration

Under the Food, Drug, and Cosmetic Act, new drugs, biologics, and devices must be approved by FDA before they may be marketed. Approval is based on data collected in clinical trials demonstrating that the product is safe and effective for its intended use. Sponsors of new products must first apply to FDA for permission to conduct the clinical trials. Sponsors (usually drug or device manufacturers) or the clinical investigators who repeatedly or deliberately fail to follow FDA rules for the conduct of the research, or who fabricate or falsify their data, may be disqualified from further participation in clinical or laboratory research involving investigational products (3). Such disqualification by FDA is similar to debarment, although FDA also may debar research entities, drug and device manufacturers, and individual scientists for serious research misconduct such as submitting false statements to the agency, conviction of a crime related to the development or approval of a drug or device, or involvement in a conspiracy to commit any such crime (4). Debarment is also authorized following conviction of a crime such as bribery, fraud, perjury, falsification or destruction of records, and similar acts related to product development. The period of debarment may be as short as one year or permanent, depending on such factors as the nature and seriousness of the offense, the extent to which management was involved (either in encouraging or participating in the criminal activity or in failing to report it), and the extent to which management tried to correct the causes of or mitigate the offense.

FDA also has authority to take administrative actions similar to those described above for PHS-funded research (e.g., immediate suspension of research in order to protect research subjects or public health and safety). In addition FDA may refuse to accept data from a clinical trial to support an application to market the product being evaluated and may even withdraw approvals previously granted (5).

Finally, FDA may seek criminal convictions or civil money penalties. When a person (individual or corporate) submits a false statement of a material fact to FDA, or knowingly fails to disclose information required to be submitted (e.g., the number and severity of adverse events observed in a clinical trial), that person may be liable for a civil penalty, under the Food, Drug, and Cosmetic Act (6). Individuals may be fined up to $250,000 for each violation, while fines against manufacturers may reach $1 million

per violation. Civil fines may be imposed in addition to other authorized civil, criminal, and administrative remedies. Criminal prosecutions are typically based on claims of wire fraud, mail fraud, or submission of false statements to a federal agency. These actions are described below.

Government-wide Debarment and Similar Exclusions

Debarment is an extended exclusion from government grants and contracts, while suspension is a temporary exclusion. Both are viewed by federal agencies as serious actions to be used only to protect the federal government's interests and are not considered to be punishment (7). The actions nevertheless have decidedly punitive effects.

In order to avoid providing government support to anyone found guilty of serious misconduct, President Reagan in 1986 ordered that the debarment of an individual or institution by one agency should have government-wide effect. The order applies as well to suspensions, disqualifications, and "voluntary exclusion agreements," which the agency negotiates with individuals or entities who are willing to settle misconduct charges to avoid the cost and disruption of hearings and appeals. Accordingly debarment of a scientist or technician for research misconduct prevents that individual from receiving research support from any federal agency for the period of debarment (8). The same restrictions apply to any research institution or corporate entity that has been suspended or debarred. As of September 1997, however, no institution had been debarred as a result of a finding by the Department of Health and Human Services (HHS) Office of Research Integrity (ORI) confirming scientific misconduct (9). A suspended, debarred, disqualified, or excluded scientist may not even participate in another scientist's federally funded research without special permission from the funding agency. Periods of debarment or exclusion for scientific misconduct typically range from 3 to 10 years but sometimes are permanent (10). A consolidated list of all agency suspensions, debarments, disqualifications, and voluntary exclusions is maintained by the General Services Administration (GSA) for enforcement purposes and is available to the public (11).

Publication of Misconduct Findings

The PHS, FDA, and NSF policies differ in their approach to publicizing misconduct findings. The PHS publishes the names of the scientists and their research institution, together with a brief summary of their misconduct and the sanctions imposed. These notices appear in both hard copy and Internet versions of the *Federal Register*, the *NIH Guide to Grants and Contracts*, and the *ORI Newsletter*. The PHS also may notify state licensing boards (if the scientist involved is a licensed health practitioner), professional associations, and journals in which the scientist has published reports of past research. FDA publishes notices of disqualifications and debarment in the *Federal Register* and also may notify sponsors of products being tested and collaborating institutions. The list of investigators who are ineligible to receive investigational drugs, or whose use of investigational products is limited, is available to the public. By contrast, NSF publishes summaries of misconduct findings and sanctions, but it does not identify either the scientists or the institutions involved (12). The summaries of NSF cases are included in semiannual reports to Congress from the NSF Inspector General.

Retraction or Correction of Publications

Both PHS and NSF require a formal correction or retraction of journal articles found to contain fabricated, falsified, or plagiarized material. Although previously these notices were submitted as letters to the editor, most biomedical journal editors now agree that retractions should be labeled as such and appear independently on a numbered page of the journal, in order to include references to the retraction or correction in standard bibliographies (13). The National Library of Medicine, for example, annotates in its computerized databases (e.g., MEDLINE) articles that have been corrected or retracted, and provides a citation to the withdrawal or correction notice (14). This practice was challenged in 1994 by a scientist who was then under investigation for scientific misconduct as a result of patients improperly enrolled, by a collaborating researcher, in a multicenter breast cancer clinical trial. The principal investigator challenged the Library of Medicine's annotations of numerous articles from the collaborative trials but was rebuffed by a federal district judge, who ruled that the entries in the Library's databases pertained to publications, not to their authors (15). The district court's ruling was affirmed by a federal appellate court and motions for reconsideration were denied. In a letter to the editor in *Science*, the acting ORI director emphasized that the annotations had not been added to the Library's databases until after there had been a formal finding that the collaborating researcher had committed scientific misconduct. He added: "Scientists should not be concerned that annotations have been in the past or will be in the future placed in databases before a misconduct investigation is completed. They have not and will not be" (16).

Recoupment of Government Funds

Federal agencies have authority to require that institutions return any public funds that have been misused. In the context of research grants, this is typically accomplished by asserting that the funds in question were used improperly, and therefore the institution was not entitled to them and must refund the money (17). When ORI closes a case with a finding of scientific misconduct, it reports its findings to the institute at NIH that awarded the grant or contract. NIH in turn may seek recoupment of the research funds involved. In 1995, for example, NIH recovered $296,478 from an institution after ORI found that a principal investigator had submitted progress reports for three years describing research he had not performed. In 1994 NIH recovered over $1 million from three institutions involved in two scientific misconduct cases. NIH recoveries of research funds are actions independent of ORI's and are not routinely reported in the *ORI Newsletter*.

In late 1996 the Department of Justice sued for restitution of over $100 million from the University of Minnesota which allegedly obtained research grants fraudulently from NIH and illegally sold an antirejection drug that FDA had not approved for marketing (18). In July 1997 a U.S. district judge reduced the amount at stake from $109 million to $60 million by dismissing the False Claims Act portions of the suit (19). That ruling was reversed on appeal (20), and the university ultimately paid $32 million to settle the case (21).

At NSF, recoupments of research support are regularly reported in the Inspector General's Semiannual Reports to Congress and commonly result from findings of scientific misconduct. Restitution often results from criminal or civil litigation but may also result from internal agency determinations.

CIVIL MONEY PENALTIES

In addition to the administrative actions described above, scientists and their employers may be subject to civil money penalties and recoupment of publicly funded research support for offenses related to the development and testing of new drugs and devices. Other agencies have similar authority.

Program Fraud Civil Remedies. The Program Fraud Civil Remedies Act of 1986 was designed to deal with the submission of false statements to the government involving claims of less than $150,000. Under that Act, anyone submitting a claim or statement to a government agency, with knowledge that the claim or statement is false, fictitious, or fraudulent (or acting in deliberate ignorance or reckless disregard of the truth or falsity of the claim), is subject to a civil penalty of up to $5,000 for each such claim and twice the amount of each claim (22).

False Claims Act. If more than $150,000 is involved, the government may proceed under the False Claims Act, 31 U.S.C. §§3729–3730, which authorizes civil fines up to $10,000 plus recoupment of three times the amount of damages suffered by the agency as a result of the false claims. Alternatively, if the person who submitted the false claim cooperates with the government's investigation, the amount assessed may be reduced to two times the amount of damages. Under this provision universities and other research entities have been induced to cooperate and plea bargain, to avoid the treble damages.

In 1994, for example, the University of Utah and the University of California agreed to repay NIH more that $1.5 million in grants allegedly obtained through false data submitted in the grant applications. The universities' alternative to settlement was to risk treble damages (totaling $3.6 million) under the False Claims Act for knowingly presenting a false or fraudulent claim for payment to a government agency (23). Federal research grants are within the Act's definition of a "claim."

Qui Tam Actions. The "qui tam" provisions of the False Claims Act permit private citizens to bring suit in the name of the United States to recover funds paid out by the government of the basis of fraudulent claims. "Qui Tam" means "who as well ..." and denotes actions initiated by private citizens or informers who sue on behalf of the government as well as for themselves. In return for prosecuting the case, or at least alerting the government to the false claims, the informer (called a "relator") is entitled to a significant portion of the amount recovered. If the government successfully prosecutes the case, the relator may receive between 15 and 25 percent of the damages recovered. If the government declines to participate in the litigation, the relator who litigates in the name of the United States — and wins — is entitled to receive up to 30 percent of the damages awarded, in addition to costs and reasonable attorney fees. With damages trebled and potentially reaching millions of dollars, the informer's share can be sizable. Critics of the qui tam provisions say that the process offers an opportunity to settle personal scores while, at the same time, collecting a windfall and posing as a public citizen.

The qui tam provisions were enacted during the Civil War in response to sales of defective supplies to the Union Army. Amendments to the law in 1986 strengthened the role of relators and resulted in a surge of qui tam cases over the next decade. In fact, recoveries under the False Claims Act increased from $2 million in 1988 to over $200 million in fiscal 1995 (24). At the same time the portion of qui tam cases involving fraud related to HHS has surpassed those at the Department of Defense, which dominated the field in the past. HHS fraud cases involve primarily Medicare, Medicaid, and similar third parties who pay for health care services and supplies (24). A growing segment, however, relates to allegations of scientific misconduct in NIH-supported research activities.

A recent case illustrates how the qui tam law operates. In 1994 a former graduate student from Cornell, Pamela Berge, filed a qui tam suit against the University of Alabama, Birmingham, and four of its faculty members for allegedly submitting false statements in grant applications to NIH. A jury returned a verdict favorable to the informer/relator, which resulted in a judgment of just under $2 million (plus costs and attorney fees). Berge's claims were based on allegations of plagiarism, or misappropriation of intellectual property, which had been investigated and found to be meritless by a series of academic, scientific, and administrative reviewers. She therefore transformed her plagiarism charges into a qui tam action on behalf of the United States, asserting that a series of annual reports and grant applications submitted to NIH by the university incorporated plagiarized material and therefore constituted multiple false claims. Following the jury verdict, a federal district court judge awarded Berge 30 percent of the $1.6 million judgment, plus costs and attorney fees. The judgment subsequently was overturned by the Fourth Circuit Court of Appeals (25), which found no evidence on which a reasonable jury could have concluded that the challenged statements were even false, much less the basis for any NIH-funding decisions. The appellate court also ruled that there was no plagiarism and that Berge's claim for misappropriation of intellectual property was preempted

by the U.S. Copyright Act. The United States Supreme Court declined Berge's petition for review; thus the Fourth Circuit Court opinion stands as the last word in this case.

The likelihood that personal grievances will generate False Claims Act litigation is increased by the bounty-hunter (qui tam) provisions and the publicity attending successful cases. In addition a Washington-based group called Taxpayers Against Fraud actively solicits and supports potential qui tam plaintiffs, with a Web site on the Internet offering referrals to counsel, loans for litigation, help with legal research, and production of amicus (friend of the court) briefs. Supporters of qui tam view this as a public service, while opponents see disgruntled employees and students being encouraged to file multimillion dollar claims in the name of good citizenship (24). A similar support group has been established in Michigan to encourage and assist "whistleblowers" alleging scientific misconduct more generally. The likelihood is that agency actions and qui tam litigation both will increase in the foreseeable future.

CRIMINAL PROSECUTION

Statutory Penalties

Submitting false statements or information to the federal government, and similar offenses involving deceit, false statements, or fabrication, may result in criminal prosecution. It is a felony knowingly to submit a false statement or false claim to a government agency (26). A statement may be false either through omission (i.e., failure to disclose a material fact) or through submission of a false statement or representation. Conviction may result in incarceration for up to five years, debarment, imposition of fines, and recoupment of grant or contract monies.

When a scientist submits false statements in an application for a research grant, or in an annual report to the granting agency, that constitutes a false claim (request for money) and is punishable under the criminal False Claims Act by imprisonment for up to five years, a fine, or both (27). When the false claim is submitted by telephone or through the mail, the scientist and research institution may also be charged with wire fraud or mail

Table 2. Penalties for Scientific Misconduct: Federal Statutes

False claims and statements	**Penalties**
Administrative remedies for false claims and statements, 31 U.S.C. §§3801–3802 (applies to claims less than $150,000)	$5000 for each false claim, plus up to twice the amount of the claim
False Claims Act (civil actions), 31 U.S.C. §3729 (false claims exceeding $150,00)	$10,000 for each false claim, plus treble amount of each claim
False claims actions by private persons (qui tam), 31 U.S.C. §3730(d); Informer ("relator") may receive up to 30% of damages, plus attorneys' fees and costs	$10,000 for each false claim, plus treble the amount of each, plus costs and attorneys' fees
False, fictitious, or fraudulent claims (criminal), 18 U.S.C. §287	Fines and/or imprisonment up to 5 years
False statements or entries (criminal), 18 U.S.C. §1001	Fines up to $10,000 and/or imprisonment up to 5 years
Related crimes	**Penalties**
Conspiracy, 18 U.S.C. §371	Depends on underlying violation(s)
Mail fraud, 18 U.S.C. §1341 and Wire fraud, 18 U.S.C. §1343	Fines (up to $250,000, depending on amount of fraud), restitution, and/or imprisonment up to five years
Presidential order (Ronald Reagan)	**Penalties**
Debarment and suspension (government-wide effect), Executive Order No. 12549; *Fed. Regist.* **51**, 6370 (1986)	Individual or entity debarred or excluded by one federal agency may not receive grants or contracts from any federal agency
Food, Drug, and Cosmetic Act	**Penalties**
Withdraw approval of abbreviated drug applications, 21 U.S.C. §335c	Approval withdrawn; drug may not be marketed
Civil penalties (fines), 21 U.S.C. §335b; (informants may receive $250,000 or one-half of penalty, whichever is less)	Fine up to $250,000 for individuals; fine up to $1 million for corporate entities, partnerships, etc.
Debarment, and suspension; also, temporary denial of approval, 21 U.S.C. §335a	For corporation, partnership, etc., exclusion from research involving investigational products for 1–10 years; for individual convicted of felony related to product development, permanent exclusion
Public Health Service Act	**Penalties**
Office of Research Integrity, 42 U.S.C. §289b	Limitations on use of grant funds, supervision, suspension/termination of grant, debarment (exclusion) from future grants/contracts

fraud, and if more than one person is involved, a conspiracy charge may be added as well (28). Nothing in the law precludes prosecution under multiple federal statutes, and each grant application, status report, or request for payment constitutes a separate claim and, thus, an additional offense.

Examples of Cases

The first scientist to be indicted for research fraud was Stephen Breuning, a psychologist who fabricated data in federally supported studies on the use of stimulant drugs to treat hyperactivity in retarded children. Following indictment for fraud and making false statements to the National Institute of Mental Health, Breuning entered into a plea agreement which resulted in a conviction. Because of the potential impact of his fraud on treatment decisions for vulnerable individuals, Breuning was sentenced to incarceration for two years (suspended except for 60 days on work-release in a halfway house), five years' probation (during which he was required to perform 200 hours of community service), exclusion from federal research support for 10 years, and reimbursement of $11,352 to his university (29,30). He was also forbidden to work as a clinical psychologist for 10 years.

In a case that involved research submitted to FDA but did not involve federal research support, Dr. Robert Fogari, a physician who fabricated results over a period of eight years in a study of investigational anti-inflammatory drugs, was convicted of criminal fraud and obstruction of justice, sentenced to four years in jail, fined over $3.8 million, and ordered to make restitution (29).

Another case, that of Barry Garfinkle, demonstrates the multiple penalties that may be imposed for misconduct in clinical trials of new drugs (31). Dr. Garfinkle was convicted in 1993 of mail fraud and making false statements to the FDA while serving as principal investigator in studies of Anafranil. The drug was being tested as a treatment of obsessive-compulsive disorder in children and adolescents. The prosecution was triggered by complaints from the study coordinator that Garfinkle had ordered her to enter false data about weekly clinical evaluations that either never took place or were conducted by the coordinator rather than a physician. In addition the study coordinator alleged serious breaches of the research protocol. The indictment included charges under the Food, Drug, and Cosmetic Act, False Claims Act, and statutes prohibiting mail fraud.

Garfinkle was sentenced to six months in a halfway house, followed by six months of home detention, 400 hours of community service, and over $200,000 in fines. Based on his conviction of multiple felonies related to drug development, Garfinkle also was permanently debarred by FDA from serving in any capacity in connection with a new drug application. In addition the FDA served notice that it would not accept or review any abbreviated drug applications prepared by, or with the assistance of, Dr. Garfinkle, and that any person with a pending or approved drug application who knowingly used Dr. Garfinkle's services would be subject to a civil money penalty.

Such prosecutions are not limited to research involving NIH or FDA. In 1993, for example, a federal appellate court upheld the conviction of a scientist for conspiracy and fraud in connection with a grant, funded by the Agency for International Development, to create a diagnostic field test for malaria (32). Following a jury trial, the researcher was convicted and sentenced to eight months in prison (five of which were suspended), three years probation, and was ordered to make restitution in the amount of $75,000. The NSF also, together with the U.S. Attorney's Office, has successfully prosecuted individual scientists and biotech companies for fraud and false statements relating to Small Business Innovative Research (SBIR) grants. Within a six month period in 1996, these prosecutions resulted in criminal fines, civil penalties, restitutions, and other savings amounting to over $6 million (33).

DAMAGE TO PERSONAL HEALTH AND PROFESSIONAL REPUTATION

Perhaps the most devastating penalties are the effects of scientific misconduct allegations on the personal and professional life of the accused. Researchers who have been accused and later exonerated report that the investigative process alone has had prolonged and significant adverse effects on their lives (34). Approximately three-fifths of the exonerated scientists who responded to a 1996 survey believed they were stigmatized by the accusations, and nearly 40 percent reported adverse effects on their professional careers, such as damage to their reputation, reduced job mobility, and diminished opportunities for presenting papers. Over three-quarters of the respondents reported negative effects on their mental health, and nearly half reported adverse effects on their physical health. Disruptions of family relationships are not unusual. These outcomes, reported by scientists who were ultimately exonerated, suggest that even more serious personal and professional consequences must follow confirmation of scientific misconduct. Scientists who are found guilty of scientific misconduct, however, have not been surveyed. In many instances, they seem simply to have left the scene.

A widely publicized example of a scientist accused and later cleared is the case of Thereza Imanishi-Kari, who coauthored a paper in 1986 with Nobel prize-winning scientist David Baltimore (35). Imanishi-Kari was accused by a coworker of faking her data. Baltimore was never accused of scientific misconduct, but his name was linked invariably with that of Imanishi-Kari in the scientific and lay press, as well as in congressional hearings. Both scientists suffered personally and professionally throughout a 10 year investigation, although neither ultimately was found to have committed scientific misconduct. Imanishi-Kari's faculty status was suspended and, with it, her eligibility for NIH grants. Baltimore was ultimately forced from his position as President of Rockefeller University and was ostracized for years by many members of the scientific community. In 1997, with his reputation in recovery, he was appointed President of California Institute of Technology. Other scientists, perhaps less conspicuously, have suffered

a similar damage to reputation, personal anguish, curtailed career, and diversion of significant emotional, intellectual, and financial resources to defend against the accusations.

Occasionally the accusation of research fraud has led to even worse tragedy. Paul Kammerer, for example, was an Austrian scientist who believed he had developed proof of the heritability of acquired characteristics. As described by Arthur Koestler in *The Case of the Midwife Toad*, Kammerer's controversial support of the Lamarckian theory of inheritance was challenged repeatedly by the followers of Darwin. One vigorous critic ultimately accused Kammerer of having faked his results. Although essential facts about the research remain murky, the effect of the accusation is clearly documented. Despairing of ever proving his innocence to the satisfaction of his critics, Kammerer went into the Vienna woods and shot himself. A note found in his pocket stated:

> Dr. Paul Kammerer requests not to be transported to his home, in order to spare his family the sight. Simplest and cheapest would perhaps be utilization in the dissecting room of one of the university institutes. I would actually prefer to render science at least this small service. Perhaps my worthy academic colleagues will discover in my brain a trace of the quality they found absent from the manifestations of my mental activities while I was alive (36, p. 13).

Kammerer's suicide occurred in 1926. Six decades later, a professor of neurology and neurosurgery at the Montreal Neurological Institute (affiliated with McGill University), together with her husband (a faculty member at another university), committed suicide following publication of anonymous allegations that she had committed research fraud (37). As her lawyer explained: "Given that her work was her life, and she felt that her ability to continue was being seriously undermined, it was obviously more than she could live with" (38).

BIBLIOGRAPHY

1. 42 U.S.C. §289b, as amended June 10, 1993, Pub. L. 103-43, Tit. I, §§161, 163, 107 Stat. 140, 142. NIH regulations are found at 42 CFR, pt. 50; NSF regulations are at 45 CFR, pt. 689. ("CFR" refers to the Code of Federal Regulations.)
2. Public Health Service, *Fed. Regist.* **56**, 27384–27391 (1991).
3. 21 CFR §§58.200, 312.70 (drugs), 812.30(b) (devices).
4. 21 U.S.C. §335a.
5. 21 U.S.C. §§335a(c), 355(e)(5), 360 e(e) (1994).
6. 21 U.S.C. §335b.
7. HHS regulations, 45 CFR §76.115(b).
8. 48 CFR Subparts 9.4 (grants) and 309.4 (contracts); 45 CFR, pt. 76 (PHS); 45 CFR, pt. 620 (NSF).
9. Office of Research Integrity, *ORI Newsl.* **5**(4), 4–5 (1997).
10. 21 U.S.C. §335a(2).
11. 45 CFR §§76.500–76.510. Available at: *www.ARNET.gov/epls/*
12. D.P. Hamilton, *Science* **255**, 1346 (1992).
13. International Committee of Medical Journal Editors, *J. Am. Med. Assoc.* **277**, 927–934 (1997).
14. National Library of Medicine, *Fact Sheet: Errata, Retraction, and Comment Policy*, NLM, Washington, DC, 1989. Available at: *www.nlm.nih.gov/pubs/factsheets/errata.html*
15. *Fisher v. NIH et al.*, 934 F.Supp. 464; *aff'd. per curium* 107 F.3d 922 (1996); *reh'g denied* (D.C. Cir. March 28, 1997). The rulings are described in *ORI Newsl.* **2**, (Sept. 1996) **7** (June 1997).
16. C. Pascal, *Science* **274** (1996).
17. National Institutes of Health (NIH), *NIH Guide to Grants and Contracts* p. 24, (March 10, 1995), Note 2, citing 45 CFR §74.27 *et seq*.
18. *United States ex rel. Zissler v. Regents of the University of Minnesota*, 992 F.Supp. 1097 (D. Minn. 1998).
19. J. Basinger, *Chron. Higher Educ.* August 8, p. A36 (1997).
20. *United States ex rel. Zissler v. Regents of the University of Minnesota*, 1998 WL 560138 (8th Cir. Sept. 4, 1998).
21. U.S. Department of Justice, *University of Minnesota Pays $32 Million to Settle Allegations of Selling an Unlicensed Drug and Mishandling NIH Grant Funds*, press release, November 17, 1998.
22. 31 U.S.C. §3802.
23. 31 U.S.C. §3729.
24. P.R. Budeiri, *The Washington Lawyer*, September/October, 24–29 (1996).
25. *United States ex rel. Berge v. Board of Trustees of the University of Alabama, Birmingham, et al.*, 104 F.3d 1453 (4th Cir.); *cert. denied*, 522 U.S. 916, 118 S. Ct. 301 (1997).
26. 18 U.S.C. §1001 (1994).
27. 18 U.S.C. §287.
28. 18 U.S.C. §§1341, 1343, 371.
29. L.S. Little, *Food Drug Cosmetic Law J.* **46**, 65–69 (1991).
30. F.A. Kaufman, *United States v. Stephen E. Breuning*, Criminal No. K-88-0135 — An example of the applicability of the federal criminal legal process to fraudulent scientific research, in AAAS-ABA National Conference of Lawyers and Scientists, *Project on Scientific Fraud and Misconduct: Report on Workshop Number Three*, AAAS-ABA, Washington, DC, 1989, pp. 165–169.
31. *United States v. Barry Garfinkle*, 29 F.3d 1253 (8th Cir. 1994); Food and Drug Administration, *Fed. Regist.* **62**(63), 15713 (1997).
32. *United States v. Kellerman*, 992 F.2d 177 (8th Cir. 1993).
33. NSF, Office of Inspector General, *Semiannual Report to Congress*, No. 14, for the period Oct. 1, 1995–March 31, 1996, pp. 35–42, 52–54.
34. Research Triangle Institute, *Survey of Accused but Exonerated Individuals in Research Misconduct Cases*, a report submitted to the HHS Office of Research Integrity, Research Triangle Institute, Research Triangle Park, NC, 1996, pp. 45, 53.
35. D.J. Kevles, *The Baltimore Case*, Norton, New York, 1998.
36. A. Koestler, *The Case of the Midwife Toad*, Random House, New York, 1971.
37. L. Ratelle, *Chron. Higher Educ.* April 27, p. A36 (1994).
40. *Id.*, quoting Eric Maldoff, a Montreal attorney.

ADDITIONAL READINGS

Federal Agency Actions, Publications

Food and Drug Administration, Barry D. Garfinkel; Denial of hearing; Final debarment order. *Fed. Regist.* **62**(63), 15713 (1997).

National Institutes of Health Responsibilities of NIH and awardee institutions for the responsible conduct of research. *NIH Guide Grants Contracts* (March 10, 1995), p. 24, Note 2.

National Library of Medicine, *Fact Sheet: Errata, Retraction, and Comment Policy*, NLM, Washington, DC, 1989. Available at: *www.nlm.nih.gov/pubs/factsheets/errata.html*

National Science Foundation, Office of Inspector General, *Semiannual Reports to Congress* (published twice yearly), NSF, Washington, DC.

Office of Research Integrity (HHS), University of Minnesota pays $32 million in qui tam suit; Drops second lawsuit. *ORI Newsl.* December 1998, p. 5.

Ronald Reagan, Executive Order 12549: Debarment and Suspension, *Fed. Regist.* **51**, 6370 (1986), codified at 3 CFR, pts. 100 and 101, Executive Orders (1986 compilation).

U.S. Department of Justice, *University of Minnesota Pays $32 Million to Settle Allegations of Selling an Unlicensed Drug and Mishandling NIH Grant Funds*, Press Release, November 17, 1998.

U.S. Public Health Service, *Policies and procedures for dealing with possible scientific misconduct in extramural research*. *Fed. Regist.* **56**, 27384–27391 (1991).

Cases

Fisher v. NIH et al., 934 F.Supp. 464, *aff'd. per curium* 107 F.3d 922 (1996), *rehearing denied* (D.C. Cir. March 28, 1997).

United States ex rel. Berge v. University of Alabama et al., 104 F.3d 1453 (4th Cir.), *cert. denied*, 118 S. Ct. 301 (1997).

United States ex rel. Zissler v. University of Minnesota, 154 F.3d 870 (8th Cir. 1998).

United States v. Garfinkle, 822 F.Supp. 1457 (D. Minn.), *remanded*, 29 F.3d 451 (8th Cir. 1994), *aff'd.* 29 F.3d. 1253 (8th Cir. 1994).

United States v. Kellermann, 992 F.2d 177 (8th Cir. 1993).

Miscellaneous

AAAS/ABA National Conference of Lawyers and Scientists, *Project on Scientific Fraud and Misconduct: Report on Workshop Number Three*, AAAS, Washington, DC, 1989.

Agnew, B., False claims: Alabama wins, but stakes soar in Minnesota case. *J. NIH Res.* **9**, 30–31 (1997).

Basinger, J., Judge reduces amount University of Minnesota could have to pay in federal lawsuit. *Chron. Higher Educ.* August 8, 1997, p. A36.

Budeiri, P.R., The return of qui tam. *Washington Lawyer*, September/October, 1996, pp. 24–29.

Herman, K.G. et al., Investigating misconduct in science: The National Science Foundation model. *J. Higher Educ.* **65**(3), 384–400 (Appendix) (1994).

Kevles, D.J., *The Baltimore Case*, Norton, New York, 1998.

Koestler, A., *The Case of the Midwife Toad*, Random House, New York, 1971.

Little, L.S., The role of the HHS Office of Inspector General in the investigation of scientific misconduct. *Food, Drug, Cosmetic Law J.* **46**(1), 65–69 (1991).

Pascal, C.B., Misconduct annotations. *Science* **274**, (November 15 (1996).

Ratelle, L., Montreal researcher commits suicide following allegations of fraud. *Chron. Higher Educ.* April 27 (1994), p. A36.

Research Triangle Institute, *Survey of Accused But Exonerated Individuals in Research Misconduct Cases*, Report submitted to the HHS Office of Research Integrity, June 30, 1996.

Singer, M., Assault on science. *Washington Post*, June 26, 1996, p. A21.

See other entries SCIENTIFIC RESEARCH, ETHICS, SCIENTIFIC MISCONDUCT; SCIENTIFIC RESEARCH, ETHICS, VALUES IN SCIENCE.

SCIENTIFIC RESEARCH, POLICY, TAX TREATMENT OF RESEARCH AND DEVELOPMENT

DAVID L. CAMERON
Willamette University College of Law
Salem, Oregon

OUTLINE

INTRODUCTION

Research and development (R&D) in the field of biotechnology typically requires a significant investment before any marketable products may be produced. In an effort to encourage greater private sector investment in research and development activities, Congress has enacted several provisions under the Internal Revenue Code, including Sections 174, 41, and 45C, that authorize tax benefits for certain types of expenditures incurred in connection with R&D activities. Specifically, Section 174 provides a current tax deduction for research or experimental expenditures while Section 41 permits a tax credit for increases in research expenditures from one tax year to another.

Importantly, a taxpayer can take advantage of both the Section 174 deduction and the Section 41 tax credit with respect to many of the same research expenditures, subject only to the limitations of Section 280C. In addition Section 45C permits a tax credit for qualified clinical testing expenses incurred in the development of so-called orphan drugs. This article explores the requirements of Sections 174, 41, and 45C, as well as the limitations under Section 280C, and the implications of these statutory provisions in the context of R&D activities in biotechnology.

THE DEDUCTION OF RESEARCH AND EXPERIMENTAL EXPENDITURES UNDER SECTION 174

Overview

Under Section 174 a taxpayer may elect to deduct currently all research and experimental expenditures made in connection with the taxpayer's trade or business (1) or to amortize the expenditures over a period of not less than 60 months (2). If the taxpayer fails to treat the expenditures under one of these methods, then the research and experimental expenses are to be capitalized (3). The decision either to deduct Section 174 expenses currently or to defer and amortize them constitutes the adoption of an accounting method, which cannot be changed without the consent of the Commissioner of the Internal Revenue Service (IRS) (4). Although the current deduction of research and experimental expenditures is permitted for regular income tax purposes, for purposes of the alternative minimum tax, research and experimental expenditures must be capitalized and amortized ratably over a 10-year period beginning with the taxable year in which the expenditures were made unless the taxpayer materially participates in the activity within the meaning of Section 469(h) (5).

The "In Connection With" Standard

Importantly Section 174 applies to research or experimental expenditures paid or incurred "in connection with" a trade or business and thus is available to taxpayers who are not yet engaged in a trade or business. In this way the "in connection with" standard of Section 174 is distinguishable from, and less stringent than, the "carrying on" standard of Section 162, which allows the deduction of trade or business expenses more generally. In *Snow v. Commissioner* (6), the Supreme Court held that Section 174 does not require that the taxpayer actually be carrying on a trade or business in order to deduct otherwise qualified research and experimental expenditures. The Supreme Court based its decision on statements contained in the legislative history that Section 174 was intended to equalize the tax treatment of "small and growing businesses" with their "large and well-established competitors" (7). As a result a start-up business may deduct research and experimental expenditures under Section 174. Ordinarily, a start-up business must capitalize its start-up costs and amortize them over not less than a five-year period (8).

The fact that Section 174 applies to allow the deduction of research or experimental expenditures by a start-up business does not completely obviate the requirement that a trade or business exist, however. The courts have concluded that research and experimental expenditures are not deductible under Section 174 if the taxpayer does not have some realistic prospect of entering into a trade or business involving the technology under development or, in fact, never eventually enters into the active conduct of a trade or business. For example, in *Harold J. Green* (9), the Tax Court stated:

> Although the Supreme Court established in *Snow* that the taxpayer need not currently be producing or selling any product in order to obtain a deduction for research expenses, it did not eliminate the "trade or business" requirement of section 174 altogether. For section 174 to apply, the taxpayer must still be engaged in a trade or business *at some time*, and we must still determine, through an examination of the facts of each case, whether the taxpayer's activities in connection with a product are sufficiently substantial and regular to constitute a trade or business for purposes of such section (10).

The requirement that deductions under Section 174 be incurred "in connection with" a trade or business also means that Section 174 is unavailable to taxpayers who merely fund the research of a third party for the development of a product when the taxpayer does not possess the intent or ability to exploit the fruits of that research on its own. These arrangements frequently involve partnerships that enter into agreements to fund research and simultaneously lease or license the rights of that research to the developing party. In such cases the courts closely scrutinize the business arrangements involved and have frequently held that the partnership is not sufficiently engaged in a trade or business to satisfy the requirements of Section 174 (11). This determination, however, must be based on an examination of all the facts of the particular situation (12). As the Tax Court has stated:

> [W]hen a partnership contracts out the performance of the research and development in which it intends to engage, all of the surrounding facts and circumstances are relevant to the inquiry into whether it has any realistic prospect of entering into a trade or business with respect to the technology under development. The inquiry includes consideration of the intentions of the parties to the contract for the performance of the research and development, the amount of capitalization retained by the partnership during the research and development contract period, the exercise of control by the partnership over the person or organization doing the research, the existence of an option to acquire the technology developed by the organization conducting the research and the likelihood of its exercise, the business activities of the partnership during the years in question, and the experience of the partners. Absent a realistic prospect that the partnership will enter a trade or business with respect to the technology, the partnership will be treated as a passive investor, not eligible for deductions under section 174 (13).

For example, in *Kantor v. Commissioner* (14), the Ninth Circuit Court of Appeals affirmed the decision of the Tax Court denying the taxpayer's claimed deductions under Section 174. In *Kantor*, PCS, Ltd. was formed to fund the adaptation of a particular computer software program for use on various types of computer systems. PCS, Ltd. entered into an R&D agreement with PCS, Inc. under

which PCS, Inc. agreed to perform the research involving the software program with PCS, Ltd. retaining ownership of the resulting software. The parties also entered into a technology transfer agreement under which PCS, Ltd. granted PCS, Inc. an option to acquire an exclusive worldwide license to market the software on payment of $5,000 and future royalties based on prospective sales. PCS, Ltd. also granted PCS, Inc. a 13-month review period in which to exercise its option.

The Ninth Circuit agreed with the Tax Court that PCS, Ltd. did not possess the objective intent nor the capability of marketing any software that might be developed. Importantly the court noted that by granting PCS, Inc. the right to market the software on payment of a nominal fee, PCS, Ltd. "made it more probable than not that the research firm [PCS, Inc.] would exercise these rights if the software that resulted from the research was at all valuable" (15). The court reached its conclusion that PCS, Ltd. did not posses the capability to market the software despite the fact that the general partner of PCS, Ltd. had significant sales and technical experience in the computer industry, was actively engaged in the research effort, negotiated an arrangement to secure financing, and negotiated licensing agreements for the marketing of the software. Rather, the court likened the general partner's activities to "those of any investor who applies his knowledge and expertise to insure that an investment is successful. . ." (16). The court concluded that permitting a deduction in such circumstances would simply allow the taxpayer to use Section 174 to deduct its capital investment.

Nevertheless, in *Scoggins v. Commissioner* (17), the Ninth Circuit Court of Appeals concluded that a research partnership was entitled to deductions under Section 174 despite the existence of facts similar to those in *Kantor*. In *Scoggins*, B&B Research and Development Partnership (B&B) contracted with Epitaxy Systems, Inc. (Epitaxy) to perform research to develop an epitaxial reactor. B&B agreed to provide Epitaxy with up to $500,000, consisting of $43,000 to be paid at the time of executing the agreement and additional amounts to be paid at the partnership's discretion. B&B granted Epitaxy a 15-month nonexclusive license to market the technology in return for royalties of 20 percent on net revenue from sales during that period. B&B also granted Epitaxy an option exercisable for one year thereafter to purchase the developed technology for $5 million.

The Ninth Circuit reversed the decision of the Tax Court and permitted the taxpayer's deductions under Section 174. Significantly the two partners in B&B, who also owned the controlling interests in Epitaxy, had previously developed and successfully marketed epitaxial reactors. The court concluded that B&B therefore possessed the technical expertise and financial ability to conduct a trade or business using the developed technology. The court also noted that the license granted to Epitaxy was nonexclusive and that Epitaxy was under no obligation to market the technology. Significant, too, was the amount of the payment to acquire the technology, $5 million, compared to the research cost. The court specifically noted that this fee was a significant impediment to Epitaxy's ability to engage in the marketing of the technology and contrasted this payment to the nominal amount to be paid in *Kantor*.

Definition of Research and Experimental Expenditures

For purposes of Section 174, "research or experimental expenditures" are limited to those expenditures incurred in connection with the taxpayer's trade or business that "represent research and development costs in the experimental or laboratory sense" (18). The regulations expand on this definition through the adoption of an "uncertainty test," conditioned on the quality of the information available to the taxpayer at the time that the expenditures are undertaken. The regulations provide that

> [e]xpenditures represent research and development costs in the experimental or laboratory sense if they are for activities intended to discover information that would eliminate uncertainty concerning the development or improvement of a product. Uncertainty exists if the information available to the taxpayer does not establish the capability or method for developing or improving the product or the appropriate design of the product. Whether expenditures qualify as research or experimental expenditures depends on the nature of the activity to which the expenditures relate, not the nature of the product or improvement being developed or the level of technological advancement the product or improvement represents (19).

Under the regulations, the word "product" is defined to include "any pilot model, process, formula, invention, technique, patent, or similar property, and includes products to be used by the taxpayer in its trade or business as well as products to be held for sale, lease, or license" (20).

While considerable uncertainty exists with respect to the precise contours of the definition of "research or experimental expenditures," the regulations are clear that the costs of obtaining a patent, including such expenditures as attorneys' fees expended to make and perfect a patent application, are included within the definition (21). Because the costs of perfecting a patent are "inextricably a part of the research and development work," such costs reasonably fall within the definition of expenditures for research or experimentation (22).

The application of Section 174 to the development of property protected by copyright law, as opposed to patent law, requires additional consideration. As previously described, research or experimental expenditures must "represent research and development costs in the experimental or laboratory sense" (23). Because such expenditures are limited to the reasonable costs incident to the development or improvement of a product including any pilot model, process, formula, invention, technique, or similar property (24), this definition by itself might exclude from Section 174 any research and development costs typically incurred in creating a work subject to copyright protection. In addition the regulations explicitly state that expenditures incurred for "[r]esearch in connection with literary, historical, or similar projects" do not fall within the definition of research or experimental

expenditures (25). In Revenue Ruling 72-395 (26), the IRS expanded on this regulatory provision and concluded that Section 174 did not apply to permit the current deduction of the costs of writing and editing, as well as design and art work for, textbooks and visual teaching aids (27). Importantly, however, the exclusion for literary, historical, and similar projects does not apply to certain expenditures incurred in connection with the development of computer software despite the fact that computer software is frequently protected under the copyright laws (28).

Although the precise contours of the definition of research or experimental expenditures are not fully described under the regulations or in other IRS pronouncements, the regulations are clear that certain costs incurred in the development of a product do not qualify as research or experimental expenditures under Section 174. For example, costs incurred for ordinary testing or inspection of materials for purposes of quality control — in addition to costs incurred for efficiency surveys, management studies, consumer surveys, advertising, or promotions — do not qualify as research or experimental expenditures (29). The regulations also specify that research or experimental expenditures do not include the acquisition costs of another's "patent, model, production or process" (30). The costs of materials or labor used in the construction, installation, acquisition, or improvement of property are also excluded from the definition of research or experimental expenditures (31). In addition Section 174 does not apply to expenditures for the acquisition or improvement of land or depreciable property (32). However, allowances for depreciation may be considered research or experimentation expenditures under Section 174 to the extent that the property is used in connection with activities involving research or experimentation (33).

A significant issue arising under Section 174 concerns the proper treatment of expenditures to a third party to develop a product or process, the expenses of which would otherwise qualify as research or experimental expenditures under Section 174 if incurred directly by the taxpayer. The regulations are clear that the definition of research or experimental expenditures includes expenditures paid or incurred for research or experimentation carried on on the taxpayer's behalf by another person or organization except to the extent that the expenditures are for the acquisition or improvement of land or depreciable property used in connection with the research or experimentation and to which the taxpayer acquires rights of ownership (34). In addition, the regulations under Section 174 provide that research or experimental expenditures made to a third party for the construction or manufacture of depreciable or amortizable property are deductible under Section 174(a) only if "made upon the taxpayer's order and at his risk" (35). Consequently no deduction is allowed if the property acquired is regularly produced or is constructed or manufactured under a performance guarantee. A performance guarantee includes any guarantee, whether express, implied, or imposed by local law, that concerns quality of production or quantity of production in relation to the consumption of raw materials and fuel, unless the guarantee is limited to engineering specifications such that economic utility is not taken into account (36).

Section 174 Election

A taxpayer's election currently to deduct all research or experimental expenditures applies to all qualifying expenditures (37). With the permission of the Commissioner, however, a taxpayer may change to the deferred method with respect to those expenditures attributable to a particular project or projects. In no event will a taxpayer be permitted currently to deduct some of the expenditures and defer and amortize other expenditures relating to the same project (38). Finally, even with the consent of the Commissioner to defer expenses attributable to a particular project, a taxpayer who originally elected currently to deduct research or experimental expenditures may not defer those expenses attributable to a new project without again obtaining the consent of the Commissioner (39).

Instead of currently deducting research or experimental expenditures under Section 174(a), a taxpayer may elect to capitalize and amortize such expenditures over not less than a five-year period under Section 174(b) (40). Under the regulations the amortization period begins with the month in which the taxpayer first realizes benefits from the expenditures. According to the regulations, the taxpayer will typically be deemed to realize benefits from any deferred expenditures in the month in which "the taxpayer first puts the process, formula, invention, or similar property to which the expenditures relate to an income-producing use" (41). Importantly, the taxpayer may select amortization periods of differing lengths for deferred expenditures attributable to different projects.

As a basic rule, deferral under Section 174 is permitted only for those expenditures that are otherwise chargeable to a capital account and that are not chargeable to depreciable or depletable property. However, such expenditures may be deferred only in part under Section 174 if they later become chargeable to depreciable or amortizable property, such as patents (42). In that case any unrecovered expenditures are to be either depreciated or amortized under Section 167 at the time that "the asset becomes depreciable in character" (43). To illustrate this situation, the regulations provide the following example:

> [F]or the taxable year 1954, *A*, who reports his income on the basis of a calendar year, elects to defer and deduct ratably over a period of 60 months research and experimental expenditures made in connection with a particular project. In 1956, the total of the deferred expenditures amounts to $60,000. At that time, *A* has developed a process which he seeks to patent. On July 1, 1956, *A* first realized benefits from the marketing of products resulting from this process. Therefore, the expenditures deferred are deductible ratably over the 60-month period beginning with July 1, 1956 (When *A* first realized benefits from the project). In his return for the year 1956, *A* deducted $6,000; in 1957, *A* deducted $12,000 ($1,000 per month). On July 1, 1958, a patent protecting his process is obtained by *A*. In his return for 1958, *A* is entitled to a deduction of $6,000, representing the amortizable portion of the deferred expenses attributable to the period prior to July 1, 1958. The balance of the unrecovered expenditures ($60,000 minus $24,000, or $36,000) is to be recovered as a depreciation deduction over the life of the patent commencing with July 1, 1958. Thus, one-half of the annual depreciation deduction based upon the useful life of the patent is also deductible for 1958 (from July 1 to December 31) (44).

Consequently the opportunity to defer research or experimental expenditures under Section 174 in connection with assets that are otherwise amortizable under Section 167 may be limited.

CREDIT FOR INCREASING RESEARCH ACTIVITIES UNDER SECTION 41

Overview

Section 41 establishes a research tax credit in connection with "qualified research expenses" and "basic research payments," frequently referred to as the incremental research credit and the basic research credit, respectively. The incremental research credit is equal to 20 percent of the excess of the qualified research expenses for the tax year in excess of a base amount, and the basic research credit is equal to 20 percent of any basic research payments in excess of a base amount (45). The purpose of the credit is to provide an incentive for increased research activities in the private sector. According to the legislative history to the research credit as originally enacted in 1981, "Congress concluded that a substantial tax credit for incremental research and experimental expenditures was needed to overcome the reluctance of many ongoing companies to bear the significant costs of staffing and supplies, and certain equipment expenses such as computer charges, which must be incurred to initiate or expand research programs in a trade or business" (46).

The incremental research credit is available in connection with the costs of any qualified research activities incurred in carrying on the trade or business of the taxpayer, including both in-house expenses and contract research expenses (47). Originally the "carrying on" standard of Section 41 generally corresponded to the same requirement found in Section 162 and thus was more restrictive than the "in connection with" standard of Section 174 (48). According to the legislative history,

> it is intended that to be eligible for the credit, research expenditures must be paid or incurred in a particular trade or business being carried on (within the meaning of sec. 162) by the taxpayer; no credit is available for expenditures for research relating to a potential trade or business which the taxpayer is not carrying on at the time the research expenditures are made. Thus, the credit is not available (either for current or carryover use) to a new entity which undertakes research with a view to using the resulting technology through future production and sales, and is not available to an ongoing business which undertakes research with a view to entering a new trade or business (49).

This requirement has been relaxed somewhat in connection with in-house research expenses. A taxpayer will be treated as meeting the trade-or-business requirement of Section 41 with respect to in-house research expenses if, at the time the expenses are paid or incurred, the principal purpose of the taxpayer in making the expenditure is to use the results of the research in the active conduct of a future trade or business of the taxpayer or of another member of the same controlled group (50).

In determining the amount of the research credit, members of a controlled group of corporations are to be treated as a single taxpayer, and the credit, if any, is to be allocated among the members in proportion to their respective increase in qualified research expenses (51). Similar aggregation and allocation rules apply in connection with the research credit available to S corporations, partnerships, estates, and trusts that are under common control (52).

Incremental Research Tax Credit

Definition of Qualified Research Expenses. The incremental research tax credit is equal to 20 percent of the qualified research expenses for the tax year in excess of a base amount (53). The definition of qualified research expenses under Section 41 is more narrow than the definition of research or experimental expenditures under Section 174. This is reflected in the three-part definition of qualified research under Section 41(d):

1. Research that satisfies the requirements of Section 174.
2. Research undertaken for the purpose of discovering information that is technological in nature, whose application is intended to be useful in the development of a new or improved business component of the taxpayer.
3. Research where substantially all of the activities constitute elements of a process of experimentation related to the development of a new or improved function, performance, reliability or quality of a business component (54).

Consequently the second and third parts of the three-part definition limit the range of activities that constitute qualified research as compared to those activities that constitute research or experimental activities for purposes of Section 174.

The second part of the three-part definition requires that the research be designed to discover information that is technological in nature. The legislative history states that qualified research must be within the basic sciences:

> The determination of whether new or improved characteristics of a business item are technological in nature depends on whether the process of experimentation to develop or improve such characteristics fundamentally relies on principles of the physical or biological sciences, engineering, or computer science — in which case the characteristics are deemed technological — or on other principles, such as those of economics — in which case the characteristics are not to be treated as technological. For example, new or improved characteristics of financial services or similar products (such as new types of variable annuities or legal forms) or advertising do not qualify as technological in nature (55).

Importantly research does not rely on principles of computer science merely because a computer is employed (56).

The proposed regulations under Section 41 provide definitions for the terms "discovering information" and "technological in nature." According to the proposed regulations, "discovering information" means "obtaining

information that exceeds, expands, or refines the common knowledge of skilled professionals in a particular field of technology or science" (57). The examples contained in the proposed regulations show clearly that research may constitute qualified research when the information is known to others but remains a closely guarded secret and is beyond the common knowledge of skilled professionals in the relevant fields (58). Research may also constitute qualified research even when the taxpayer abandons the project and attempts to develop the technology prove unsuccessful (59). Finally, research may be qualified only in part where research activities advance to the point where further analysis is within the common knowledge of skilled professionals in the relevant field (60). Relying on the language cited from the legislative history, the proposed regulations also provide that information is "technological in nature" if "the process of experimentation used to discover such information fundamentally relies on principles of physical or biological sciences, engineering, or computer science" (61).

The third part of the three-part definition requires that the research constitute a process of experimentation related to the development of a new or improved function, performance, reliability or quality of a business component. In defining the term "process of experimentation," the legislative history refers to

> a process involving the evaluation of more than one alternative designed to achieve a result where the means of achieving that result is uncertain at the start. This may involve developing one or more hypotheses, testing and analyzing those hypotheses (through, for example, modeling or simulation), and refining or discarding the hypotheses as part of a sequential design process to develop the overall component (62).

Although this definition of the term "process of experimentation" suggests that the result itself must be uncertain, subsequent legislative history has emphasized that only the means of reaching the result need be uncertain. "Thus, even though a researcher may know of a particular method of achieving an outcome, the use of the process of experimentation to effect a new or better method of achieving that outcome may be eligible for the credit. . ." (63). As examples of processes of experimentation, the legislative history specifically notes experiments undertaken by chemists or physicians in developing and testing a new drug and the work of engineers in designing a new computer system or an improved or new integrated circuit.

The proposed regulations under Section 41 follow the legislative history in describing a process of experimentation (64). Interestingly the focus of both the legislative history and the proposed regulations on uncertainty in defining a process of experimentation suggests a certain similarity to the uncertainty test under Section 174 (65). Nevertheless, the proposed regulations provide, as an illustration, that expenditures incurred to resolve uncertainty in manufacturing an improved widget may be treated as research or experimental expenditures under Section 174 but not be undertaken to obtain knowledge that exceeds, expands, or refines the common knowledge of skilled professionals in the relevant technological fields necessary to satisfy the requirements under Section 41 (66). In addition the proposed regulations provide that, in testing and analyzing one or more hypotheses, the taxpayer must design a "scientific experiment" that, "where appropriate to the particular field of research, is intended to be replicable with an established experimental control" (67). Finally, the statutory definition of qualified research requires that substantially all of the activities that constitute the process of experimentation must relate to the function, performance, reliability, or quality of a business component (68). The proposed regulations provide that the "substantially all" requirement is satisfied only if 80 percent or more of the research activities, measured on a cost or other consistently applied basis, constitute elements of a process of experimentation (69).

Importantly the test of whether particular research is to be treated as qualified research is determined with respect to each business component. A "business component" is defined to include any "product, process, computer software, technique, formula, or invention" held by the taxpayer for sale, lease, or license or used by the taxpayer in its trade or business (70). The proposed regulations under Section 41 provide that the research credit is not available for research activities relating to the development of a manufacturing or other commercial production process unless the activities satisfy the requirements of Section 41 without taking into account the research activities related to the development of the product (71). Similarly the research credit is not available for research activities relating to the development of a product unless the activities satisfy the requirements of Section 41 without taking into account the research activities related to the development of the manufacturing or other commercial production process (72).

The tests to establish eligibility for the credit are applied to each business component or sub-component under the so-called "shrinking-back" concept:

> [T]he requirements for credit eligibility are applied first at the level of the entire product, etc. to be offered for sale, etc. by the taxpayer. If all aspects of such requirements are not met at that level, the test applies at the most significant subset of elements of the product, etc. This shrinking back of the product is to continue until either a subset of elements of the product that satisfies the requirements is reached, or the most basic element of the product is reached and such element fails to satisfy the test (73).

This "shrinking back" concept allows the taxpayer to obtain the benefits of the credit with respect to any portion of its research activities undertaken in the development of a business component that satisfies the requirements of Section 41 (74).

In *Norwest Corporation* (75), the Tax Court considered the definition of qualified research in the context of the development of internal use computer software. The court interpreted Section 41(d) as imposing the following four separate tests:

1. The Section 174 Test (Test 1), which requires that the research expenditures qualify as expenses under Section 174.

2. The Discovery Test (Test 2), which limits the type of information discovered to that which is technological in nature.
3. The Business Component Test (Test 3), which requires that the taxpayer's activities provide some level of functional improvement to a business component of the taxpayer.
4. The Process of Experimentation Test (Test 4), which requires that initial uncertainty concerning the technical ability of the taxpayer to develop the product be eliminated through the development, testing, and analyzing of one or more hypotheses as part of a sequential design process to develop the overall component.

The Tax Court provided content to each of these tests by relying on the legislative history of Section 41 and the policy objectives of the research tax credit.

With respect to Test 2, the Discovery Test, the court concluded that in order to be eligible for the credit, an objective of the taxpayer's activities must be the creation of new knowledge in the field in which the taxpayer is working:

> The legislative history of Section 41 dictates that the knowledge gained from the research and experimentation must be that which exceeds what is known in the field in which the taxpayer is performing the research and experimentation — in this case the computer science field. The fact that the information is new to the taxpayer, but not new to others, is not sufficient for such information to come within the meaning of discovery for purposes of this test. The purpose of the R&E credit was to stimulate capital formation and improve the U.S. economy — not merely the taxpayer's business (76).

The court referred to the legislative history of Section 41 to conclude that the discovery of information must concern the principles of the hard sciences which could result by either "expanding" or "refining" those principles (77). The court also concluded that the Discovery Test of Section 41 differs from the uncertainty test of Section 174. The court pointed out that the uncertainty test under the Section 174 regulations was not adopted until 1994, eight years after the introduction of the Discovery Test under Section 41 in 1986. In addition the court cited the legislative history to Section 41 to suggest that, in amending Section 41 in 1986, "Congress sought to tighten the requirements for obtaining the R&E credit" (78). The court reasoned that because Congress did not change the requirements of Section 174 at that time, the congressional purpose could only be achieved by viewing the two tests as different. Finally, the court viewed the uncertainty test of Section 174 and the Discovery Test of Section 41 as relating to the discovery of different types of information. According to the court the regulations under Section 174 refer to "uncertainty concerning the development or improvement of a product" while Section 41 relates to information that is "technological in nature" and which "fundamentally relies on principles of the hard sciences."

In reaching its decision, the Tax Court also considered Tests 3 and 4. Unfortunately, the court provided only a limited description of Test 3, the Business Component Test. The court only noted that the taxpayer's activities must provide some level of functional improvement, at a minimum, to a business component of the taxpayer. The court provided a greater explanation of its view of Test 4, the Process of Experimentation Test. After quoting from the legislative history, the court stated that this test requires a more structured method of discovery than that required under Section 174, a process in which one or more hypotheses must be developed, tested, and analyzed. According to the court this requirement is to be applied in concert with the shrinking back test until the 80 percent standard is satisfied or the most basic element of the product is reached and that element fails to satisfy the standard. The court concluded by noting that the shrinking back test must be examined on a case-by-case basis to determine which activities are part of the same process or product and which are sufficiently discrete as to warrant separate evaluation.

Exclusion of Certain Research Expenses. Costs incurred in connection with several types of research activities are statutorily excluded from the definition of qualified research expenses under Section 41. Research expenditures are excluded if the research relates only to style, taste, cosmetic, or seasonal design changes (79). More importantly perhaps, qualified research expenses do not include expenses incurred in connection with research conducted after commercial production of a component has started (80). According to the legislative history, "commercial production" is achieved when "the component has been developed to the point where it either meets the basic functional and economic requirements of the taxpayer for such component or is ready for commercial sale or use" (81). The proposed regulations provide that the following activities are deemed to occur after the beginning of commercial production of a business component: pre-production planning for a finished business component, tooling-up for production, trial production runs, troubleshooting involving detecting faults in production equipment or processes, and debugging or correcting flaws in a business component (82). In addition the legislative history states that the credit is not available for the costs of additional clinical testing of a pharmaceutical product after the product is made commercially available, except when the testing is necessary to establish new functional uses for the existing product. For example, "testing a drug currently used to treat hypertension for a new anti-cancer application, and testing an antibiotic in combination with a steroid to determine its therapeutic value as a potential new anti-inflammatory drug, are eligible for the credit" (83).

Qualified research expenses also do not include the costs of research designed to reproduce an existing business component (84). This provision is intended to exclude the costs of "reverse engineering" activities from eligibility for the credit and applies to the reproduction of an existing component by another person based on a physical examination of the business component or on plans, blueprints, detailed specifications, or publicly available information (85). In Private Letter Ruling 9346006 (a Technical Advice Memorandum), the IRS invoked this

exclusion and denied the taxpayer's claimed research credit in connection with the development of the generic form of certain drugs that had previously received FDA approval and for which information concerning active and inactive ingredients was publicly available (86). However, this exclusion does not apply if a taxpayer examines a competitor's product in developing its own component through a process of otherwise qualified experimentation.

Section 41 also excludes from the definition of qualified research expenses the costs of research that is fully funded by another entity (87). Under the regulations, research does not constitute qualified research to the extent that it is funded by a grant, contract, or otherwise by another person, including any governmental entity. However, amounts payable under any agreement that are contingent on the success of the research are considered as paid for the results of the research and are not treated as funding (88). In addition research is considered fully funded if the taxpayer retains no substantial rights in the products of the research under the terms of the agreement providing for the performance of the research (89).

The exclusion for fully-funded research was the subject of Private Letter Ruling 9410007 (a Technical Advice Memorandum). In this ruling, the taxpayer was engaged in fixed price contracts with the U.S. government to conduct research and develop certain types of equipment. Under the contracts the taxpayer retained title and rights to any inventions developed as a result of the research, including patents and copyrights, subject to a "non-exclusive, nontransferable, irrevocable" license in favor of the government. The IRS rejected the taxpayer's claimed tax credit under Section 41 for expenditures incurred in conducting the research required under the contracts. Because the contracts provided for progress payments that were subject to little risk of termination or withholding by the government, the IRS viewed the payments for the contracted research as "expected and likely in the normal course of events." As a result the IRS concluded that, because the payments were not contingent on the success of the taxpayer's research, the research was fully funded. In addition the IRS concluded that the research was fully funded because the taxpayer retained no substantial rights in the research. The IRS reasoned that the taxpayer's rights to use and transfer the technology, copyrights, and technical data resulting from the research were subject to "significant restriction" by the U.S. government (90).

The exclusion from the definition of qualified research for research that is considered fully funded was also the subject of review in *Lockheed Martin Corp. v. United States* (91). As previously noted, the regulations provide that research will be treated as fully funded where the taxpayer retains no substantial rights in the research under the agreement to provide the research. In *Lockheed*, the taxpayer performed research under a number of defense contracts with the federal government and claimed a tax refund of over $63 million under Section 41. The court rejected the taxpayer's claim that the "substantial rights" requirement under the regulations was invalid. The court also rejected the claim that a taxpayer fails to satisfy the substantial rights requirement only when the taxpayer retains *no* rights to the research. Instead, the court looked to Section 1235 for guidance and concluded that the government's unlimited right to use and disclose the research results, as well as the considerable restrictions on the taxpayer's ability to use the research results in the form of security classifications and export restrictions, prevented the taxpayer from claiming that it retained substantial rights in the research. Consequently none of the taxpayer's research expenses qualified for the research credit.

The definition of qualified research expenses also excludes the costs of preparing various types of surveys or studies (92). Under the proposed regulations, this exclusion applies to efficiency surveys; activities (e.g., studies) related to management functions or techniques, market research, market testing, or market development (e.g., advertising or promotions); routine data collections; or routine or ordinary testing or inspection of materials or business components for quality control. Management functions and techniques include the preparation of financial data and analysis, development of employee training programs and management organization plans, and management-based changes in production processes (e.g., rearranging work stations on an assembly line) (93).

Finally, qualified research expenses do not include the costs of research designed to adapt an existing component to a particular customer's needs (94); research involved in the preparation of certain types of computer software (95); research conducted outside the United States, Puerto Rico, or any possession of the United States (96); and research in the social sciences or humanities (97).

Determination of the Incremental Research Tax Credit. Provided that the research at issue satisfies the definition of qualified research, Section 41 allows a tax credit in an amount equal to 20 percent of the qualified research expenses for the taxable year in excess of a base amount (98). Qualified research expenses for the taxable year include both in-house and contract research expenses (99). In-house research expenses include wages paid to employees engaged in qualified research or directly supervising or supporting activities that constitute qualified research (100), the cost of supplies used in the conduct of qualified research (101), and the cost of computers and computer time used in qualified research efforts (102). Contract research expenses, on the other hand, are generally limited to 65 percent of any amount paid or incurred by the taxpayer to a person other than an employee for qualified research (103). However, if the contracting party that will conduct the research is a qualified research consortium, 75 percent of any amounts paid or incurred by the taxpayer to the consortium for qualified research on behalf of the taxpayer and one or more unrelated taxpayers will be treated as qualified research expenses eligible for the credit (104). Contract research expenses that are prepaid are considered as paid or incurred during the period in which the qualified research is actually conducted (105).

With respect to contract research expenses, the regulations specifically require that the qualified research be performed "on behalf of the taxpayer" and that the payments not be contingent on the success of the research (106). The

former requirement is satisfied even if the taxpayer retains only a nonexclusive right to the research results (107). In *Norwest Corp.* (108), the IRS argued that the taxpayer's right under a software development agreement to a "perpetual, nontransferable, nonexclusive and . . . royalty-free license" to use the developed software did not satisfy the regulatory requirements. The IRS suggested that a difference existed between rights to research results and rights to any final product. The court rejected this argument stating that "the right to use the results of the research without paying for that right is at least a right to the research results as that term is applied [under the regulations] — although it may or may not constitute 'substantial rights in the research' within the purview of the regulations" (109). The IRS also maintained that the taxpayer's ability to terminate the development agreement at selected times violated the regulatory requirement that payments not be contingent on the success of the research. The court rejected this argument as well, noting that the taxpayer had no ability under the agreement to recover any payments that it might have previously made.

Importantly a contract research expense will not be a qualified research expense if the product or result of the research is intended to be transferred to another in return for a license or royalty and the taxpayer does not use the product of the research in the taxpayer's trade or business (110). In such a situation, the taxpayer will not be deemed to be engaged in a trade or business to satisfy the "carrying on" requirement of Section 41(b). The legislative history was emphatic about this point:

> [U]nder the trade or business test of new section [41], the credit generally is not available with regard to a taxpayer's expenditures for "outside" or contract research intended to be transferred by the taxpayer to another in return for license or royalty payments. (Receipt or royalties does not constitute a trade or business under present law, even though expenses attributable to those activities are deductible from gross income in arriving at adjusted gross income.) In such a case, the nexus, if any, between research expenditures of the taxpayer and activities of the transferee to which research results are transferred (e.g., any use by an operating company, that is a general partner in a limited partnership which make the research expenditures, of the research results in the operating company's trade or business) generally will not characterize the taxpayer's expenditures as paid or incurred in carrying on a trade or business *of the taxpayer.* (Under appropriate circumstances, nevertheless, the nexus might be deemed adequate for purposes of the section 174 deduction elections.) If, however, the taxpayer used the product of the research in a trade or business of the taxpayer, as well as licensing use of the product by others, the relationship between the research expenditures of the taxpayer (i.e., those research expenditures paid or incurred after such time as the taxpayer is considered to be carrying on the trade or business in which such expenditures are paid or incurred) and the taxpayer's trade or business in which the research expenditures are paid or incurred generally would be sufficient for credit purposes (111).

As previously noted, Section 41 allows a tax credit in an amount equal to 20 percent of the qualified research expenses for the taxable year in excess of a base amount (112). The base amount is determined by multiplying the average annual gross receipts of the taxpayer for the four taxable years preceding the taxable year for which the credit is being determined by the fixed-base percentage (113). With certain limited exceptions, the term "gross receipts" means "the total amount, as determined under the taxpayer's method of accounting, derived by the taxpayer from all its activities and from all sources (e.g., revenues derived from the sale of inventory before reduction for cost of goods sold)" (114). The "fixed-base percentage" is the ratio of the taxpayer's aggregate qualified research expenses for the taxable years between December 31, 1983, and January 1, 1989, to the taxpayer's aggregate gross receipts for that same period (115). Special rules for the determination of the fixed base percentage apply with respect to start-up companies, under which the fixed-base percentage is typically three percent (116). Two statutory limitations also apply, however, in the determination of the taxpayer's fixed-base percentage and base amount. In no event can the taxpayer's fixed-base percentage exceed a maximum of 16 percent or the base amount be less than 50 percent of the qualified research expenses for the current taxable year (117).

The purpose behind the determination of the tax credit as an amount in excess of a firm specific percentage of gross receipts is described in the following excerpt from the legislative history to the Omnibus Budget Reconciliation Act of 1989:

> Although the committee believes it is important to readjust the base amount annually in a way that does not undercut the incentive effect of the credit (which occurs when a firm's base is adjusted solely by reference to its own prior levels of research spending), the committee also determined it was appropriate that the base adjustment reflect firm-specific factors. By adjusting each taxpayer's base to its own experience, the committee wanted to make the credit widely available at the lowest possible revenue cost.
>
> Because businesses often determine their research budgets as a fixed percentage of gross receipts, it is appropriate to index each taxpayer's base amount to average growth in its gross receipts. By so adjusting each taxpayer's base amount, the committee believes the credit will be better able to achieve its intended purpose of rewarding taxpayers for research expenses in excess of amounts which would have been expended in any case. Using gross receipts as an index, firms in fast-growing sectors will not be unduly rewarded if their research intensity, as measured by their ratio of qualified research to gross receipts, does not correspondingly increase. Likewise, firms in sectors with slower growth will still be able to earn credits as long as they maintain research expenditures commensurate with their own sales growth.
>
> Adjusting a taxpayer's base by reference to its gross receipts also has the advantage of effectively indexing the credit for inflation and preventing taxpayers from being rewarded for increases in research spending that are attributable solely to inflation (118).

Alternative Incremental Research Tax Credit

Section 41 also provides for an alternative incremental research tax credit that a taxpayer may elect (119). As described above, the incremental research tax credit is equal to 20 percent of the qualified research expenses in excess of a base amount. The base amount is determined

by multiplying the fixed-base percentage by the taxpayer's average annual gross receipts for the four taxable years preceding the taxable year for which the credit is being determined. Because the fixed-base percentage is the ratio of the taxpayer's aggregate qualified research expenses for the taxable years from December 31, 1983, through January 1, 1989, to the taxpayer's aggregate gross receipts for that same period, a taxpayer may not be entitled to the incremental research credit if the growth in the taxpayer's gross receipts has been significantly greater than the growth in its qualified research expenses. For example, assume that a taxpayer had a fixed-base percentage of 10 percent because gross receipts for the taxable years between December 31, 1983, and January 1, 1989, were $10 million and qualified research expenses for that same period were $1 million. If the taxpayer's average annual gross receipts for the four years prior to the taxable year for which the credit is being claimed increased by 20 percent over the average of the taxpayer's annual gross receipts for the 1984 through 1988 taxable years ($2.4 million = 120 percent × [$10 million ÷ 5 years]), but the taxpayer's qualified research expenses for the taxable year increased by only 10 percent as compared to the average annual qualified research expenses over the 1984 through 1988 period ($220,000 = 110 percent × [$1 million ÷ 5 years]), no credit would be available because the taxpayer's qualified research expenses of $220,000 would not exceed the base amount of $240,000 (10 percent of $2.4 million).

To alleviate this problem, the alternative incremental research tax credit dispenses with the fixed-base percentage of the incremental research tax credit and determines the amount of the credit based on the extent to which the qualified research expenses for the taxable year exceed fixed percentages of the taxpayer's average annual gross receipts for the four taxable years preceding the taxable year for which the credit is being determined (the "Section 41(c)(1)(B) amount") (120). The alternative incremental research tax credit is equal to the sum of the following three amount (121):

1. 2.65 percent of the qualified research expenses for the taxable year to the extent that the expenses exceed 1 percent of the Section 41(c)(1)(B) amount but do not exceed 1.5 percent of such amount.
2. 3.2 percent of the qualified research expenses for the taxable year to the extent that the expenses exceed 1.5 percent of the section 41(c)(1)(B) amount but do not exceed 2 percent of such amount.
3. 3.75 percent of the qualified research expenses for the taxable year to the extent that the expenses exceed 2 percent of the section 41(c)(1)(B) amount.

Thus a taxpayer who is not entitled to a credit under the standard incremental research credit may be entitled to relief under the alternative incremental research credit.

Basic Research Tax Credit

Section 41 also permits a basic research tax credit in the amount of 20 percent of any "basic research payment" made during the taxable year in excess of a base amount (122). The basic research credit was enacted in its current form to provide incentives for corporate support of basic scientific research:

> By contrast to other types of research or product development, where expected commercial returns attract private investment, basic research typically does not produce sufficiently immediate commercial applications to make investment in such research self-supporting. Because basic research typically involves greater risks of not achieving a commercially viable result, larger-term projects, and larger capital costs than ordinary product development, the Federal Government traditionally has played a lead role in funding basic research, principally through grants to universities and other nonprofit scientific research organizations. In addition, the research credit as modified by the [Tax Reform Act of 1986] provides increased tax incentives for corporate funding of university basic research to the extent that such expenditures reflect a significant commitment by the taxpayer to basic research (123).

Basic research for purposes of the credit is defined as "any original investigation for the advancement of scientific knowledge not having a specific commercial objective. . ." (124). A basic research payment includes any amount paid in cash by a corporation to a qualified organization for basic research provided (*1*) the basic research is performed by the qualified organization and (*2*) the payment is made pursuant to a written agreement (125). Qualified organizations include colleges and universities, tax-exempt scientific research organizations, and certain tax-exempt organizations operated primarily to promote scientific research by colleges and universities (126).

As previously noted, the credit is equal to 20 percent of the basic research payments in excess of a base amount. The purpose of calculating the tax credit as a percentage of basic research payments in excess of a the base amount is to ensure that the credit is used to encourage increased taxpayer support of basic research and not to encourage taxpayers simply to switch donations from general university giving to forms of support for which the credit is available (127). This base amount is referred to as the "qualified organization base period amount" and is equal to the minimum basic research amount plus the maintenance-of-effort amount (128). In determining the qualified organization base period amount, the minimum basic research amount is an amount equal to the greater of (*1*) 1 percent of the average amount of any in-house and contract research expenses paid or incurred over the base period or (*2*) the amount of basic research payments treated as contract research expenses under Section 41(e)(1)(B) during the base period (129). For calendar-year taxpayers, the base period is the three-year period from 1981 to 1983 (130). For taxpayers not in existence during the base period, the minimum basic research amount is not to be less than 50 percent of the basic research payments for the taxable year (131).

The maintenance-of-effort amount is equal to the average of the nondesignated university contributions paid by the taxpayer during the base period in excess of the nondesignated university contributions paid by the taxpayer during the taxable year (132). Nondesignated university contributions are equal to any amount paid by

the taxpayer to a qualified organization as defined under Section 41 for which a charitable contribution deduction was allowable under Section 170 and which was not taken into account in determining the basic research credit or as a basic research payment (133). Consequently any reduction in the amount of charitable contributions to qualified organizations from the average amount of contributions made during the base period will offset basic research payments eligible for the basic research credit.

LIMITATIONS ON SECTION 174 DEDUCTIONS

Because qualified research expenses and basic research payments under Section 41 may also be deductible as research or experimental expenditures under Section 174, the Code requires that Section 174 deductions be reduced by the amount of any credit taken under Section 41 (134). The legislative history provides the following example of this requirement:

> For example, assume that a taxpayer makes credit-eligible research expenditures of $1 million during the year, and that the base period amount is $600,000. The taxpayer is allowed a tax credit equal to 20 percent of the $400,000 increase in research expenditures, or $80,000. . . .Under the provision, the taxpayer's deduction is reduced by the $80,000 credit, leaving a deduction of $920,000 (135).

In addition, if research and experimental expenditures are capitalized rather than currently deducted, the capitalized amount must be similarly reduced by the amount of any research credit available under Section 41 (136).

ORPHAN DRUG CREDIT UNDER SECTION 45C

Section 45C of the Code creates the so-called orphan drug credit, which permits a credit equal to 50 percent of the qualified clinical testing expenses paid or incurred for the taxable year (137). The orphan drug credit is so named because it permits a tax credit for certain expenses incurred in the development of drugs used to treat those rare diseases or conditions that affect fewer than 200,000 persons in the United States or that affect more than 200,000 persons but for which the developer of such a drug would have no reasonable expectation of recovering the cost of developing or marketing the drug from its sales in the United States (138). Such diseases and conditions include Huntington's disease, myoclonus, amyotrophic lateral sclerosis (ALS or "Lou Gehrig's disease"), Tourette's syndrome, and Duchenne's dystrophy, a form of muscular dystrophy (139).

Qualified clinical testing expenses are those amounts that would satisfy the definition of qualified research expenses under Section 41(b), with certain modifications (140). One such modification permits 100 percent of any contract research expenses to fall within the definition of qualified clinical testing expenses rather than only 65 percent of such expenses that fall within the definition of qualified research expenses (141). Nevertheless, qualified clinical research expenses do not include any amount otherwise funded under any grant or contract by another person or governmental entity (142). Under the regulations, if the taxpayer conducting the clinical testing for another person retains no substantial rights in the testing, the taxpayer's testing expenses are treated as fully funded (143). Incidental benefits such as increased experience in the field of human clinical testing do not constitute substantial rights in the clinical testing. When the taxpayer conducting the clinical testing retains substantial rights in the testing, the testing expenses are reduced to the extent of any payments and the fair market value of any property to which the taxpayer becomes entitled by conducting the clinical testing (144).

Importantly qualified clinical testing expenses are limited to human clinical testing (145). Human clinical testing requires the use of human subjects to determine the effect of the designated drug on humans necessary to receive approval under Section 505(b) of the Federal Food, Drug, and Cosmetic Act or be licensed under Section 351 of the Public Health Services Act (146). A human subject is an individual who is a participant in research, either as a recipient of the drug or as a control, and may be either a healthy individual or a patient (147). The clinical testing must also occur within the United States unless an insufficient testing population exists within the United States and the testing is performed by a U.S. person or any other person unrelated to the taxpayer (148).

Because qualified clinical testing expenses will also constitute qualified research expenses under Section 41, such expenses are not taken into account for purposes of the research credit under Section 41 if the taxpayer elects the orphan drug credit for the taxable year (149). Nevertheless, such qualified clinical testing expenses are taken into account in determining base period research expenses for purposes of applying Section 41 in any subsequent taxable year (150).

BIBLIOGRAPHY

1. IRC §174(a); Reg. §§174-1 and 1.174-3.
2. IRC §174(b); Reg. §§1.174-1 and 1.174-4.
3. Reg. §1.174-1.
4. Rev. Rul. 58-356, 1958-2 CB 104; Rev. Rul. 83-138, 1983-2 CB 50.
5. IRC §§56(b)(2)(A)(ii) and 56(b)(2)(D); Priv. Ltr. Rul. 9746002 (a Technical Advice Memorandum).
6. 416 U.S. 500 (1974).
7. *Snow v. Commissioner*, 416 U.S. 500, 503–504 (1974).
8. IRC §195.
9. 83 TC 667 (1984).
10. Id. at 686–687 (emphasis in original). Compare *Scoggins v. Commissioner*, 46 F3d 950 (9th Cir. 1995) (taxpayer's objective intent and capability to enter into a business in connection with its research activities demonstrated a "realistic prospect" of subsequently entering into business such that research expenditures were deductible); *Best Universal Lock Co.*, 45 TC 1 (1965), acq. (allowing deductions under §174 for research and experimentation expenses incurred in developing an isothermal air compressor because the expenditures were incurred in connection with taxpayer's business, even though the new product was unrelated to taxpayer's past line of products);

M. Bush, TC Memo 1994-523, 68 TCM 974 (deductions in connection with research and development expenditures permitted despite the taxpayer's failure to sell products during the taxable year at issue); *O.B. Kilroy*, TC Memo 1980-489, 41 TCM 292 (taxpayer engaged in trade or business of exploiting inventions could deduct research or experimentation expenditures under §174 despite fact that gross receipts from the activity were negligible); and Rev. Rul. 71-162, 1971-1 CB 97 (research or experimental expenditures incurred in developing products unrelated to taxpayer's current product line or manufacturing processes may be deductible under §174); with *Mach-Tech, Ltd. Partnership & Serv-Tech, Inc.*, 95-2 USTC ¶50,375 (5th Cir. 1995) (disallowing deductions under §174 because the taxpayer was neither engaged in, nor had any realistic prospect of, entering into a trade or business); *Mayrath v. Commissioner*, 357 F2d 209 (5th Cir. 1966) (disallowing deductions under §174 for costs incurred in developing new techniques for housing construction because no indication of profit motive, as taxpayer was a professional inventor of farm machine products, not construction products); *W.J. Piszczek*, TC Memo 1998-307, 76 TCM 338 (applying the regulatory factors under Regulation §1.183-2(b) to conclude that the taxpayer was not engaged in activities to produce a wind-powered ethanol distillery with the profit motive necessary to satisfy the trade or business requirement of §174); *P.E. Sheehy*, TC Memo 1998-183, 75 TCM 2309 (taxpayer provided no evidence that partnership was actively involved in a trade or business involving the development or manufacture of recyclable plastic containers); *Utah Jojoba I Research*, TC Memo 1998-6, 75 TCM 1524 ("For an investing partnership successfully to claim research and experimental deductions, there must be a realistic prospect that the technology to be developed will be exploited in a trade or business of the partnership claiming deductions under section 174 ... Mere legal entitlement to enter into a trade or business does not satisfy this test."); *H.I. Shaller*, TC Memo 1984-584, 49 TCM 10 (research relating to ocean surf energy did not rise to the level of a trade or business); *Gyro Eng'g Corp.*, TC Memo 1974-288, 33 TCM 1343 (taxpayer not engaged in trade or business of research or inventing). For decisions prior to *Snow* considering the existence of a trade or business, see *O.B. Kilroy*, TC Memo 1973-7, 32 TCM 27; *J.H. Cunningham*, TC Memo 1968-242, 27 TCM 1219; *Stanton v. Commissioner*, TC Memo 1967-137, 26 TCM 618, aff'd, 399 F2d 326 (5th Cir. 1968); *C.H. Schafer*, TC Memo 1964-156, 23 TCM 927; *E.G. Bailey*, TC Memo 1963-251, 22 TCM 1255.

11. *Harris v. Commissioner*, 16 F3d 75 (5th Cir. 1994); *P.D. Martyr*, TC Memo 1990-558, 60 TCM 1115, aff'd sub nom. *Gatto v. Commissioner*, 1 F3d 826 (9th Cir. 1993); *Kantor v. Commissioner*, 998 F2d 1514 (9th Cir. 1993); *United Fibertech, Ltd. v. Commissioner*, 976 F2d 445 (8th Cir. 1992); *Nickeson v. Commissioner*, 962 F2d 973 (10th Cir. 1992); *Diamond v. Commissioner*, 930 F2d 372 (4th Cir. 1991); *Zink v. United States*, 929 F2d 1015 (5th Cir. 1991); *R.C. Jay*, TC Memo 1988-232, 55 TCM 933, aff'd sub nom. *Ben-Porat v. Commissioner*, 908 F2d 976 (9th Cir. 1990); *Property Growth Co. v. Commissioner*, 89-2 USTC ¶9479 (8th Cir. 1989); *Spellman v. Commissioner*, 845 F2d 148 (7th Cir. 1988); *Levin v. Commissioner*, 832 F2d 403 (7th Cir. 1987); *Independent Elec. Supply, Inc. v. Commissioner*, 781 F2d 724 (9th Cir. 1986); *S. Drobny*, 86 TC 1326 (1986); *H.J. Green*, 83 TC 667 (1984); *Utah Jojoba I Research*, TC Memo 1998-6, 75 TCM 1524; *Cactus Wren Jojoba, Ltd.*, TC Memo 1997-504, 74 TCM 1133; *3-Koam Co.*, TC Memo 1997-148, 73 TCM 2415; *S.H. Glassley*, TC Memo 1996-206, 71 TCM 2898; *Digital Accounting Technology, Ltd.*, TC Memo 1995-339, 70 TCM 178; *LDL Research & Dev. II, Ltd.*, TC Memo 1995-172, 69 TCM 2411, aff'd, 124 F3d 1338 (10th Cir. 1997); *Estate of G.B. Cook*, TC Memo 1993-581, 66 TCM 1523; *P.A. Stankevich, Jr.*, TC Memo 1992-458, 64 TCM 460; *Software 16*, TC Memo 1992-247, 63 TCM 2876; *E. Stauber*, TC Memo 1992-128, 63 TCM 2258; *Double Bar Chain Co., Ltd.*, TC Memo 1991-572, 62 TCM 1276; *Scientific Measurement Sys. I, Ltd.*, TC Memo 1991-69, 61 TCM 1951; *J.P. Coleman*, TC Memo 1990-357, 60 TCM 123, upheld on reh'g, TC Memo 1990-511, 60 TCM 889; *C.F. Alexander*, TC Memo 1990-141, 59 TCM 121, aff'd without pub. op. sub nom. *Stell v. Commissioner*, 999 F2d 544 (9th Cir. 1993); *Medical Mobility Ltd. Partnership I*, TC Memo 1993-428, 66 TCM 741; *Active Lipid Dev. Partners, Ltd.*, TC Memo 1991-522, 62 TCM 1046; *N.F. Ben-Avi*, TC Memo 1988-74, 55 TCM 199; *R. Rosenberg*, TC Memo 1987-441, 54 TCM 392; *T.S. Reinke*, TC Memo 1981-120, 41 TCM 1100; Priv. Ltr. Rul. 9604004 (a Technical Advice Memorandum); Field Service Advice 1999-839 (undated), available in LEXIS, 1999 TNT 70-12. But see *Scoggins v. Commissioner*, 46 F3d 950 (9th Cir. 1995); *Smith v. Commissioner*, 937 F2d 1089 (6th Cir. 1991).
12. Several courts have held that this issue involves the application of law to fact and, thus, call for de novo review on appeal. *Scoggins v. Commissioner*, 46 F3d 950 (9th Cir. 1995); *Nickeson v. Commissioner*, 962 F2d 973 (10th Cir. 1992); *Zink v. United States*, 929 F2d 1015 (5th Cir. 1991).
13. *E. Stauber*, TC Memo 1992-128, 63 TCM 2258. In addition, taxpayers who merely invest in an entity engaged in research and development activities are precluded from deducting their investments as research or experimental expenditures. *Safstrom v. Commissioner*, 95-1 USTC ¶50,030 (9th Cir. 1994); *Cleveland v. Commissioner*, 297 F2d 169 (4th Cir. 1961).
14. 998 F2d 1514 (9th Cir. 1993).
15. *Kantor v. Commissioner*, 998 F2d 1514, 1519 (9th Cir. 1993). See also *Diamond v. Commissioner*, 930 F2d 372 (4th Cir. 1991) (based on a review of the financial arrangements between the parties, the court concluded that "if a money-making business should materialize, there exists no reasonable expectation that [the research entity] will permit it to be exploited by one of the partnerships"); *Spellman v. Commissioner*, 845 F2d 148 (7th Cir. 1988) (concluding that an option price of only $20,000 gave the profit potential of any trade or business that might result to the research entity).
16. *Kantor v. Commissioner*, 998 F2d 1514, 1520 (9th Cir. 1993). See also *Levin v. Commissioner*, 832 F2d 403 (7th Cir. 1987) (concluding that the partnership was not engaged in a trade or business where the general partner visited a food machinery plant for the first time when escorted to one after the partnerships were formed).
17. 46 F3d 950 (9th Cir. 1995).
18. Reg. §1.174-2(a)(1).
19. Reg. §1.174-2(a)(1).
20. Reg. §1.174-2(a)(2).
21. Reg. §1.174-2(a)(1).
22. GCM 39527.
23. Reg. §1.174-2(a)(1).
24. See Reg. §§1.174-2(a)(1) and 1.174-2(a)(2).
25. Reg. §1.174-2(a)(3)(vii). See also Rev. Rul. 71-363, 1971-2 CB 156; *Hakim v. Commissioner*, TC Memo 1974-46, 33 TCM 223, aff'd, 512 F2d 1379 (6th Cir. 1975).
26. 1973-2 CB 87.

27. See also *H.F. Crouch*, TC Memo 1990-309, 59 TCM 938; *A. Quinn*, TC Memo 1974-64, 33 TCM 310; Priv. Ltr. Rul. 7004169560A (a Technical Advice Memorandum); GCM 39527.
28. *Yellow Freight Sys., Inc. of Del. v. United States*, 92-1 USTC ¶50,029 (Cl. Ct. 1991).
29. Reg. §§1.174-2(a)(3) and 1.174-2(a)(4); *Utah Jojoba I Research*, TC Memo 1998-6, 75 TCM 1524; *Cactus Wren Jojoba, Ltd.*, TC Memo 1997-504, 74 TCM 1133.
30. Reg. §1.174-2(a)(3)(vi).
31. Reg. §1.174-2(b)(4).
32. Reg. §1.174-2(b)(1); *Ekman v. Commissioner*, TC Memo 1997-318, 74 TCM 72, aff'd, 99-1 USTC ¶50,580 (6th Cir. 1999); *P.F. Sheehy*, TC Memo 1996-334, 72 TCM 178.
33. Reg. §1.174-2(b)(1).
34. Reg. §§1.174-2(a)(8) and 1.174-2(a)(9) examples (1) and (2). See also Rev. Rul. 73-324, 1973-2 CB 72; Rev. Rul. 73-20, 1973-1 CB 133; Rev. Rul. 69-484, 1969-2 CB 38.
35. Reg. §1.174-2(b)(3). See Priv. Ltr. Rul. 8614004. Despite satisfaction of the requirements of Regulation §1.174-2(b)(3), deductible research or experimental expenditures do not include "the costs of the component materials of the depreciable property, the costs of labor or other elements involved in its construction and installation, or costs attributable to the acquisition or improvement of the property." Reg. §1.174-2(b)(4).
36. Reg. §1.174-2(b)(3).
37. IRC §174(a)(3); Reg. §1.174-3(a).
38. Reg. §1.174-3(a).
39. Rev. Rul. 68-144, 1968-1 CB 85.
40. Reg. §1.174-4(a)(2).
41. Reg. §1.174-4(a)(3).
42. IRC §174(b)(1)(C); Reg. §§1.174-4(a)(3) and 1.174-4(a)(4).
43. Reg. §1.174-4(a)(4).
44. Reg. §1.174-4(a)(4).
45. IRC §41(a).
46. Staff of the Joint Comm. on Tax'n, General Explanation of the Economic Recovery Tax Act of 1981, 120 (1981).
47. IRC §41(b). The credit is not available to taxpayers who merely fund research activities, typically of their wholly owned corporations, and do not acquire rights to the product of that research. See *Safstrom v. Commissioner*, TC Memo 1992-587, 64 TCM 971, aff'd, 95-1 USTC ¶50,030 (9th Cir. 1994).
48. See Reg. §1.41-2(a)(1); *N.F. Ben-Avi*, TC Memo 1988-74, 55 TCM 199; *E.H. Allen*, TC Memo 1988-166, 55 TCM 641.
49. Staff of the Joint Comm. on Tax'n, General Explanation of the Economic Recovery Tax Act of 1981, 122 (1981); H.R. Rep. No. 201, 97th Cong., 1st Sess. 112 (1981).
50. IRC §41(b)(4). See also Reg. §1.41-2(a)(4) (extending the availability of the research tax credit to certain partnerships and joint ventures not carrying on a trade or business to which the research relates).
51. IRC §41(f)(1)(A); Reg. §1.41-8.
52. IRC §§41(f)(1)(B) and 41(g); Reg. §§1.41-8 and 1.41-9.
53. IRC §41(a)(1).
54. IRC §§41(d)(1) and 41(d)(3); Prop. Reg. §1.41-4(a)(2).
55. H.R. Rep. No. 426, 99th Cong., 1st Sess. 180 (1985). See also S. Rep. No. 313, 99th Cong., 2d Sess. 696 (1986); Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 133 (1987).
56. Prop. Reg. §1.41-4(a)(7); H.R. Rep. No. 841, 99th Cong., 2d Sess. II-71, n.3 (1986); Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 133, n.23 (1987).
57. Prop. Reg. §1.41-4(a)(3). For illustrations of situations in which this requirement is satisfied, see Prop. Reg. §1.41-4(a)(8) examples (1), (5), (6), and (7); for illustrations of situations in which this requirement is not satisfied, see Prop. Reg. §1.41-4(a)(8) examples (2), (3), and (4).
58. Prop. Reg. §1.41-4(a)(8) example (6).
59. Prop. Reg. §1.41-4(a)(8) example (5).
60. Prop. Reg. §1.41-4(a)(8) example (8).
61. Prop. Reg. §1.41-4(a)(4).
62. Staff of the Joint Committee on Taxation, General Explanation of the Tax Reform Act of 1986, 133 (1987). See also H.R. Rep. No. 426, 99th Cong., 1st Sess. 180-181 (1985); S. Rep. No. 313, 99th Cong., 2d Sess. 696 (1986).
63. H.R. Conf. Rep. No. 825, 105th Cong., 2d Sess. 1548-1549 (1998) (describing the provisions of the Tax and Trade Relief Extension Act of 1998, Pub. L. No. 105-277, §1001, 112 Stat. 2681, 2681-888 (1998)); Staff of the Joint Committee on Taxation, General Explanation of Tax Legislation Enacted in 1998, 236 (1998).
64. Prop. Reg. §1.41-4(a)(5).
65. See Reg. §1.174-2(a)(1).
66. Prop. Reg. §1.41-4(a)(8) example (4).
67. Prop. Reg. §1.41-4(a)(5)(ii). See also Prop. Reg. §§1.41-4(a)(5)(iii) and 1.41-4(d).
68. IRC §41(d)(3)(A).
69. Prop. Reg. §1.41-4(a)(6).
70. IRC §41(d)(2); Prop. Reg. §1.41-4(b)(1).
71. Prop. Reg. §1.41-4(b)(1).
72. Prop. Reg. §§1.41-4(c)(2)(iii) and 1.41-4(c)(10) example (3).
73. Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 134 (1987).
74. Prop. Reg. §1.41-4(b)(2); Prop. Reg. §1.41-4(b)(3).
75. 110 TC 454 (1998).
76. Norwest Corp., 110 TC 454, 493 (1998).
77. Norwest Corp., 110 TC 454, 494 (1998) (referencing H.R. Rep. No. 841, 99th Cong., 2d Sess. II-71, n.3 (1986)).
78. Norwest Corp., 110 TC 454, 493 (1998). See also *United Stationers, Inc. v. United States*, 163 F3d 440 (7th Cir. 1998).
79. IRC §41(d)(3)(B). See also Priv. Ltr. Rul. 9522001 (a Technical Advice Memorandum).
80. IRC §41(d)(4)(A); Prop. Reg. §1.41-4(c)(2).
81. Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 136 (1987).
82. Prop. Reg. §§1.41-4(c)(2)(ii) and 1.41-4(c)(10) example (3). See Norwest Corp., 110 TC 454 (1998).
83. Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 136 (1987). See Prop. Reg. §§1.41-4(c)(10) examples (1) and (2).
84. IRC §41(d)(4)(C); Prop. Reg. §1.41-4(c)(4). See Prop. Reg. §1.41-4(c)(10) example (5).
85. Prop. Reg. §1.41-4(c)(4); Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 137 (1987); Market Segment Specialization Program, "Manufacturing Industry" (May 1, 1998), available in LEXIS, 1999 TNT 81-23.
86. See also Field Service Advice 1999-1023 (dated Oct. 22, 1993), available in LEXIS, 1999 TNT 81-49; Market Segment

Specialization Program, "Manufacturing Industry" (May 1, 1998), available in LEXIS, 1999 TNT 81-23.

87. IRC §41(d)(4)(H); Reg. §1.41-5(d); Prop. Reg. §1.41-4(c)(9).
88. Reg. §1.41-5(d)(1). See also Reg. §1.41-2(e)(2).
89. Reg. §1.41-5(d)(2).
90. But see *Fairchild Indus., Inc. v. United States*, 94-1 USTC ¶50,164 (Fed. Cl. 1994), rev'd, 71 F3d 868 (Fed. Cir. 1995) (permitting a research credit in connection with expenses incurred under a government defense contract where the taxpayer bore the economic risk of loss; "[t]he inquiry turns on who bears the research costs upon failure, not on whether the researcher is likely to succeed in performing the project").
91. 98-2 USTC ¶50,887 (Fed. Cl. 1998).
92. IRC §41(d)(4)(D); Prop. Reg. §1.41-4(c)(5). See also Prop. Reg. §1.41-4(c)(10) example (6).
93. Prop. Reg. §1.41-4(c)(5); Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 136-137 (1987). See also Norwest Corp., 110 TC 454 (1998).
94. IRC §41(d)(4)(B); Prop. Reg. §1.41-4(c)(3). See also Prop. Reg. §1.41-4(c)(10) example (4); Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 136 (1987).
95. IRC §41(d)(4)(E); Prop. Reg. §§1.41-4(c)(6) and 1.41-4(e).
96. IRC §41(d)(4)(F); Prop. Reg. §1.41-4(c)(7).
97. IRC §41(d)(4)(G); Prop. Reg. §1.41-4(c)(8).
98. IRC §41(a).
99. IRC §41(b)(1).
100. IRC §41(b)(2)(A)(i); Reg. §1.41-2(d). See also Market Segment Specialization Program, "Manufacturing Industry" (May 1, 1998), available in LEXIS, 1999 TNT 81-23; Market Segment Specialization Program, "Computers, Electronics, and High Tech Industry" (March 15, 1997), available in LEXIS, 98 TNT 199-12; *Apple Computer, Inc.*, 98 TC 232 (1992), acq. in part; *Sun Microsystems, Inc.*, TC Memo 1995-69, 69 TCM 1884, acq.; *E.V. Fudim*, TC Memo 1994-235, 67 TCM 3011; Priv. Ltr. Rul. 8835002 (a Technical Advice Memorandum). Priv. Ltr. Rul. 9018003 (a Technical Advice Memorandum); Internal Revenue Service, Industry Specialization Program Coordinated Issue Paper (Data Processing Industry) Qualifying Wages under Section 41 in Determining Tax Credit for Increasing Research Activities, available in LEXIS, 96 TNT 199-12; Internal Revenue Service, Industry Specialization Program Coordinated Issue Paper (All Industries) Qualifying Wages Under Section 41 in Determining Tax Credit for Increasing Research Activities (February 16, 1999), available in LEXIS, 1999 TNT 74-86.
101. IRC §41(b)(2)(A)(ii); Reg. §1.41-2(b). See Market Segment Specialization Program, "Manufacturing Industry" (May 1, 1998), available in LEXIS, 1999 TNT 81-23.
102. IRC §41(b)(2)(A)(iii).
103. IRC §41(b)(3); Reg. §1.41-2(e).
104. IRC §41(b)(3)(C)(i). A qualified research consortium is a tax-exempt organization organized and operated primarily to conduct scientific research. IRC §41(b)(3)(C)(ii).
105. IRC §41(b)(3)(B); Reg. §1.41-2(e)(4).
106. Reg. §1.41-2(e)(2).
107. Reg. §1.41-2(e)(3).
108. 110 TC 454 (1998).
109. Norwest Corp., 110 TC 454, 519 (1998).
110. Reg. §1.41-2(a)(1).
111. Staff of the Joint Comm. on Tax'n, General Explanation of the Economic Recovery Tax Act of 1981, 122–123 (1981); H.R. Rep. No. 201, 97th Cong., 1st Sess. 113 (1981); S. Rep. No. 144, 97th Cong., 1st Sess. 78 (1981).
112. IRC §41(a). For examples of the calculation of the research tax credit, see Market Segment Specialization Program, "Manufacturing Industry" (May 1, 1998), available in LEXIS, 1999 TNT 81-23, and Market Segment Specialization Program, "Computers, Electronics, and High Tech Industry" (March 15, 1997), available in LEXIS, 98 TNT 199-12.
113. IRC §41(c)(1).
114. Prop. Reg. §§1.41-3(c)(1) and 1.41-3(c)(2).
115. IRC §41(c)(3)(A).
116. IRC §41(c)(3)(B).
117. IRC §§41(c)(2) and 41(c)(3)(C).
118. H.R. Rep. No. 247, 101st Cong., 1st Sess. 1199–1200 (1989).
119. IRC §41(c)(4).
120. IRC §41(c)(4)(A).
121. IRC §41(c)(4).
122. IRC §§41(a)(2) and 41(e)(1)(A).
123. Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 131 (1986).
124. IRC §41(e)(7)(A).
125. IRC §41(e).
126. IRC §41(e)(6).
127. H.R. Rep. No. 426, 99th Cong., 1st Sess. 178 (1985); S. Rep. No. 313, 99th Cong., 2d Sess. 695 (1986); Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 131 (1987).
128. IRC §41(e)(3).
129. IRC §41(e)(4)(A).
130. IRC §41(e)(7)(B).
131. IRC §41(e)(4)(B).
132. IRC §41(e)(5)(A).
133. IRC §41(e)(5)(B).
134. IRC §280C(c)(1). See H.R. Rep. No. 795, 100 Cong., 2d Sess. 452 (1988).
135. H.R. Rep. No. 795, 100 Cong., 2d Sess. 453 (1988).
136. IRC §280(c)(2).
137. IRC §45C(a); Reg. §1.28-1(a).
138. IRC §45C(d)(1) (referring to §526 of the Federal Food, Drug, and Cosmetic Act, Pub. L. No. 97-414, §2(a), 96 Stat. 2050 (1983), as amended, codified at 21 USC §360(bb)); Reg. §§1.28-1(c) and 1.28-1(d)(1)(i).
139. Reg. §1.28-1(d)(1)(i). See H.R. Rep. No. 426, 99th Cong., 1st Sess. 230 (1985); Staff of the Joint Comm. on Tax'n, General Explanation of the Tax Reform Act of 1986, 141 (1987).
140. IRC §45C(b)(1); Reg. §1.28-1(b)(1).
141. IRC §45C(b)(1)(B)(ii); Reg. §1.28-1(b)(2).
142. IRC §45C(b)(1)(C); Reg. §1.28-1(b)(3)(i).
143. Reg. §1.28-1(b)(3)(ii).
144. Reg. §1.28-1(b)(3)(iii)(A).
145. IRC §45C(b)(2)(A); Reg. §1.28-1(c)(2).
146. Reg. §1.28-1(c)(2).
147. Reg. §1.28-1(c)(2).
148. IRC §45C(d)(2); Reg. §1.28-1(d)(3).
149. IRC §45C(c)(1).
150. IRC §45C(c)(2).

ADDITIONAL READINGS

Generally

C.E. Falk, Tax Planning for the Development and Licensing of Patents and Know-How, 557 Tax Mgmt. (BNA) (1996).

J.E. Maule, Tax Credits: Concepts and Calculation, 506 Tax Mgmt. (BNA) (1994).

J.E. Maule and L.M. Starczewski, Deductions: Overview and Conceptual Aspects, 503 Tax Mgmt. (BNA) (1995).

J.E. Maule and L.M. Starczewski, Deduction Limitations: General, 504 Tax Mgmt. (BNA) (1995).

P.W. Oosterhuis and J.S. Stanton, Research and Development Expenditures, 42-3rd Tax Mgmt. (BNA) (1987).

P.F. Postlewaite, D.L. Cameron and T. Kittle-Kamp, Federal Income Taxation of Intellectual Properties and Intangible Assets, Warren, Gorham, and Lamont, 1997.

Research and Development Expenditures Under Section 174

J. Bankman, The Structure of Silicon Valley Start-Ups, 41 UCLA L. Rev. 1737 (1994).

J.E. Bischel, Deduction and Allocation of Research and Development Expenditures; The Final Chapter or Just Another Installment? 16 Int 1 Tax J. 225 (1990).

G. Foreman, Research and Experimental Expenditures Sec. 174 and Credit for Increasing Research Activities Sec. 41, 60 CPA J 69 (1990).

D.S. Goldberg, Recent Approaches to the Trade or Business Requirement of Section 174: Unauthorized Snow Removal, 8 Va. Tax Rev. 861 (1989).

D.S. Hudson, The Tax Concept of Research or Experimentation, 45 Tax Law. 85 (1991).

F. Murias, New R&E Regs. Are More Rational Than Prior Rules, but Key Terms Need Clarification, 80 J. Taxation 44 (1994).

B.M. Seltzer, D.A. Golden, and M.E. Monahan, Maximizing Opportunities Under the New Research and Experimentation Regulations, 47 Tax Exec. 102 (1995);.

S.P. Starr and D.S. Shapiro, Current Tax Issues in Research and Experimentation, 52 Inst. on Fed. Taxation 5 (1994).

H.W. Wolosky, IRS Eases Position on R&D, 26 Prac. Acct. 43 (June 1993).

The Research Tax Credit Under Section 41

B.A. Billings and G.A. McGill, The Effect of Base Changes on the Incremental Research and Experimentation Tax Credit, 54 Tax Notes 1155 (1992).

B.A. Billings and J.D. Schroeder, Research and Experimentation as an Investment Alternative: The Effect of the Incremental Tax Credit, 61 CPA J. 60 (1991).

D.L. Champi, Tax Policy and National Security Strategy After Operation Desert Storm: The Role of the Research and Experimentation Credit in Enhancing Expenditures on Basic Research and Technology Development Made by the United States Defense Technology Base, 45 Tax Law. 195 (1991).

Coopers and Lybrand Tax Policy Economics Group, Economic Benefits of the R&D Tax Credit, 78 Tax Notes 1019 (1998).

N.J. Crimm, A Tax Proposal to Promote Pharmacologic Research, to Encourage Conventional Prescription Drug Innovation and Improvement, and To Reduce Product Liability Claims, 29 Wake Forest L. Rev. 1007 (1994).

D.S. Hudson, "The Tax Concept of Research or Experimentation," 45 Tax Law. 85 (1991).

K.J. Lester, Note, Availability of the Research Tax Credit for Government Funded Research: *Fairchild Industries v. United States*, 50 Tax Law. 873 (1997).

National Association of Manufacturers, The R&D Tax Credit: Lasting Gains in Research and GDP, available in LEXIS, 98 TNT 167–22, (April 13, 1998).

P.W. Oosterhuis, International R&D and Technology Transfer Arrangements, 73 Taxes 905 (1995).

B.J.W. Raby and W.L. Raby, Seven Tests for Software Development Research Projects, 80 Tax Notes 583 (1998).

Staff of the Joint Committee on Taxation, General Explanation of the Tax Reform Act of 1986 (1987).

Staff of the Joint Committee on Taxation, General Explanation of Tax Legislation Enacted in 1998 (1998).

S. Stratton and B. Massey, Major Changes to Research Credit Rules Sought at IRS Reg Hearing, 83 Tax Notes 623 (1999).

P.A. Stoffregen, Giving Credit Where Credit Is Due: A Brief History of the Administration of the R&D Credit, 66 Tax Notes 403 (1995).

J.S. Wong, *Fairchild Industries, Inc.:* Was the Court Fair in Denying Research Credit? 64 Tax Notes 1477 (1994).

J.S. Wong, Research Credit Allowed for Some Internal-Use Computer Software, 80 Tax Notes 851 (1998).

H. Watson, The 1990 R&D Tax Credit: A Uniform Tax on Inputs and a Subsidy for R&D, 49 Natl Tax J. 93 (1993).

See other entries FEDERAL REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE, FDA, ORPHAN DRUG ACT; STRATEGIC ALLIANCES AND TECHNOLOGY LICENSING IN BIOTECHNOLOGY.

STRATEGIC ALLIANCES AND TECHNOLOGY LICENSING IN BIOTECHNOLOGY

JOSH LERNER
Harvard University and National Bureau of Economic Research
Boston, Massachusetts

OUTLINE

INTRODUCTION

This article explores the roles of strategic alliances and licensing in the biotechnology industry. Two areas that have attracted considerable attention from academics and practitioners are highlighted. First, the article considers the reasons for the prevalence of alliances and licensing agreements in the biotechnology industry. The hypotheses and evidence about the structuring of

these agreements are then considered. Corroboratory evidence from recent field research is then summarized. The final section highlights additional issues for future research.

At the outset, it is worth mentioning the complexity of these agreements, which reflects the costly and uncertain nature of biotechnology projects. The complexity and unpredictability of the research presents challenges in drafting enforceable agreements that specify the contributions of each party in the face of all contingencies. A great deal of innovation has consequently been devoted to the design of these contracts, which makes it difficult to generalize about this phenomenon.

Despite these difficulties, the understanding of strategic alliances is critical to those who wish to understand the biotechnology industry, or high technology industries more generally. The availability of equity from public investors for new high technology firms has been variable, with biotechnology a particularly extreme case. The financing activities of biotechnology firms between 1978 and 1995 are summarized in Figure 1 and Table 1 (1,2). During periods with little financing activity, young high technology firms suffer tremendous financial stresses and have few alternatives to raise capital other than strategic alliances.

Furthermore, the economic importance of technology alliances has been increasing. Panel A of Table 2 shows the number of such alliances has been growing in a variety of industries. While obtaining a comprehensive view of alliance financing is exceedingly difficult, tabulations suggest that alliances are the dominant source of external financing for R&D by young firms in many industries, including advanced materials, information technology, and telecommunications (3). Surveys of corporate research managers suggest that alliances will be an increasingly important mechanism through which R&D is financed in the years to come (4).

Nowhere is this trend clearer than in biotechnology, where alliances with pharmaceutical firms have become in recent years the single largest source of financing for biotechnology firms, accounting for several billion dollars of funds annually. Panel B of Table 2 illustrates the growth in the number of alliances involving U.S. biotechnology firms and other private-sector entities. The economic importance of these transactions is also shown by the willingness of firms to spend substantial amounts litigating them and the size of the damage awards: for example, Genentech and Eli Lilly's dispute over their alliance to develop human growth hormone, which led to the filing of at least six suits between 1987 and 1993. One indication of the importance of these agreements is the

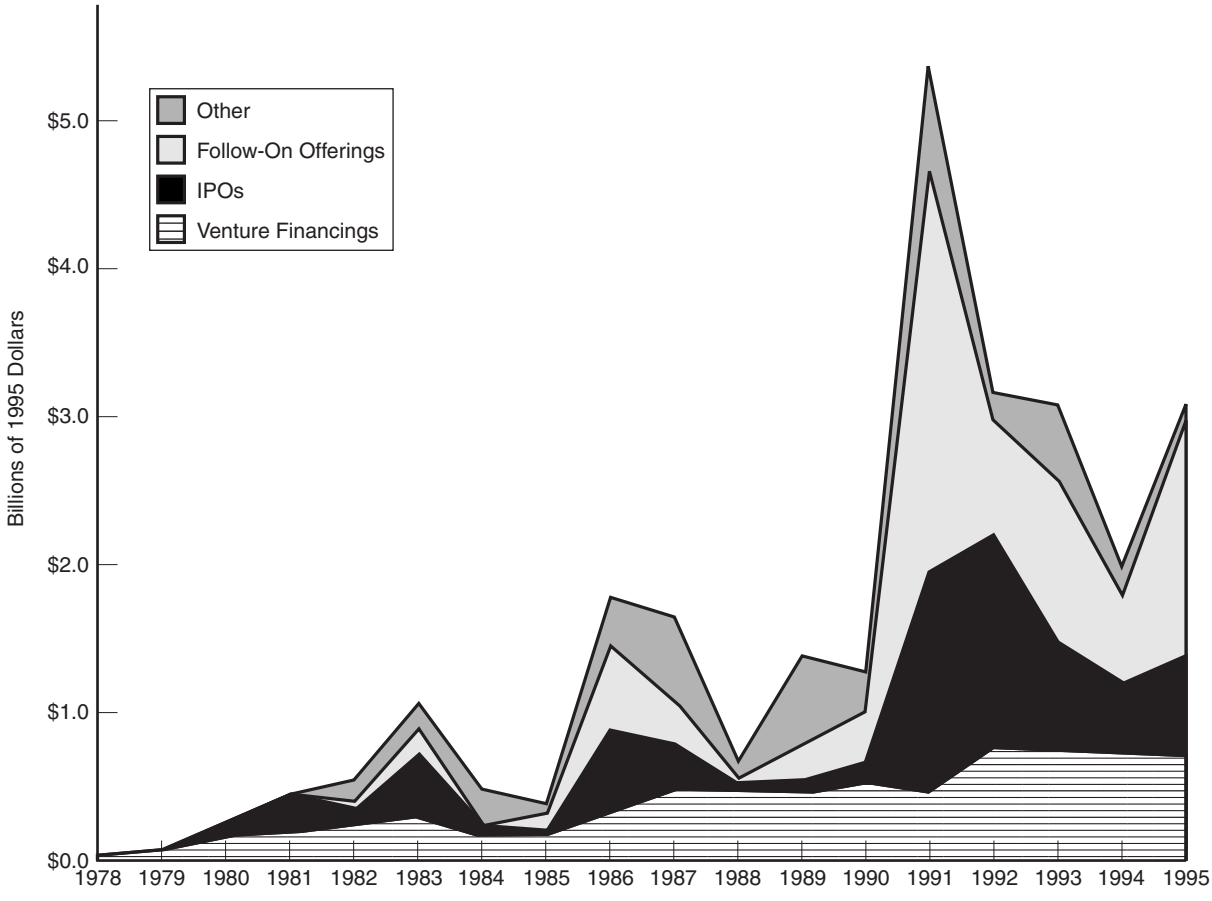

Figure 1. External financing of the U.S. biotechnology industry. The chart depicts the amount raised by U.S. new biotechnology firms through private venture financings, initial public offerings, follow-on public equity offerings, and other sources. (Alliance-related financings are excluded.)

Table 1. External Financing of the U.S. Biotechnology Industry

Year	Amount (millions of 1995 dollars) raised through: Venture capital	IPOs	Follow-on offerings	Private placements	Debt and convertibles	RDFOs	Total	Biotech equity index
1978	23	0	0	0	0	0	23	1.16
1979	71	0	0	0	0	0	71	1.39
1980	161	87	0	0	0	0	248	2.26
1981	185	263	0	13	0	0	461	1.99
1982	247	98	57	0	0	147	549	2.29
1983	292	423	182	0	0	172	1070	1.87
1984	179	55	0	0	0	251	485	1.20
1985	190	10	118	10	0	57	384	1.75
1986	342	538	581	0	148	174	1782	1.60
1987	481	306	309	0	442	113	1651	0.89
1988	467	52	44	35	0	74	671	1.06
1989	469	73	259	24	350	210	1386	1.09
1990	514	152	340	29	130	118	1282	1.14
1991	467	1482	2734	220	301	182	5385	2.48
1992	768	1432	788	12	55	118	3172	2.27
1993	763	716	1092	313	197	11	3091	1.48
1994	737	451	608	184	0	0	1980	1.02
1995	716	670	1605	100	0	0	3091	1.62

Source: The methodology for the construction of the venture financing and biotechnology index series is described in Lerner (1). The IPO and follow-on offering series are from unpublished databases of Recombinant Capital. The private placement, debt and convertible, and RDFO compilations are based on the compilations of Shane (2), extended in time and comprehensiveness through searches of a wide variety of sources.

Note: The table summarizes the total raised (in millions of 1995 dollars) by U.S. new biotechnology firms through several major sources: venture capital investments in private firms, initial public offerings (IPOs), follow-on public equity offerings, private placements by financial investors in public firms, debt and convertible security issues, the issuance of shares in R&D financing organizations (RDFOs), and the sum of these offerings. Alliance-related financings are excluded. The table also indicates the year-end level of an inflation-adjusted index based on the valuation of publicly traded biotechnology firms in this period, normalized to be equal to one on January 1, 1978.

fact that in 1995, the dollar volume of commitments to new alliances in the biotechnology industry was almost equal to venture capital disbursements *in all industries* ($3.4 billion vs. $3.7 billion).

Two limitations of this article should be acknowledged at the outset. First, I do not attempt to duplicate the guides that explain the intricacies of the alliance process to practitioners. Numerous excellent volumes exist (5–7), that document the legal and institutional considerations associated with undertaking biotechnology alliances and technology licenses at much greater depth than could be done here. Second, my focus is primarily on the empirical evidence about these alliances. While a number of works about the theory of technology alliances and licensing are mentioned in passing, the primary focus is on the empirical research.

WHY ARE ALLIANCES AND LICENSING SO IMPORTANT IN BIOTECHNOLOGY?

Academic technology transfer officers and executives at small biotechnology companies often face the challenge of commercializing early-stage biotechnologies with tremendous promise. But a variety of considerations make it difficult to raise financing from traditional sources — such as banks and public investors — for some of the most potentially profitable and exciting technologies. As a result in many cases they are required to turn to strategic alliances for financing. These difficulties can be sorted into four critical factors: uncertainty, asymmetric information, the nature of firm assets, and the conditions in the relevant financial and product markets.

The first of these four problems, uncertainty, is a measure of the array of potential outcomes for a company or project. The wider the dispersion of potential outcomes, the greater the uncertainty. By their very nature, young biotechnology companies are associated with significant levels of uncertainty: Only a relative handful of drugs actually become commercial products, and a small subset of these proves to be profitable. The extent of intellectual property protection that a new biotechnology product will receive is also often very uncertain. High uncertainty means that entrepreneurs and investors cannot confidently predict what the company will look like in the future. Uncertainty affects the willingness of investors to contribute capital, the desire of larger firms to license unproved technologies, and the decisions of firms' managers.

The second factor, asymmetric information, is distinct from uncertainty. Because of his day-to-day involvement with the technology, a scientist knows vastly more about his discovery's prospects than prospective investors. Various problems develop in settings where asymmetric information is prevalent. For instance, the entrepreneur may take detrimental actions that outsiders cannot observe: perhaps undertaking a riskier strategy than

Table 2. Interfirm Alliances in Three Industries, and by Biotechnology Firms

Panel A: Interfirm alliances by U.S. firms in three industries, 1980–1994

	Number of New Alliances Publicized, by Nationality of Firms		
Year	U.S.–U.S.	U.S.–Europe	U.S.–Japan
1980	42	40	15
1981	48	30	26
1982	57	54	39
1983	51	37	51
1984	88	60	55
1985	86	82	52
1986	118	78	47
1987	133	95	53
1988	141	98	39
1989	122	86	44
1990	121	66	34
1991	106	53	51
1992	155	89	43
1993	192	104	45
1994	235	145	40

Panel B: Interfirm alliances by U.S. biotechnology firms, 1981–1997

		Payments through alliances (millions of 1995 dollars)	
Year	Number of new filed alliances	Precommercial payments promised in new alliances	Actual payments during year to 49 leading firms
1981	30		9
1982	35		111
1983	31		152
1984	42		210
1985	57		149
1986	63		184
1987	62		415
1988	64		298
1989	71		205
1990	81		851
1991	115	741	647
1992	75	931	392
1993	113	1373	806
1994	66	1772	
1995	171	3421	
1996		2334	
1997		4352	

Source: The first panel is from National Science Board (3). The number of new alliances and precommercialization payments series in the second panel are from Recombinant Capital and its unpublished databases. The actual payments series is from Shane (2). It has not been extended beyond 1993 or to include additional firms.

Note: The first panel presents the number of publicized alliances by U.S. firms in three industries — information technology, biotechnology, and advanced materials — between 1980 and 1994. The second panel examines only alliances involving U.S. biotechnology companies between 1981 and 1995 filed with the U.S. Securities and Exchange Commission or state regulatory bodies who make such information public. Presented are the number of new filed alliances each year, the sum of all promised precommercialization payments in the filed alliances that year (the sum of the nominal payments is expressed in millions of 1995 dollars), and the actual payments to a sample of 49 of the largest biotechnology firms in each year (in millions of 1995 dollars).

initially suggested or not working as hard as the investor expects. The entrepreneur might also invest in projects that build up his reputation at the investors' expense.

Asymmetric information can also lead to selection problems. The scientist who makes a potentially important discovery may exploit the fact that he knows more about the project or his abilities than his investors do. Licensees may find it difficult to distinguish between truly revolutionary technologies and impractical ones. Without the ability to screen out unacceptable projects, outsiders are unable to make efficient and appropriate decisions choices regarding where to invest. These problems have been particularly severe in biotechnology, due to the scientific complexity of the development of new products and processes.

The third factor is the nature of the assets. Firms that have tangible assets — such as machines, buildings, land, or physical inventory — may find financing easier to obtain or may be able to obtain more favorable terms. The ability to abscond with the firm's source of value is more difficult when it relies on physical assets. When the most important assets are intangible, raising outside financing or entering into strategic alliances may be more challenging. In the biotechnology industry, firms have tended to rely on patent protection to protect assets. Those firms that have relied on trade secrets or informal "know-how" have found attracting investors or entering into licensing agreements to be very difficult. For instance, trade secrets offer exceedingly narrow intellectual property protection, only protecting against misappropriation: "the acquisition of a trade secret by a person who knows or has reason to know that the trade secret was acquired by improper means" (8). Thus a firm cannot sue a rival who discovers its trade secret independently or through "reverse engineering" (the disassembly of a device to discover how it works). This is unlike patent protection, which allows the awardee to prosecute others who infringe, regardless of the source of the infringers' ideas. Pooley (9) notes that very few "naked" trade secret licenses are observed, suggesting that the information covered only through this very narrow property right is difficult to transfer in an arm's-length exchange. Further evidence of the importance of broad patent protection in biotechnology is found in Lerner (10).

Market conditions also play a key role in determining the difficulty of financing firms. Both the capital and product markets may be subject to substantial variations. The supply of capital from public investors and the price at which this capital is available may vary dramatically. These changes may be a response to regulatory edicts or shifts in investors' perceptions of future profitability. The availability of equity from public investors for new biotechnology firms has been particularly variable. As Figure 1 indicates, the amount raised by publicly traded biotechnology firms in follow-on offerings (measured in 1995 dollars) went from $340 million in 1990 to $2.7 billion in 1991, then fell again to $788 million in 1992. Since young biotechnology firms face enormous costs while developing new products, they are typically very aggressive in raising capital. Practitioner accounts suggest that during the periods when there are few public equity issues, the markets are essentially "closed" to biotechnology firms.

As a result of all these problems, biotechnology firms often have little choice but to turn to corporations for financing. A pharmaceutical firm or other corporation with a related line of business can overcome many of these information problems, by undertaking extensive due diligence prior to the transaction and monitoring the firm afterwards. As discussed in Lerner (11), this type of intensive oversight is also provided by venture capitalists, though the resources that venture capitalists invest in the typical company are much smaller than those that a major corporation will devote to a substantial technology alliance. The corporation may have assets, such as sales forces and manufacturing know-how, which young biotechnology firms lack and yet are essential to the successful introduction of a new product. Small, research-intensive firms frequently rely on alliances with larger corporations to avoid having to construct these capabilities, which may take years to develop. Furthermore, the ongoing operations of the corporate partner may enable it to overcome some of the problems associated with the intangible nature of the biotechnology firm's assets.

Described in this manner, these problems may appear to be quite abstract. But they have very real implications for academic technology managers or corporate executives seeking to commercialize early-stage biotechnologies. They may find investors unwilling to invest the time and resources to examine early-stage technologies, or offering only modest payments in exchange for large stakes in innovations that the scientists, technology transfer officers, and company executives believe to be quite valuable. In the remainder of this section, I will summarize the evidence regarding these claims.

This set of suggestions has been most directly examined in two recent Ph.D. dissertations. Both Shane (2) and Majewski (12) examine the decision of biotechnology firms to raise capital through the public markets and through alliances. Through their different methodologies, these works show that firms turn to alliance financing when asymmetric information about the biotechnology industry is particularly high. During these periods — which are measured through such proxies as the variance of the returns of biotechnology securities — firms are likely to delay the time until their next equity issuance, and to rely on alliances rather public offerings as a source of external financing. The authors argue that the greater insight on the part of the pharmaceutical company into the nature of the biotechnology firm's activities allows it to make successful investments at times when uninformed public investors are deterred by information problems.

A challenge to this view is Pisano's (13) examination of the outcome of biotechnology firms' research projects that are and are not developed with the help of alliances with pharmaceutical firms. He examines the probability that a biotechnology company successfully develops a drug being pursued through an alliance, as opposed to one developed by the firm itself. His conclusion that a "lemons problem" leads biotechnology firms to only undertake alliances with pharmaceutical companies that involve inferior technologies, however, seems difficult to reconcile with the large number and dollar volume of these transactions. (A lemons problem can arise when, because of his day-to-day involvement with the firm, an entrepreneur knows more about his company's prospects than investors, suppliers, or strategic partners. Because the counterparty in the exchange is at an informational disadvantage, the potential partner or investor may refuse to enter into a transaction at all.) One way to reconcile these observations is the possibility that biotechnology firms pursue projects through alliances that have lower expected probabilities of success, but whose ultimate payouts are greater.

A related set of work has looked at the decision of the larger company to undertake an alliance. Pisano (14) examines pharmaceutical firms' choices between developing new drugs in-house versus through alliances in the

case of 92 drug development projects. He demonstrates that the insights of transaction cost economics (15) and incomplete contracting theory (16) are highly relevant here. In brief, both sets of work highlight the difficulties that two parties may have in undertaking contracts when there is considerable uncertainty. (Among the barriers to writing optimal contracts are the difficulty of negotiating detailed contracts and the problem of being unable to foresee all contingencies.) In particular, when there are few small biotechnology firms working in the field, the pharmaceutical company is more likely to undertake the project in-house. This is consistent, Pisano argues, with the theoretical suggestions that "hold up" problems — that is, efforts by one of the parties to renegotiate the contract on terms more favorable to itself after the agreement is signed — will be greater in this setting. Pisano also finds evidence that firm-specific factors are critical in the decision to undertake alliances.

The determinants of the firm-specific differences in the rate of alliance formation by various pharmaceutical and biotechnology firms are the other major focus of the writings on why firms form alliances. Arora and Gambardella (17) are two of the few economists to examine these firm-specific factors. (Most researchers of these questions have come from an organizational or sociological perspective.) They (17) examine four strategies — alliances with biotechnology firms, research collaborations with universities, purchases of minority interests in biotechnology firms, and acquisitions of these firms — by large pharmaceutical and biotechnology firms. They present a model that suggests that these activities will be complements: If a firm pursues one of these activities, it should be likely to undertake all of them. Furthermore they suggest that these activities should be disproportionately pursued by firms with a greater internal knowledge of biotechnology. On the basis of their analysis of 81 firms, the authors conclude that the four activities are indeed complements, and are disproportionately undertaken by firms with large existing stocks of biotechnology patents. Similar conclusions emerge from a later study by Arora and Gambardella (18). [These studies should be viewed in light of the critique by Athey and Stern (19) of studies of complementarities.] Other studies that have highlighted the importance of firm-specific factors in biotechnology alliance include Argyres and Liebskind (20) and Roberts and Mizouchi (21).

Another approach to firm-specific factors has characterized the organization literature. In particular, these works have highlighted the importance of networks of firms. In industries with rapidly changing and evolving technologies, this literature argues, knowledge is diffused across a variety of firms. As a result firms seek to learn through the formation of alliances. This literature highlights the importance of alliances early in an industry's evolution. These can form the foundation for repeated relationships and further learning.

Powell, Koput, and Smith-Doerr (22) highlight the path-dependent nature of alliance formation in biotechnology. The decision by both biotechnology and pharmaceutical firms to establish alliances seems to be driven less by characteristics such a biotechnology firm's age or growth rate, but rather by earlier experience with such alliances. The authors argue that these firms learn about how to manage and absorb knowledge from their initial strategic alliances. The authors further suggest that firms that are more "centrally" located in alliance networks are more likely to expand their alliance activities.

Some of the consequences of alliances are highlighted in Stuart, Hoang, and Hybels (23) and Koput, Powell, and Smith-Doerr (24). These studies show that biotechnology firms with prestigious sponsors (e.g., financing from well-established venture capital groups and collaborations with prominent pharmaceutical companies) benefit from these relationships. In the former paper, it is shown that biotechnology firms with prestigious partners are likely to complete an IPO sooner and to command a higher market capitalization at the time of the offering. The latter paper documents that companies with such partners are more likely to enter into other collaborations, whether with other elite or non-elite firms.

HOW ARE ALLIANCES STRUCTURED?

The second broad question that this article considers is the manner in which alliances are structured. Unlike the formation of alliances, which as highlighted above has been explored by both economists and sociologists, economists have played a leading role in the study of this issue. Given the dynamic and complex nature of R&D alliances, it is not surprising that there are multiple potential explanations for their structure. This section will begin by highlighting three classes of relevant theoretical research.

The first of these is incomplete contracting theory. A wide range of models, beginning with Grossman and Hart (16) and Hart and Moore (25) and summarized in Hart (26), consider incomplete contracting between a principal and an agent. A typical assumption is that it is impossible for the two parties to write a verifiable contract that could be enforced in a court of law that specifies the effort and final output of the two parties. This is because there are many possible contingencies, all of which cannot be anticipated at the time the contract is drafted. Because of this nonverifiability problem, these models argue that it is optimal for ownership of the project to be assigned to the party with the greatest marginal ability to affect the outcome. This party, who will retain the right to make the decisions that cannot be specified in the contract, should also receive any surplus that results from the project. Because of this incentive, the party will make the decisions that maximize — or come close to maximizing — the returns from the project. An alternative, though complementary, view suggests that firms may write "excessively" incomplete contracts. In particular, Bernheim and Whinston (27) show that in settings where one set of behaviors cannot be contracted upon, it may be optimal to leave other aspects of ownership unsettled. Such "strategic ambiguity" is more likely in settings with greater uncertainty, consistent with the predictions of the Grossman-Hart-Moore class of models.

Aghion and Tirole (28) adapt this general model to a R&D alliance between two firms. As long as the R&D-performing firm has the initial bargaining power or does not face capital constraints, the results are as discussed above: The control rights are assigned to the party whose marginal contribution to the project's success is greatest. When the financing firm has the initial bargaining power and the R&D firm is capital constrained, however, a different pattern may emerge. In particular, if it is optimal for the property rights to be transferred to the R&D firm, the best outcome will not be achieved: The financing firm will be willing to transfer ownership, but the cash-constrained R&D firm will not have enough resources to compensate the financing firm. As a result, an inefficient allocation of the property rights occurs, with the financing firm retaining the rights to the invention.

Biotechnology research has numerous features that resemble the setting depicted in the theoretical literature on incomplete contracts. Biotechnology projects — particularly early-stage efforts — are highly complex and uncertain, making it very difficult to specify the features of the product to be developed. As one biotechnology executive relates:

> Redefining the work when the unexpected happens, as it invariably will, [is essential]. Research is by its very nature an iterative process, requiring constant reassessment depending on its findings. If there is a low risk of unexpected findings requiring program reassessment, then it is probably not much of a research program. (29, pp. 220–221)

Similarly the complexity and unpredictability of the research presents challenges in drafting an enforceable agreement that specifies the contributions of the R&D firm. In particular, firms that contract to perform R&D in alliances frequently have ongoing research projects of their own, in addition to the contracted efforts. In case of a dispute, it may be very difficult for the financing firm to prove that the R&D firm has employed alliance resources to advance projects that are not part of the alliance.

At the same time biotechnology alliances present a more complex picture than many incomplete contract models. Typically these models assume a one-time contracting process between the two parties. Actual alliances reveal more complex contracting patterns. For instance, pairs of firms undertake repeated sets of alliances on different topics. These prior interactions may allow firms to develop reputational capital, and at least partially address some of the contracting problems. Second, these models often assume a vertical relationship: The agent contributes all the effort. In actuality, the relations between some alliance parties are likely to have horizontal elements. For instance, some of the alliances are between pairs of biotechnology concerns, each of which may contribute knowledge. Third, typically parties in these models bargain over a very reduced set of parameters, with a single ownership right being divided between the parties. Real-life agreements, as noted above, are much more complex. Finally, in many models, such as Aghion and Tirole (28), there is no consideration of the impact of asymmetric information on the negotiation of the alliance: Both parties are assumed to be fully informed at the time the agreement is signed. Rather, the major concern of the negotiating parties is reducing the potential for suboptimal effort after the agreement is signed. In actuality, the situation may be more complex. In particular, the financing party may not be fully informed about the prospects for the project. This may lead to contractual terms that seek to force the informed party to reveal if he has additional information.

This class of models suggests a variety of empirical implications. In general, among alliances where informational asymmetries are greater, the period during which the R&D firm maintains active control over the project should be longer. For instance, alliances that focus on drug development, which is a much more expensive and complex process than the development of diagnostics and other biotechnology applications, are likely to have much greater informational asymmetries. The incomplete contracting hypothesis consequently suggests that these alliances should be associated with longer contract lives.

A second and quite contrasting explanation is the need to provide monitoring. An essential assumption of the incomplete contracting literature is that addressing the agent's (R&D firm's) behavior after the contract is signed is very difficult for the principal (financing firm). In particular, because the contractual agreements cannot foresee every contingency, even if the financier observes problematic behavior on the part of the agent, he cannot address the behavior in a court of law.

The costly monitoring hypothesis, on the other hand, assumes that such behavior can be addressed, but to do so is expensive. Provisions that allow greater oversight and control on the part of the principal are expensive to negotiate and oversee. As a result these terms should be only included in contracts when the danger of opportunistic behavior is high or the costs of oversight modest. The costly monitoring hypothesis [as articulated, for instance, in Smith and Warner (30) and Williamson (15)] predicts that because the ease of monitoring and incentives to pursue opportunistic behavior vary, the optimal degree of restrictiveness will differ across contracts.

One manifestation of these trade-offs should be in the period that the corporation commits to finance the R&D of the smaller firm. If corporations could costlessly monitor the firm, they would monitor and infuse cash continuously. If the project's expected value fell below some value, the corporation would halt funding of the project. (The R&D firm might still wish to continue because the funds are enabling the firm to continue its existing projects, or because the management team itself is enjoying private benefits from the funding.)

The renegotiation or extension of a strategic alliance, however, is a costly process. Particularly within the financing firm, there is likely to be the need for extensive analysis and review prior to the modification of an existing agreement. The renegotiation process is likely to involve a variety of internal and external legal advisors. As a result alliances are negotiated for distinct periods, not continually reviewed and funded.

This view suggests that the duration of these alliances should be a function of the degree of potential agency

problems and the cost of providing monitoring. For example, alliances that focus on the complex drug development process are likely to face much greater informational asymmetries. In these settings the costly monitoring hypothesis suggests a greater need for monitoring and shorter alliance lives.

A third explanation is the need for avoiding the costs of financial distress. An extensive corporate finance literature has documented these costs, which can be divided into three classes:

- The liquidation of a biotechnology firm can be highly destructive of value. In particular, it is often difficult to sell "naked" patent awards without the associated know-how and trade secrets. Without the researchers who developed the concept, most patent awards are likely to be worth little. More general evidence of the substantial indirect costs of financial distress is found in Lang and Stultz (31) and Opler and Titman (32).
- Even if the biotechnology firm is not liquidated, concerns about financial distress may cause the firm to be unable to pursue value-creating investment opportunities. Models such as Stultz (33) and Froot, Scharfstein and Stein (34) suggest that if information asymmetries at times preclude external financing or else make it very costly, firms may be unable to pursue value-creating projects.
- Even if financial distress imposes few costs on the firm's shareholders or society as a whole, managers may still seek to avoid distress. In particular, Smith and Stultz (35) formally show that a risk-averse manager who owns shares in a company is likely to engage in socially undesirable levels of risk management. Managers may fear that the bankruptcy of the firm for which they work will be very costly, both in terms of personal wealth and future earning potential (due to the reputational consequences).

This danger of financial distress is very real in the biotechnology industry: Numerous firms have been forced to liquidate or radically trim back promising research programs because of an inability to access external financing.

Motivated by any one of these reasons, the managers of R&D firm may seek to limit the firm's potential exposure to financial distress. The managers may see undertaking one or more corporate alliances as an attractive mechanism to this end. In exchange for giving up much of the eventual profits from an innovation, the R&D firm receives a guaranteed stream of payments from the financing firm. The R&D firm's desire to engage in such "risk management" is likely to be an increasing function of the potential costs of financial distress and the probability that such an event will occur. (Presumably a risk-neutral financing firm would acquiesce to such an agreement in exchange for a lower royalty other concessions.)

While the variation in the extent of the social or managerial costs of financial distress is difficult to observe, it is possible to identify characteristics of the R&D firm that are likely to be associated with a higher probability of distress. For instance, firms whose research focuses on costly drug development are likely to face a greater probability of financial distress, and may be more willing to enter into alliances that guarantee protracted financial payments.

Pisano (36) first examined these questions in a pioneering work. He studied 195 collaborative agreements between biotechnology and pharmaceutical firms, and asked where purchases of equity by the pharmaceutical company were used alongside contractually specified governance rights (e.g., the pharmaceutical company's right to obtain periodic briefings on the progress of the biotechnology firm). Consistent with the costly monitoring view, he found that in settings with greater information problems and information asymmetries (e.g., when the project was R&D intensive or the alliance entailed multiple projects), the pharmaceutical company was more likely to purchase equity as a part of the agreement. The added rights associated with equity ownership can be seen as strengthening the pharmaceutical company's ability to control the biotechnology firm.

Lerner and Merges (37) examine the determinants of control rights within a sample of 200 alliances. They analyze the share of 25 key control rights allocated to the financing firm by regressing the assigned number of rights on independent variables denoting the project stage and financial conditions, as well as controls for a variety of alternative explanations. Consistent with the framework developed by Aghion and Tirole (28), the greater the financial resources of the R&D firm, the fewer control rights are allocated to the financing firm. For instance, a one standard deviation increase in shareholders' equity at the mean of the independent variables leads to an 11 percent drop in the predicted number of control rights assigned to the financing firm. Evidence regarding the relationship between control rights and the stage of the project at the time the alliance is signed is less consistent with existing theory. Projects in their early stages at the time of alliance formation actually assign significantly less control to the R&D firm.

Lerner and Tsai (38) explore the impact of the financing environment at the time was signed on the success of agreements. They show that in periods where financing availability was strong, the agreements were more successful, whether measured by the probability that the drug advanced to the next stage in the clinical trials or was approved. They show that the effect was more pronounced in those agreements where the biotechnology company received little of the control, as Aghion and Tirole predict. This helps address concerns that the result is driven by shifts in an unobserved third factor. Lerner and Tsai also examine the likelihood of renegotiation. If it would maximize innovative output to assign control to the small biotechnology company, though this allocation of control is precluded by financial market conditions, then there should be evident a distinct pattern in renegotiations. In particular, when financing conditions improve for biotechnology firms, it is those agreements assigning the bulk of the control to the major pharmaceutical firm that should be disproportionately renegotiated. The empirical results are consistent with this pattern.

There is also a small but growing empirical sociology literature on the governance of inter-firm transactions. (My thanks to the referee for highlighting this literature. Also related is the much larger literature in the area of social networks, which addresses the communication properties of networks in situations characterized by embedded exchange.) These issues were framed by a theoretical essay by Granovetter (39), who argued that the structure of the alliance network in any given industry plays an important role in diffusing information about the reliability of various industry participants. Access to this information, in turn, will play an important role in influencing how a new alliance will be governed. For instance, a major pharmaceutical company already engaged in a number of alliances will find that its existing relationships provide access to reliable and cheap information about other actors. Information flowing through connections such as shared third parties (i.e., when two firms have an alliance with the same third organization) are quite effective, which allow well-connected contracting parties greater flexibility in structuring collaborative arrangements. The empirical findings in this literature, most importantly Podolny (40) and Gulati (41), are consistent with Granovetter's reputation model. Two of the more robust findings in this literature are that alliances are less likely to have an equity component when two firms have previously formed an alliance or when they are proximately located in the network of prior deals.

Research into contract structure is at an earlier stage than that about the determinants of alliance formation. Thus, the extent of uncertainty about the key drivers of contract structure is not surprising. The proliferation of information on alliances available through database companies such as Recombinant Capital (much of which is publicly accessible on the Internet at *http://www.recap.com*) and the Securities and Exchange Commission's EDGAR database (available at *http://www.sec.gov*) should encourage future researchers.

WHAT IS THE EVIDENCE FROM FIELD RESEARCH?

Another important source of information on biotechnology alliances is case study research. The conclusions of these cases are often not as "neat" as statistical analyses that are crafted to examine a particular question, but field-based analyses can generate a variety of insights. This section highlights the experiences of three companies that have been examined in case studies. While not exhaustive of the case study literature on biotechnology alliances, they suggest the richness of insights that field-based research can provide.

These three young companies all were developing advanced human therapeutics and grappling with the challenges posed by alliances. The biotechnologies pursued by the three firms are quite different: antigen-based allergy drugs (ImmuLogic Pharmaceutical Corporation), advanced drug delivery mechanisms (ALZA Corporation), and monoclonal antibody-based treatments of inflammation (Repligen Corporation). There were considerable differences in the location and sophistication of strategic partners and the stage of development of the technologies.

One point that these cases raise — consistent with the literature discussed above — is how the allocation of control rights is determined both by concerns about behavior after the alliance is signed and by relative bargaining power. One alliance that may be considered successful in many respects was Repligen's May 1992 alliance with Eli Lilly regarding a very early-stage effort to develop a monoclonal antibody-based treatment of inflammation after heart attacks (43). The net-of-market return for Repligen in the three-day window around the announcement of the transaction in May 1992 was +9 percent, and that of Lilly, +2 percent. [These increases can be compared to the +2.1 percent reaction to 55 announcements of R&D initiatives by high-technology firms found by Chan et al. (42).] The early-stage project succeeded in getting its lead product candidate into Phase I trials in just 13 months. (After extending the project in June 1995, however, Lilly canceled its involvement three months later, citing shifting internal priorities.)

In the Repligen-Lilly alliance, three control rights were the subjects of protracted negotiations. The first was the management of clinical trials: the right to decide which drugs would be pursued and when. A second was the control over the marketing strategy, an arena in which Lilly had extensive experience and Repligen only a slight acquaintance. Finally, both parties wished to control the process development and ultimate manufacturing of the drug. Repligen compared favorably on various financial measures to other biotechnology firms at the time that the alliance was signed. Similarly the firm had outperformed an index of biotechnology securities over its history by over 40 percent. (Stock price performance is measured from the close of the day of Repligen's initial public offering to avoid including the "underpricing" of the offering — i.e., the discount at which the underwriters sold the shares to the original investors. Repligen's beta did not differ materially from that of other biotechnology firms.) At the same time, investment banking analysts had expressed concern about the financial pressures that might result if Repligen's earlier alliance with Merck was terminated.

The terms of the alliance that emerged from the negotiations appeared to assign the control rights to the parties whose behavior would have the greatest impact on the product development effort. Repligen was allowed a great deal of control over developing the lead product candidate, an area where it had considerable experience, but tangential product development activities were subject to extensive review by Lilly. Lilly was assigned control over all aspects of marketing, while Repligen was assigned all manufacturing control rights, unless it encountered severe difficulties with regulators.

Other alliances illustrate the importance of the relative bargaining power of the two parties. An example was the January 1978 alliance between ALZA and Ciba-Geigy (44,45). At the time of the alliance, ALZA faced a major financial crisis. The firm had little more than $1 million in the bank, was spending $2 million more per month than it was receiving in revenues, had nearly exhausted its bank credit line, was in violation

of several loan covenants, and was precluded from a sale of equity to the public by unfavorable market conditions and the perception that ALZA had been excessively optimistic in its earlier communications with investors and analysts.

The alliance assigned almost total control to the Swiss pharmaceutical giant. Ciba-Geigy was given a supermajority on the joint board that reviewed and approved potential research projects, the right to license and manufacture any of ALZA's current or future products, the ability to block any other alliances that ALZA proposed to enter into, and 8 of the 11 seats on ALZA's board of directors. In addition the Swiss pharmaceutical giant received a new class of preferred shares. If converted into common stock, the new preferred shares would represent 53 percent of the equity in ALZA. Until conversion, however, Ciba-Geigy had 80 percent of the voting rights, an allocation that allowed it to employ ALZA's tax losses.

At the same time it is reasonable to believe that concerns about the postalliance behavior of ALZA also motivated Ciba-Geigy to demand strong control rights. ALZA's leaders had displayed little ability to direct the firm's research effort over the course of the 1970s. This may have led Ciba-Geigy to conclude that the benefits of allocating control rights to ALZA's management were limited. Despite the strict control rights contractually assigned to Ciba-Geigy, there were frequent disputes between the two firms as ALZA researchers sought to either circumvent the pharmaceutical firm's middle management or ignored their instructions outright. Frustrated by these problems, Ciba-Geigy agreed to terminate the alliance and sell back its equity to ALZA in November 1981.

A contrasting illustration is presented by ImmuLogic (46). In March 1991 the firm was considering either entering into an alliance or raising equity in an initial public offering. One concern that led the firm to decide to go public was that a potential strategic partner might exploit its relatively weak financial condition. In other words, ImmuLogic feared that a pharmaceutical company might obtain numerous concessions on key governance and financial issues by protracting the negotiations until ImmuLogic was close to running out of capital. It consequently deferred negotiating an alliance to develop and market its allergy drugs until the firm went public in May 1991. The firm announced an alliance with Marion Merrell Dow in December 1991, which allowed ImmuLogic to retain numerous control rights, such as an equal role in planning marketing strategy in the United States: *In Vivo* magazine hailed the transaction as "push(ing) the limit of the biotech deal . . . a partnership in fact as well as name" (quoted in Ref. 46, Teaching Note 5-293-118, p. 7). Just as ALZA's relinquishment of almost total control to Ciba-Geigy was in large part a consequence of its weak financial position, ImmuLogic's ability to obtain these control rights reflected its financial strength.

These cases also emphasize two issues that are not highlighted in the theoretical literature. One is the interaction between the allocation of control rights and the financial terms of the transactions. For instance, in the negotiations that led to Repligen's retention of control over manufacturing, the firm agreed to an alteration in its compensation. Repligen accepted a lower royalty than originally envisioned, 5 percent of the sales price, but agreed to supply the drug to Lilly at a price (about 15 percent of the sales price) above what it believed its true manufacturing cost would be. Repligen agreed to reduce the price that it charged Lilly if it was able to manufacture the drug for less, but only if its cost was below 8 percent of the sales price.

A second interesting and unexplored aspect is the apparent signal that the allocation of control rights provided to potential investors and other outsiders. Both ImmuLogic and Repligen highlighted their retention of key control rights in the press releases announcing the transactions described here. Their ability to obtain these rights attracted favorable comments in the trade press and analyst reports alike. These patterns suggest a richer set of interactions than theoretical treatments of these issues imply.

WHERE IS FURTHER RESEARCH NEEDED?

This article has sought to suggest the importance, richness, and complexity of alliances and licensing in the biotechnology industry. While much has been learned from the economic and sociological research into biotechnology alliances over the past decade, much more remains to be discovered. This final section will highlight two issues that deserve to be a particular focus of attention.

The first of these relates to the structure of the payments between the financing and the R&D firm. The design and implementation of incentive schemes in general is a major focus of the finance and economics literature, but payments in alliances have been little examined except in theoretical works (47,48). This lack of attention is a reflection of the difficulty in analyzing them. The payments typically are of several types: an initial up-front payment, a purchase of equity (which the financing firm may be able to force the R&D firm to repurchase if the alliance is unfruitful) or warrants, commitments to contract for R&D on specific topics, milestone payments contingent on the achievement of technological and marketing objectives or the renewal of the agreement, and a royalty on the eventual sales generated by the product. Assessing the expected net present value of these payments is very difficult. The magnitude and timing of eventual sales that the project will generate are difficult to anticipate. The amount of the R&D to be contracted for is often ambiguous. Alliances may also include contingent payments for remote outcomes. (The rationale for their inclusion is that firms frequently report — and analysts tabulate when assessing firms — the sum of all precommercialization payments from new alliances, whether the funds are likely to be received or not. These contractually specified contingent payments may thus convey important strategic benefits, even if the probability of payment is very low.) Clearly, this is a difficult but important area for research.

A second question relates to the impact of this contracting regime on the rate and direction of technological innovation in the biotechnology industry. Zucker, Darby,

and Brewer (49), in their analysis of the impact of academic research on the development of the biotechnology industry, suggest that academic licensing practices are one of the reasons for geographic clustering of innovative biotechnology firms around top-tier universities. Theoretical work by Gans and Stern (50) provides reasons to believe that the impact of contracting on the pace of innovation may be complex and multidimensional. The impact of intercorporate licensing on innovation is an important issue for future research, given the number and financial significance of these transactions. While a few initial steps along these lines have been taken (51,52), much more remains to be done.

ACKNOWLEDGMENTS

Harvard University and National Bureau of Economic Research. I thank an anonymous referee for helpful comments. Parts of this essay are based on Lerner and Merges (37) and Lerner and Tsai (38). Financial support was provided by the Division of Research at Harvard Business School.

BIBLIOGRAPHY

1. J. Lerner, *J. Financial Econ.* **35**, 293–316 (1994).
2. H.L. Shane, Asymmetric Information and Alliance Financing in the Biotechnology Industry, in *Three Essays in Empirical Finance in High-Technology Firms*, Unpublished Ph.D. Dissertation, Wharton School, University of Pennsylvania, Philadelphia, 1995.
3. National Science Board, *Science and Technology Indicators — 1998*, U.S. Government Printing Office, Washington, DC, 1998.
4. Industrial Research Institute, *Industrial Research Institute's Trends Forecast for 1998*, Industrial Research Institute, Washington, DC, 1997.
5. *Corporate Alliances: Strategies in Biotechnology*, KMPG Peat Marwick, Biotechnology Industry Organization, and Pittiglio, Rabin, Todd & McGrath, San Francisco, CA, 1993.
6. B. Cunningham, *Issues and Trends in Biotech Corporate Partnering*, Cooley, Godward, Castro, Huddleston & Tatum, Palo Alto, CA, 1994.
7. J.W. Schlicher, *Licensing Intellectual Property: Legal, Business, and Market Dynamics*, Wiley, New York, 1996.
8. R.M. Milgrim, *Milgrim on Trade Secrets*, Matthew Bender, New York, 1993.
9. J. Pooley, *Trade Secrets: A Guide to Protecting Proprietary Business Information*, American Management Association, New York, 1989.
10. J. Lerner, *Rand J. Econ.* **25**, 319–333 (1994).
11. J. Lerner, *J. Finance* **50**, 301–318 (1995).
12. S.E. Majewski, Causes and Consequences of Strategic Alliance Formation: The Case of Biotechnology, Unpublished Ph.D. Dissertation, Department of Economics, University of California at Berkeley, 1998.
13. G.P. Pisano, *R&D Performance, Collaborative Arrangements and the Market-for-Know-How: A Test of the 'Lemons' Hypothesis in Biotechnology*, Working Paper No. 97–105, Harvard Business School, Boston, MA, 1997.
14. G.P. Pisano, *Admin. Sci. Q.* **35**, 153–176 (1990).
15. O.E. Williamson, *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, Free Press, New York, 1985.
16. S.J. Grossman and O.D. Hart, *J. Political Econ.* **94**, 691–719 (1986).
17. A. Arora and A. Gambardella, *J. Ind. Econ.* **38**, 361–379 (1990).
18. A. Arora and A. Gambardella, *J. Econ. Behav. Organ.* **24**, 91–114 (1994).
19. S. Athey and S. Stern, *An Empirical Framework for Testing Theories About Complementarity in Organizational Design*, Working Paper No. 6600, National Bureau of Economic Research, Cambridge, MA, 1998.
20. N.S. Argyres and J.P. Liebeskind, *Acad. Manage. Rev.*, (in press).
21. E.B. Roberts and R. Mizouchi, *Int. J. Technol. Manage.* **4**, 43–61 (1989).
22. W.W. Powell, K.W. Koput, and L. Smith-Doerr, *Admin. Sci. Q.* **41**, 116–145 (1996).
23. T.E. Stuart, Ha Hoang, and R.C. Hybels, *Admin. Sci. Q.* **44**, 315–349 (1999).
24. K.W. Koput, W.W. Powell, and L. Smith-Doerr, Interorganizational Relations and Elite Sponsorship: Mobilizing Resources in Biotechnology, Unpublished working paper, University of Arizona, Tucson, 1997.
25. O.D. Hart and J. Moore, *Econometrica* **56**, 755–785 (1988).
26. O.D. Hart, *Firms, Contracts, and Financial Structure*, Oxford University Press, New York, 1995.
27. B.D. Bernheim and M.D. Whinston, *Am. Econ. Rev.* **88**, 902–932 (1998).
28. P. Aghion and J. Tirole, *Q. J. Econ.* **109**, 1185–1207 (1994).
29. J.P. Sherbloom, in R. Dana Ono, ed., *The Business of Biotechnology: From the Bench to the Street*, Butterworth-Heinemann, Stoneham, MA, 1991, pp. 213–224.
30. C.W. Smith, Jr. and J.B. Warner, *J. Financial Econ.* **7**, 117–161 (1979).
31. L.H.P. Lang and R.M. Stultz, *J. Financial Econ.* **32**, 45–60 (1997).
32. T. Opler and S. Titman, *J. Finance.* **49**, 1015–1040 (1994).
33. R.M. Stultz, *J. Financial Econ.* **26**, 3–28 (1990).
34. K.A. Froot, D.S. Scharfstein, and J.C. Stein, *J. Finance* **48**, 1629–1658 (1993).
35. C.W. Smith, Jr. and R.M. Stultz, *J. Financial Quant. Anal.* **20**, 391–405 (1985).
36. G.P. Pisano, *J. Law, Econ., Organ.* **5**, 109–126 (1989).
37. J. Lerner and R.P. Merges, *J. Ind. Econ.* **46**, 125–156 (1998).
38. J. Lerner and A. Tsai, Do Equity Financing Cycles Matter? Evidence from Biotechnology Alliances, Unpublished working paper, Harvard Business School, Boston, MA, 1999.
39. M. Granovetter, *Am. J. Sociol.* **91**, 481–510 (1985).
40. J. Podolny, *Admin. Sci. Q.* **39**, 458–483 (1994).
41. R. Gulati, *Admin. Sci. Q.* **40**, 619–652 (1995).
42. S.H. Chan et al., *J. Financial Econ.* **46**, 199–221 (1997).
43. D. Kane and J. Lerner, *Repligen Corporation: January 1992*, Harvard Business School Case No. 9-294-082 (and Teaching Note 5-295-137), Harvard Business School, Boston, MA, 1994.
44. R. Angelmar and Y. Doz, *Ciba-Geigy/Alza Case Series (including Advanced Drug Delivery Systems: Alza and Ciba-Geigy, A through F; Alza Corporation, A and B; Ciba-Geigy Limited: Pharmaceutical Division, A through C)*, Unnumbered INSEAD case studies, Fontainbleau Cedex, France, 1988.

45. J. Lerner and P. Tufano, *ALZA and Bio-Electro Systems*, Harvard Business School Cases No. 9-293-124 through 9-293-127 (and Teaching Note 5-296-060), Harvard Business School, Boston, MA, 1993.
46. J. Lerner, *ImmuLogic Pharmaceutical Corporation*, Harvard Business School Cases No. 9-293-066 through 9-293-071 (and Teaching Note 5-293-118), Harvard Business School, Boston, MA, 1994.
47. N. Gallini and B.D. Wright, *Rand J. Econ.* **21**, 237–252 (1990).
48. M. Kamien and Y. Tauman, *Q. J. Econ.* **101**, 471–493 (1986).
49. L.G. Zucker, M.R. Darby, and M.B. Brewer, *Am. Econ. Rev.* **88**, 290–306 (1998).
50. J. Gans and S. Stern, Incumbency and R&D Incentives: Licensing the Gale of Creative Destruction, Unpublished working paper, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, 1998.
51. W. Shan, G. Walker, and B. Kogut, *Strategic Manage. J.* **15**, 387–394 (1994).
52. D.C. Mowery, J.E. Oxley, and B.S. Silverman, *Strategic Manage. J.* **17**, 77–91 (1996).

See other entries MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT; SCIENTIFIC RESEARCH, POLICY, TAX TREATMENT OF RESEARCH AND DEVELOPMENT.

T

TRANSFERRING INNOVATIONS FROM ACADEMIC RESEARCH INSTITUTIONS TO INDUSTRY: OVERVIEW

MICHAEL D. WITT
Technology Enterprise Development Company
Troy, Michigan

SUSAN K. LEHNHARDT
Morrison & Foerster LLP
New York, New York

OUTLINE

ACADEMIC RESEARCH INSTITUTIONS

Congress has made a national priority of bringing academe and industry together because of the beneficial effect of biomedical technology transfers on U.S. competitiveness and public health (1).

The importance of this interaction has evolved over time. Traditionally academe viewed its mission as above the fray of industry. Industry, likewise, relied on itself for research and development of new products and viewed academe as the "ivory tower" and not producing anything of commercial value. Since the 1960s academic research centers have attracted the best and brightest scientists and researchers to their laboratories. Large numbers of available and qualified applicants enabled research facilities to be rigorous and selective in their review of medical school and graduate school candidates. The students who survived this process were skilled, exceptionally talented, and dedicated to achieving success in the academic environment. With substantial federal support for research, mostly at the basic research level, a productive, unique, yet eccentric atmosphere developed where research was not inhibited by commercial or financial restraints.

Although this research capability is still largely in place, tremendous fiscal constraints for the last decade or more have eroded the ability of scientists to pursue basic research without at least some consideration of practical, commercial applications of their research. In addition federal reimbursement levels for medical care at academic medical centers (which also funds teaching, capital improvements and medical research) is increasingly restrictive.

The fiscal drain on academic science is exacerbated by heightened bureaucratic demands on medical research center personnel. Independent yet federally mandated peer review organizations, quality assurance programs, boards of medical examiners, state, federal and private insurance cost containment programs, and malpractice suits, each detract from the institutional mission, and cumulatively extract a profound professional toll on academic scientists. As a result frustration and anger is endemic, even among the most idealistic and dedicated.

In light of these developments, scientists still willing to work in the non-profit sector are turning to industry for research support (2). Research scientists who long eschewed corporate contacts are now more willing to seek out and perform corporate sponsored research.

Many corporations, similarly squeezed by considerable international competition, restrictive federal tax policies, product liability costs, and reduced access to the public financing markets, are looking anew at research centers (3). Corporations sense that internal research and development (R&D) might be productively supplemented with outside research efforts and technology, initially developed at government expense (3).

The resulting relationships between academics and industry require considerable patience. The parties also need to understand each other. Their motivation, stress, pressures and conflicting obligations are quite different.

Academics is Different than Industry

Scientists in the nonprofit sector are accustomed to responding to academic pressures. They are relatively unaccustomed to conducting exploratory research at the applied level as do their industry counterparts. The rate and method of academic research, the selection of research objectives, limited resources, and accountability are fundamentally different.

When an academic scientist and his peer group are convinced of a given result, the result often raises other questions and additional effort, but the given result, even if appropriate, is generally not developed into a product — in industry, this often is only a starting point. The result must be developed into a product, and tested further to the full satisfaction of management and the regulatory and licensing agencies. The scope of the commercial-grade scientist's work may be less purely inventive at times,

but the cost and effort required to create a product, then testing it, may be several hundred to thousand-fold the cost of the initial discovery.

Qualitative and Quantitative Differences in Research

Job security in the academic world requires quality teaching, writing, research, and a financial base provided by grants and government research funds. Reduced grant funding threatens this job security. It gives rise to tremendous pressure to generate research funds, and causes scientists to seek funding from any available resource. One source of research funding is the commercial sector. The commercial sector is willing to pay for research, and to pay for intellectual property that may lead to a product. Historically academicians were not financially rewarded for turning their research results into products. Thus they tended to carry the research effort only to the "proof of concept" stage, then moved on to their next activity. Policy changes at most academic centers now allow scientist to receive financial rewards from commercial successes (4).

As nonprofit organizations, academic research institutions are primarily funded by the federal government. Since the federal government has a major interest in promoting the development of innovations, it implemented legislation to facilitate the transfer of new technology from academic research institutions to industry, ensuring that scientists have an added incentive to turn their research results into products. On December 12, 1980, nearly 20 years ago, Congress enacted the Bayh-Dole Act to "reform U.S. patent policy related to government-sponsored research" (4, p. 3). The Bayh-Dole Act has two primary purposes:

- To enable universities, small business, and not-for-profit corporations to "patent and commercialize their federally funded inventions"
- To enable "federal agencies to grant exclusive licenses for their technology to provide more incentive to businesses" (4, p. 3)

The Bayh-Dole Act also has eight key regulations (4):

1. At the time of funding, an agency must make the funding university aware of its intention to hold title to an invention due to extraordinary circumstances or conditions.
2. Within two months of the date the inventor makes the university aware of the invention in writing, the university has to inform the proper federal agency of any invention created with the use of federal funds. This is the disclosure date.
3. To maintain ownership of an invention, the university basically has to inform the agency of its decision to keep the title within two years of the disclosure date. When the one year statutory period is initiated by public use, sale, or publication from which "valid" patent protection can be secured in the United States, the agency may reduce the "period of election to not more than 60 days prior to the end of the statutory period."
4. In order to use its invention, the university must give the U.S. government a paid-up, nontransferable, nonexclusive license (confirmatory license).
5. The government has "march-in rights" over the invention. This means that the government may relieve the university of its title to an invention if the university does not try to develop an invention or if there is a need to "alleviate health or safety concerns" (4,5).
6. The university has to give preference to small businesses when granting licenses for the use of the invention.
7. The university has to make sure that the invention will be developed primarily in the U.S. when it "grants an exclusive license."
8. The university has to share a portion of the royalties with the inventors (4–6). Hopefully this Act and its implementation will exert subtle pressure on the academic researcher to work with companies to develop products.

With the implementation of legislation such as the Bayh-Dole Act, it is not difficult to understand that the historic pressure on the academician to publish, teach, and receive grant funding has increased. Nevertheless, the fundamental differences between academe and industry create challenges for any commercial industry wishing to build on discoveries found in the research center.

Governance and Decision Making

Traditionally research institutions have benefited from corporate and personal charitable donations. These funds are now less available. The trustees of the research centers have responded by requiring management to pursue other forms of support.

To their credit, research institutions have created a variety of constructive programs aimed at addressing the concerns of industry regarding the scope and commercially useful nature of academic center research. In fact some universities have expended considerable funds to facilitate transferring technologies and programs to meet the specifications of the Bayh-Dole Act. Under these programs certain units and personnel are instructed to manage activities relating to inventions" (4).

If the approaches adopted by research centers are well managed, and carefully selected, then the research resulting from the collective endeavors of its scientists will more often lead to products, which in turn, should lead to fees and royalties to the researcher center and its scientists.

Most of these programs are designed to support the educational and basic research mission of the academic center. This is required by their nonprofit charter. The programs identify potential commercial products within the facilities, protect the inventions, and then identify and/or build companies for commercializing the discovery. Researchers are not forced into taking a role in the commercialization process. Through this mechanism academic freedoms traditionally enjoyed by the academician are maintained. It is not clear, however,

that in the long term the fundamental mission of the research center will not be subverted, a major risk for and concern of academic governing bodies.

In order to benefit from academic inventions, both academe and industry must develop sensible and affordable cooperative relationships that meet the goals of both parties. While an academician may be aware of commercial tasks required to create a product, commercially oriented research projects are not a typical part of academic research. Through education and by understanding the cultural, administrative and professional constraints on industry, the academic can contribute to the commercial success of the project. One approach which works for some corporations is to recognize the high level of skill and quality research conducted at the academic research center, and effectively integrate them into the corporate R&D process.

PROGRAMS AT UNIVERSITIES TO COMMERCIALIZE INNOVATIONS

Goals of Research Institution Technology Transfer Programs

Licensing executives at most technology transfer offices (TTOs) will state that their primary mission is to enhance the flow of innovations from the institution's laboratories to the commercial sector. This mission statement usually derives from their nonprofit charter, as well as a long history of teaching and research, not for private gain but for the greater good of the community that the research center serves. The logic is that by transferring technologies to the private sector, people will benefit from the innovations, and thus the general welfare of the state is enhanced.

This mission statement also reflects pressure from community leaders to refrain from competing with taxpayers. By conferring nonprofit status on the universities, the federal, state, and local tax codes restrict research facilities from competing with commercial entities. While there is continual and lengthy debate at the trustee level about amounts of unrelated business income derived from for-profit activities conducted under the nonprofit charter, rarely are the amounts of income and commercial activity a serious problem for the research center.

The second most often stated goal of the technology transfer program (TTP) is to make money for the university provided that the activity is consonant with the nonprofit and educational goals of the institution. So long as such is the mission of the institution, the TTP and its staff have complete freedom to conduct business as they wish.

Academic medical centers have been most often criticized for the results of their efforts to maximize health care reimbursement. In pursuit of this effort, hospital corporations restructured, developing series of for-profit and nonprofit corporations. Some of the nonprofit businesses were travel agencies, laundries, janitorial services, or power stations, often far afield from the basic mission and directly in competition with local area businesses. Community reaction and abhorrence held much of this in check. Some state and local tax authorities occasionally successfully levied property taxes on the research centers, pursuant to the logic that the research center was not benefiting the public welfare by offering such services.

The general threat of loss of nonprofit status, which rarely is a serious issue for well-planned research centers, continues to keep the research center focussed on its obligations to its community.

Other stated missions of the TTO include improving patient care by developing innovations that benefit the public, and to reward (and thus retain) talented faculty by allowing them to share in the financial benefits of fees and royalties from their discoveries. Scientists rightly observed that they could make more money by working for private companies, so universities responded by allowing them to consult, for fees and equity, and to share in royalty and fee income. While this raises questions of conflict of interest, most institutions allow and encourage such activity so long as scientists comply with institutional conflict of interest policies.

Probably the most important benefit of transferring a new technology to a local company is to increase community good will. This transfer often translates into donations and additional corporate sponsorship. Many of the research centers are also large area employers, thus further strengthening the local economy.

Effectiveness

Financial performance of the TTOs are difficult to assess. Some of the larger institutions have well-established programs and successfully generate fee income. These programs tend to have substantial revenues — much of those fees derive, however, from only several innovations which are, in turn, used to support an aggressive licensing and technology transfer operation. The surveys conducted by AUTM (Association of University Technology Managers, a nonprofit organization formed to assist university intellectual property administrators in the effective transfer of technology to the public) (4), indicate that some universities have had success with activities involving inventions and the report released by the U.S. General Accounting office on Technology Transfer: The Administration of the Bayh-Dole Act by Research Universities, further indicates that many universities believe that the Bayh-Dole Act is accomplishing its objectives (4).

Licensing Operations

Modern TTPs take a variety of forms. The most common program is a licensing operation (LO). The operational role of this operation is to arrange for intellectual property (IP) protection of discoveries made by university faculty, and to negotiate, prepare, and monitor license agreements with outside companies.

Initially such activities were part of the contracts and grants office. The general role of this office was (and is) to negotiate contracts and maintain relationships with private and government granting agencies.

As federal laws changed to allow title to discoveries to vest in research institutions, the contracts and grants

officers were additionally required to arrange for the protection of rights and to create royalty and fee income for the institution by transferring such rights to companies via license agreements. These influences and pressures required different skills and approaches, and the TTO resulted.

TTOs were initiated with high hopes and aspirations. The revenues that were expected to roll in did not. To compound the problem, decisions to patent inventions were made without sound business rational. Patent costs soared.

Additionally the skill set necessary to manage the complex licensing process typically resides at companies accustomed to licensing — there were not many seasoned business executives employed by universities in this role at that time.

It became apparent that the institution undertaking such a program required substantial institutional commitment, a sound business approach, and a long-term view to the process, at least 10 years. Some of the more fortunate institutions capitalized on early discoveries that were highly profitable, encouraging them to pursue other licenses. If these early successes are removed from their revenue streams, it is still clear that they now are investing tremendous resources on the assumption that their current licenses will be profitable. The result of this will not be known for some time, when many hundreds of licensed innovations mature. Although, some of the early licensing programs are beginning to yield substantial returns, it is not a widespread phenomenon (4).

Program Resource Requirements

Underfunding. University TTOs tend to suffer from a variety of dilemmas. Most institutions underestimate the length of time required for their patents to mature and yield returns. If they have a 10 year time horizon, they generally will be able to withstand the criticisms and pressures to perform by university administration.

Most institutions underfund and understaff their TTOs. Usually the cost of maintaining an effective program is greater than the university can afford. The result is that the licensing officers are responsible for handling hundreds of patent disclosures in a variety of technical areas, and the overall effort is diminished.

Patent Process. Because of this lack of personnel and financial resources, licensing officers tend to conduct minimal market, business, and/or technical research to validate their initial decision to patent an invention. This results in waste and inefficiency in the patenting process.

It also results in portfolio of patents that may be too limited for full commercial utility for some valuable discoveries, and an overly optimistic assessment for most of the inventions. The most sensible approach is to invest the requisite time and effort to conduct a business and technical assessment of the invention prior to embarking on the patent process.

Technology Audits. Institutions should take stock of invention inventory through a thorough technology audit. Each invention is assessed for its market potential, feasibility, time lag to product, and regulatory and financial requirements. The inventions are then ranked according to institutional priorities, and resource allocation requirements. They may also be ranked according to their commercial potential.

This invention assessment allows the institution to make sound business judgments about funding, its level of interest in seeing the products reach the marketplace, and its appetite for pursuing licensing or new enterprise development.

Any subsequent financial decisions with respect to the inventions are weighed against other institutional commitments, and its priority ranking. This is especially valuable in assessing whether to proceed with the patent process, and how seriously the effort should be pursued. Subsequent marketing efforts are also measured against their priority on the list. Those below a certain level, weighed in light of other resource demands, can be returned to the inventor. In this way, the most technologies considered most valuable to the institution are properly and rationally protected.

Marketing Effort. The scope of some discoveries justifies new company formation, with the requisite involvement of capable management and proper funding. For a variety of reasons, however, most successful technologies, perhaps more than 95 percent, are suitable only to be out-licensed. Selecting the proper licensees, unfortunately, requires a considerable amount of time and effort. This effort, due to funding, time, and personnel constraints, is generally beyond the scope of most TTOs. The result then is predictable: The technology licensee is usually the company, any company, that first makes an offer, any offer, to the university.

The more sensible approach is to develop a marketing plan for the technology, beginning with the data that resulted from the technology audit. This information needs to be supplemented with fresh technical, market, business, and regulatory analyses and summarized in both nonconfidential and confidential disclosures. This effort aids the TTO staff in its search for appropriate licensees, which are targeted in a defined marketing program. This defined marketing effort is calculated, if successful, to result in serious and appropriate partners, and new corporate relationships for the institution. Even if the technology is not purchased by the potential licensees, a well-reasoned and sensible approach to a serious company will create a favorable impression. This may well lead to subsequent opportunities — it definitely improves the likelihood that a prospective licensee will take the TTO seriously.

Drafting and Maintaining Licenses. Once a licensee has evaluated the technology, and wishes to enter into a license agreement to acquire the rights to the technology, the license negotiation process is initiated. The drafting and negotiation process is rarely routine. Many treatises have been written, and numerous license agreement forms have been generated, all of which serve as useful tools for the experienced and inexperienced TTO personnel.

Licensing Check List. A typical license agreement should address at least the following:

- Exclusive versus nonexclusive
- Restrictions in freedom to License and Sublicense
- Contamination — e.g., commingled funding by federal and/or other industrial money
- Collateral agreements may restrict rights
- Confidentiality agreements — any confidential information used from any/other restricted sources?
- Any nondisclosure agreements that encumber the invention?
- Collaborations with other scientists on site —
  - any visitors that were employees of other institutions that signed institutional
  - patent agreements
  - off-site collaborations
  - any scientists who have left.
- Financing/Royalty Clause Consideration
- License issue fee/varies —
  - fully paid up/lump sum
  - running royalty, royalty cap, or minimum/maximum
  - equity in lieu of or addition to royalties
  - credits given against royalties for earlier expenditures
  - reimburse for research expenses
- Milestones/upfront payments/termination fees
- Background rights —
  - define patent/other intellectual property with care
  - follow on patents, improvement patents
  - improvements
  - subservient to basic patents
  - retain option to these background rights
  - negotiate a separate royalty rate, short time
  - window for improvements (6–18 months)
  - subject to the rights of other parties
  - define field of use
- Know-How —
  - carefully define and transfer it
  - its delivery is very amorphous — if well defined, then know-how transfer is easier
  - treat via field of use
  - typically nonexclusive
  - if exclusive, only if it can be protected, and only if it is no longer needed
  - retain the rights to use know-how
  - consider using separate agreement
- Infringement and patent protection
- Indemnity
- Disclaim warranty provisions
- Territory Infringement
- Field of use
- March in rights and due diligence
- Use periodic payments, with reversion of rights
- Automatically for nonperformance
- If patents are assigned, pre-execute grant back of patent rights subject to conditions

Once a licensing relationship is in place, it must be regularly and objectively reviewed. Maintaining and monitoring licenses, communicating with licensees, and enforcing the terms of the agreement are time-consuming, and must be planned for in advance, and in light of their ranking in the priority list. Those agreements that have gone awry must be analyzed as to the reasons for their failure, utilizing in-house and independent peer review personnel. Nonperforming or minimally productive licenses should be terminated.

If the licensee does not proceed in a timely and businesslike manner, the technology should revert to the institution, pursuant to defined and clear terms that do not hamper the licensee's ability to raise funding or pursue the technology. Most often, if the company is not willing to commit to financial benchmarks, it is the wrong licensee.

Licensing Agreement Controls. Proper controls to ensure quality license agreements must be in place. Standardized forms are usually a useful starting point, and should be used where possible. Associates preparing the licenses must be subject to performance reviews. In order for the process to result in an appropriate business result, the expectations and demands on both the licensee and licensor must be adhered to, requiring the licensing office to ensure that the institutional commitments are met and that the office is operating in a businesslike fashion.

Patent Decisions. With respect to the decision whether to apply for a patent, claiming certain inventions, the typical licensing officers do not conduct technical analyses of the technology, relying instead on the inventor to do so. With some unusual inventors, this might be adequate provided that they know the industrial side of their research. Most, however, do not have the appropriate skill set. A careful search for and review of relevant technical literature by a scientist other than the inventor may reveal very sound reasons why the patent expenditure does not make any sense. For example, if a use patent is filed on a known compound, but the literature reveals that the use applied for may give rise to an adverse reaction that conclusively eliminates its commercial value, then the licensing office should not expend its resources to patent the invention. Unfortunately, most of these important pieces of information appear when the USPTO examines the patent application, or, more often, when a licensee conducts its due diligence, and it becomes obvious that the money spent on the patent was wasted. It also dilutes the effectiveness of the office staff, who could better use their time on other licensing projects.

If the TTO's decision is not to file the patent application, the university should relinquish its rights to the invention and return such rights to the inventor. It should do so as soon as practicable. It should also describe why it has

declined the opportunity to file a patent application and give the details of its business rational.

A corollary to the technical literature search is a careful analysis of the prior art in the patent literature. While this usually is done well by competent patent counsel, the licensing officer can glean valuable information about competing patent estates that impinge on the value of the technology. Learning to operate the computer search software programs is easy; learning to use them well is very, very difficult, and is best done by someone who is in a dedicated service role to the licensing associate.

Another problem is created when the licensing operation fails to plan for the cost of enforcing patents. Usually the decision to prosecute an infringer is made in a hurried manner, without fully appreciating the potential costs and benefits. Prosecuting infringers is a very expensive proposition.

Other Agreements. A typical hazard for the licensee and licensor are "hidden documents" that appear at the last minute, or are discovered very late in the patenting, licensing, or due diligence process. These hidden documents may contain restrictions that limit the university's ability to grant free title to the licensee. Consulting agreements and material transfer agreements may have been executed without appropriate terms and conditions. Federal funding, commingled funding, university collaborations, whether formal or informal, nonconfidential and confidential disclosure agreements, and joint inventorship of patents, each raise case-specific problems that affect the value of the technology. Analyzing these issues and optimizing all aspects of ownership is most effective when done before the TTO approaches or is approached by a potential licensee.

Faculty Communication. The licensing officers must have early and frequent communication with their inventors. By involving them in the decision-making process, the faculty are recruited into the process, learn about the rationale for decisions made, and become a valuable resource. If decisions are made without their involvement, the faculty will be alienated, and resentful of business and administrative decisions, which at many points will detract from the office's effectiveness.

CONFLICT OF INTEREST DILEMMAS IN BIOMEDICAL RESEARCH

Industry–academe collaborations have costs. These collaborations create conflicts among the researchers, the institutions, industry, and the researcher's academic and financial interests. These conflicts of interest threaten the objectivity of science, the integrity of scientists and institutions, and the safety of medical products.

One such example of a conflict of interest between the researcher, the institution, and the industry can be observed from the Synthroid Marketing Litigation case (7). In this case researcher Dr. Betty Dong "discovered" that there were less expensive alternatives to the Synthroid medication (8). Dr. Dong's research was funded by the University of California, San Francisco, and Knoll Pharmaceutical Company, which manufacturers a thyroid medication called Synthroid. When Dr. Dong attempted to publish the results of her study, the representatives of Knoll informed her that she was barred from doing so because she had signed a contract agreeing to publish her results subject to the approval of the company when she initially began the project (9).

Claiming the right to academic freedom, Dr. Dong took the Knoll Pharmaceutical Company to court (10). This case illustrates a wide range of conflicts. Was it permissible for the company to suppress Dr. Dong's study for its own reasons? Was it permissible for Dr. Dong as a university employee and beneficiary of funding from Knoll to comply with the suppression of her article for seven years? If Dr. Dong were to be penalized for breaching her contract with the company, would the university be liable as well, or would she be treated as an independent contractor? Hopefully these questions will be answered once the case has been resolved (11).

To avoid conflicting interests that undermine scientist's integrity such as was the case with the Dong study, adherence to uniform federal standards should be mandatory. Federal rules are necessary to require disclosure of conflicts, limit the most troublesome forms of conflict, and create uniformity in ethical standards across the country.

Federal Legislation Supporting Industry–University Collaborations

Federal legislation, since 1980, has facilitated industry–university collaborations, and speeded promising new products from the laboratory to the market. This legislation, however, has brought academe and industry together without adequately regulating the consequences of the interactions.

The Stevenson-Wydler Technology Innovation Act of 1980 (12) established a policy of "stimulating improved utilization of federally funded technology developments by state and local governments and the private sector" (13). The Act created an Office of Research and Technology Applications, whose primary purpose was to investigate projects that could be utilized by government or private industry (14). Each Agency implemented this requirement by having it own version of such Office (i.e., the Office of Technology Transfer at the National Institutes of Health, NIH, handles this function for the Department of Health and Human Services, DHHS). The Act also created the Center for the Utilization of Federal Technology, established within the National Technical Information Services, to provide industry with a central source of information on federally owned or developed technologies with potential commercial application (15).

Later in 1980, Congress accelerated technology transfers by amending the patent and trademark laws, and for the purpose of supporting small business. As mentioned previously, the Bayh-Dole Act (16) gave inventors in small business firms and nonprofit organizations the power to retain ownership rights to patents protecting inventions developed with federal funding (17) (Licenses are intended to be granted to small businesses in the United States. If this is not possible, the licensing institution is to use its

best efforts to ensure that the licensee, whether foreign or U.S. based, manufactures the product in the United States, and for consumption in the United States). Prior to the Bayh-Dole Act, the federal government owned the rights to most federally-supported inventions and for-profit firms wishing to develop federally supported inventions had to wade through a bureaucratic maze to obtain a license (18). Congress recognized that the federal government had been unsuccessful in nurturing the development of new products to the market, and that it was in the public interest to bring innovative ideas into clinical practice without unnecessary delays (19). A policy statement by the Reagan Administration extended the Bayh-Dole coverage beyond small business firms and nonprofit organizations (20). The Bayh-Dole Act enabled institutions and their investigators to license patents from federally-supported work to companies interested in developing the products for market. This has allowed the institutes and researchers easier access to money, both for research and for personal profits, and opened up to industry a large market for potential commercial advantage (21).

Liability Risks of Research Institutions and Investigators

Despite the progress of the current federal legislation supporting industry–university collaborations, the collaborators still have to contend with the potential for a conflict of interest leading to legal liability. This potential is derived from several sources. First, state tort law imposes liability on institutions for the misconduct of their employees. A research institution could be held liable for the negligence, misrepresentation, or fraud of its investigator or employee. For example, if a researcher misrepresents the quality of an invention that is commercialized, and the company relies on false claims, the university could be held responsible. Considerable damages may be assessed resulting from the delay in marketing the product and/or in wasted or misdirected investments.

Second, research institutions have various obligations under state nonprofit corporation laws and federal tax laws, any breach of which could jeopardize the institution's nonprofit status, or subject the institution to enforcement actions by the state attorney general. Typically these laws require the directors to operate the research institution in a manner consistent with its charitable purpose; forbid certain director conflicts, interlocking or interested director transactions; and require the directors to preserve and prudently invest corporate assets. Although research institutions need to closely adhere to the current regulations, the court has recently given them some flexibility with respect to managing an invention the has been assigned to the institution by an employee. In *Kucharczyk v. Regents of the University of California* (22) addresses the ability of inventors to influence negotiations for the sale of their inventions.

The Regents had negotiated a licensing agreement with Nycomed for the use of a patented medical technique developed by Dr. John Kucharczyk and Dr. Michael Moseley (22). The doctors assigned their rights to the patent to the Regents which then sold these rights to Nycomed for $25,000. Fifty percent of the sale price went to inventors, Drs. Kucharczyk and Moseley in compliance with the Bayh-Dole Act (4,22). The doctors then filed suit against the Regents of the University and Nycomed alleging that the defendants "acted improperly to deprive plaintiffs of their rightful share of the financial rewards of the patented medical technique they developed" (4). In their lawsuit, the plaintiffs claimed that the medical technique invented by them was "worth substantially more than $25,000."

The court ruled that the actions of the University of California and Nycomed did not constitute a breach of contract since the University doctors had contracted out their right to sue when they assigned their inventors rights to the University. The court further ruled that the plaintiffs might have a claim against Nycomed for fraud and interference with contractual relations, because there was sufficient evidence of Nycomed's suggestion to the University doctors that their medical technique had substantial profit-making capacity, and that this information was omitted from the negotiations with the University.

CONCLUSION

With increasing sophistication and skill, technology transfer offices are successfully commercializing useful inventions and generating fee income. The recent court decisions of *Kucharczyk v. Regents* and *Synthroid Marketing Litigation* threaten the delicate balance within and bring unpredictability to the contracting process. Many challenges lie ahead, including conflicting university, faculty, and societal interests; competition; change in structure and focus of the university; congressional intervention; and availability of personnel, among other factors. An enlightened technology transfer office will move past these "speedbumps" and continue down this creative and energetic road.

BIBLIOGRAPHY

1. 35 U.S.C. §§200; 15 U.S.C. §§3701.
2. *Kucharczyk v. Regents of the University of California*, No. C 94-3886, 1999 U.S. Dist. LEXIS 6905 (N.D. Cal. May 6, 1999).
3. J. Tesk, *J. NIH Res.* 3-13–3-14 (1991).
4. *Technology Transfer: Administration of the Bayh-Dole Act by Research Universities* (GAO/RCED-98-126), May 7, 1998.
5. 37 C.F.R. §401.6 (1999).
6. 35 U.S.C.S. §§200–212 (1999).
7. *In re Synthroid Marketing Litigation*, No. 97 C 6017, 1999 U.S. Dist. LEXIS 11195 (N.D.Il. July 19, 1999).
8. J. Kopito, *Synthroid Litigation* (November 30, 1997), Available at: *http://www.medicinegarden.com/archive/00004309.htm*
9. K. Davidson, *USCF Angers Faculty by Quashing Research; Drug Maker Funded Study, Then Exercised Right to Suppress It* (April 26, 1996).
10. L. Marsa, *The Rising Health Costs of Capitalism's Invasion of the Science Lab* (December 20, 1998).
11. Mark Patterson, *Conflicts of Interest in Scientific Expert Testimony*, 40 Wm. and Mary L. Rev 1313, 1346–1350 (1999). *Synthroid Marketing Litigation*, Available at: *www.synthroidsclaims.com/*

12. Stevenson-Wydler Technology Innovation Act, Pub. L. No. 96-480, 94 Stat. 2311 (1980) (codified at 15 U.S.C. §§3701).
13. Pub. L. No. 96-480, 3 (1980) (codified at 15 U.S.C. §3702(3)).
14. Pub. L. No. 96-480, 11 (1980) (codified at 15 U.S.C. §3710(b) & (c)).
15. Pub. L. No. 96-480, 11 (1980) (codified at 15 U.S.C. §3710(d)). (Now operated by the National Technical Information Service, 15 U.S.C.S. §3710(d) [(Deering Supp. 1991)].
16. Pub. L. No. 96-517 (codified at 35 U.S.C. §§200).
17. 35 U.S.C.S. §202(a) (Deering Supp. 1991).
18. E.C. Walterscheid, *Harvard J. Law Tech.* **103**(3), 131–134 (1990).
19. S. Rep. No. 480, 96th Cong., 1st Sess. 19 (1979).
20. Presidential Memorandum to the Heads of Executive Departments and Agencies, Subject: Government Patent Policy. 1983 Pub. Papers 248; Feb. 18, 1983 and Exec. Order No. 12,591, 35 C.F.R. §200 (1987), reprinted in 15 U.S.C.A. §3710, app. at 256-58 (West Supp. 1992).
21. Brown, *Business Goes to College, Forbes* 196, October 11 (1982).
22. 1999 U.S. Dist. Lexis 6905.

See other entries MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT; UNIVERSITY-INDUSTRY RESEARCH RELATIONSHIPS, ETHICS, CONFLICT OF INTEREST.

TRANSGENIC ANIMALS: AN OVERVIEW

RIVERS SINGLETON, JR.
University of Delaware
Newark, Delaware

OUTLINE

INTRODUCTION

To present an overview of animal use in modern biotechnology is a difficult task because of both the multitude of procedures and the types of animals involved. Procedures vary from small-scale laboratory research programs, involving a range of different animals from across the evolutionary spectrum to biotechnology product production facilities that may involve herds or colonies of a single animal species. Early biotechnology was restricted to procaryotic organisms, such as insulin production by strains of *Escherichia coli* constructed by recombinant DNA techniques. Since those early 1970s protocols, however, the diversity and complexity of organisms used in biotechnology now varies greatly across taxonomic lines from relatively simple procaryotes to nematodes to fish to mammals. Furthermore, since a major goal of biotechnology is to create organisms possessing genetic properties of other organisms, a pivotal result of these procedures is that taxonomic lines can be blurred as genetic elements of one organism are introduced into another.

The recent development of facile techniques to clone mammals further complicates the view of animals in nature and their role in biotechnology. Previous technologies created chimeric (An organism consisting of two or more tissues of different genetic composition, produced as a result of mutation, grafting, genetic engineering, or the mixture of cell populations from different zygotes.) organisms by combining genetic attributes from differing taxonomic lines. Because cloning bypasses genetic recombination, the technology leads to monophyletic (Of or concerning a single taxon of animals, relating to, descended from, or derived from one stock or source.) organisms, which may have highly unusual genetic traits, that are essentially unique in the biological world.

The point of the previous paragraphs is that both animals and procedures gathered under the rubric of "biotechnology" vary greatly. Because of this complexity, the focus of this article, in large measure, is on transgenic organisms as a model for some of the ways biotechnology uses animals. Thus a goal is not to provide a comprehensive overview of all animal use but rather to provide a paradigm by which the reader may gain insight into other arenas of biotechnology animal use. Some related issues have been discussed elsewhere (1,2).

HUMAN AND NON-HUMAN ANIMALS: SOME HISTORICAL REFLECTIONS

It is a truism to state that humans and animals have always interacted with each other. All living plants and animals are linked together through a shared evolutionary origin. While we are all part of nature's evolutionary web, however, our interactions with the nonhuman world are unique in nature. As human beings have domesticated the world around us, we have interacted with a diverse array of organisms across broad taxonomic lines in pervasive and complex ways. Indeed, our domestication of nature for utilitarian purposes appears a uniquely human activity. Consider, for example, that wheat domestication

is viewed as the hallmark of human civilization, yet humans have arguably domesticated dogs for a longer time (3–7). [For a nontechnical but very readable account of canine evolution, see Budiansky (8).] Since those initial domestication forays at our emergence as modern human species, we have radically altered the permanent form and function, through selective breeding, of an immense variety of plants and animals.

Domestication has been driven primarily for human utilitarian benefit, although many organisms have undoubtably benefited from this close human association. While dogs frequently serve as human companions and pets, historically they have also been invaluable for herding sheep, pulling sleds, and other utilitarian purposes. For millennia an immense variety of animals served as both food supply and labor savers (it is not coincidental that a unit of work expenditure is "horse power") for human beings.

Animals have played another vitally important role in human culture. By studying animals, humankind gained an immense knowledge about the natural world. In antiquity Galen speculated about human physiology (often erroneously) and based his observations on animal dissection and vivisection. In the sixteenth century, William Harvey's brilliant description of pulmonary, cardiac, and circulatory physiology was deeply rooted in a variety of animal observations and experiments. As scientific knowledge exploded during the next four centuries, animal study played an important role in that expansion of human knowledge. Since the Renaissance vivisection and animal experimentation have increased human understanding of both basic physiology and pathological disease processes. It is reasonable to conclude that many advances of modern medicine would have been impossible without animal use. Indeed, we can conclude that much of our knowledge of fundamental biology would not have been achieved without recourse to animal experimentation and vivisection (9,10).

The historical streams of both animal domestication and experimentation are important to understand the place of animals in biotechnology. A definition of technology is "the application of scientific discoveries to the production of goods and services that improve the human environment" (11). Biotechnology thus is simply using biological systems, biological processes, or exploiting living organisms as part of the process of producing "goods and services that improve the human environment." The National Agricultural Library defines biotechnology as:

> . . . a set of powerful tools that employ *living organisms* (or parts of organisms) to make or modify products, improve plants or "animals," or develop microorganisms for specific uses. Examples of the "new biotechnology" include the industrial use of recombinant DNA, cell fusion, novel bioprocessing techniques, and bioremediation (12; p. 1 emphasis added).

Thus animals — *living organisms* — are central to biotechnology in several vital aspects. The basic science upon which biotechnology is structured would not exist in the absence of animal experimentation and vivisection. Selective breeding continues to produce a variety of animals, ranging from shrimp to cattle, of great commercial importance. Modern techniques of molecular biology, such as marker-assisted selection, have enhanced domestic breeding programs to make them more effective and efficient.

Equally important, however, is the notion that — for many people — animal use in biotechnology is a simple extrapolation of human domestication of nature. Domestic breeding can enhance only genetic traits that naturally occur in an organism. Creation and use of transgenic organisms can "leap-frog" these genetic limitations, however, and introduce traits not normally found in particular species. Consequently transgenic animals represent and illustrate notions of domestication and experimentation as well as the general utility of animals in biotechnology.

BIOTECHNOLOGY AND DOMESTIC BREEDING

Traditional Breeding Perspectives

As noted previously, domestication of plants and animals for human utility and companionship has been a characteristic of human beings since our early evolution as a species. The power of selective breeding to bring about radical and relatively stable alterations in the form and function of animals was so well recognized by the nineteenth century that Charles Darwin devoted the first chapter of *Origin of Species* to the subject as a model for natural selection. Modern concepts of molecular biology, such as marker-assisted selection, combined with traditional selective breeding practices have greatly enhanced the power of domestic selection.

In traditional breeding practices, as Darwin noted, a breeder identifies a desirable physical trait in an individual organism within a population, such as increased milk production in a dairy cow. The exemplary animal is then used for breeding purposes. Those progeny exhibiting the desired trait are in turn used as further breeding stock, yielding — after a period of several generations — a population of animals that expresses the desired trait, namely a herd of cows with increased milk production.

Marker-Assisted Selection

A variety of new molecular methods — such as restriction fragment length polymorphism (RFLP) and polymerase chain reaction (PCR) — now allow breeders to identify DNA sequences associated with specific animals. These sequences may or may not be responsible for the particular trait the breeder desires to enhance. For breeding purposes, importance rests in expression of the genotypic "marker sequence" in progeny carrying the desired phenotypic trait; for example, in the case of increased milk capacity, a "marker sequence" of DNA should always be found in progeny expressing increased milk production.

Although the notion rests on numerous molecular methods for cutting, isolating, and analyzing specific segments of DNA, which have evolved over the past several decades, the concept of marker-assisted selection appears to have achieved practical application during the 1990s. For example, a Medline© database shows

initial papers, specifically mentioning marker-assisted selection, beginning to appear around the early to mid-1990s. Since then, the number of papers has increased dramatically. Citation analysis also suggests a curious bias in the ways marker-assisted selection is used. A Medline© database search of 80 papers published since 1994 showed approximately 60 percent of the publications dealing with animals. However, another search of 250 publications in several agricultural databases showed only 20 percent dealing with animals. While both searches are anecdotal in nature, they suggest that the concept of marker-assisted selection is receiving wide application in enhancing agricultural plants.

Currently there is an international effort underway to generate a detailed description of the molecular genomic structures in a diverse array of organisms; the Human Genome Project is one part of this effort. As we increase our understanding of these genomic sequence details, and the biological function of particular genetic sequences is clarified, the power of marker-assisted selection will be greatly accelerated and enhanced. Nevertheless, despite advantages introduced by molecular techniques, selection-based breeding programs will always be limited to genetic attributes inherent within a species. There is, for example, a statistical distribution of milk production within cows; all that domestic selection can do is to skew that distribution toward a desired goal. To breach that genetic restriction, the new techniques of transgenic organisms must be employed.

BIOTECHNOLOGY AND TRANSGENIC ORGANISMS

What is a Transgenic Animal?

A simple definition of a transgenic animal is one "to which copies of a gene sequence have been artificially added" (13). However, there is difficulty in such definitions. The *Hastings Center Special Supplement* on animal biotechnology concluded that a transgenic organism "carries and expresses genetic information not normally found in that species of organism" but also noted that such a definition was literal and restrictive. Thus the definition was broadened "to include the purposeful amplification, spread, or dissemination of a *gene* within a *species* at a rate much faster than would have occurred in the absence of artificial interventions" (1; emphasis added). A view of animals possessing *desired* genetic properties, which in turn express novel phenotypes, emerges from this definition. These animals have been intentionally *designed* using modern biotechnological tools.

Some of these terms need further clarification. Our notion of a *gene* emerges from the central dogma of molecular biology:

$$\text{DNA} \Rightarrow \text{RNA} \Rightarrow \text{protein}$$

In traditional biological terms, DNA represents an organism's "genotype," and protein expresses its "phenotype." We can consider a gene as a section of a DNA molecule that provides biological information and is ultimately transcribed and translated into a protein molecule; proteins, in turn, perform various cellular activities. Changes in DNA molecular structure will alter cellular function because of the resulting changed protein. Thus a transgenic animal carries a novel sequence of DNA [referred to as the *transgene* (13)]. If the transgene is stably incorporated into the animal's chromosomal DNA and its products functionally expressed, the animal — *and its progeny* — will possess an altered phenotype.

Species is more difficult to define, and Ernst Mayr (14) noted that biologists have understood the term in at least three different ways. Historically biologists viewed species from an *essentialist* perspective, which held that "each species is characterized by its unchanging essence and separated from all others by a sharp discontinuity (14, p. 256)." Charles Darwin helped change that view to a more *nominalistic* concept that rejected notions of "essential character" and conceived of species as groups of organisms that shared common attributes with a common descriptive name. Finally, while some biologists might argue with it, the modern notion of species, namely "a reproductive community of populations (reproductively isolated from each other) that occupies a specific niche in nature" (14, p. 273), is acceptable by most of the biological science community.

How are Transgenic Animals Created?

How do we go about this process of "purposefully amplifying, spreading, and disseminating" a gene within a species? Before describing transgenic technology, two brief reflections are important. First, transgenic technology with eucaryotic (also eucaryote, A single-celled or multicellular organism whose cells contain a distinct membrane-bound nucleus.) organisms is a logical and conceptual extrapolation of the recombinant DNA technologies with procaryotes in the 1960s. Intentionally creating a transgenic organism in a laboratory or factory is deeply rooted in the recombinant DNA work of the 1960s (1,2,13). Paul Berg's colleagues created an early transgenic organism when they used restriction enzymes and plasmid vectors to insert genetic elements from Simian Virus 40 into *Escherichia coli*. A significant difference between modern technology and these early techniques is that the latter were unidirectional. One could only introduce genetic material from a foreign source into bacterial (procaryotic) systems. Modern technology allows the manipulation of genetic information between virtually any plant or animal.

Second, clarification of two experimental distinctions is important. A *knockout experiment* is one that creates a mutation in an organism's own genome. Some genetic element native to an organism is inactivated so that the resulting progeny lack the functional capability associated with that gene or genes. In true *transgenic experiments* novel genetic elements, not normally found in that organism, are inserted into an organism's native genome. Thus this type of experiment creates an organism that "carries and expresses genetic information not normally found in that species of organism." Despite the different outcomes, both experiments use similar technological approaches.

A somewhat typical knockout experiment, which involved creation of a mouse with a defective *fosB*

mutation (15), illustrates the technology. The knockout experiment discussed here illustrates some of the general technology used to create transgenic organisms. The specific experiment is also important because of the unexpected outcome (see below), which illustrates the serendipitous nature of science. Some background on the *fosB* gene and its protein product is important to understand the experiment and its somewhat unusual results. The FosB protein is one of many transcription factors found in cells; they facilitate the phrase in the central dogma of molecular biology:

$$\text{DNA} \Rightarrow \text{RNA}$$

The *fos* genes, which produce these proteins, are activated during a variety of adaptive neuronal responses in several brain regions. Despite extensive work that correlated *fos* gene products with mRNA production, their role in nervous system function and development remains unclear (16). Thus Brown et al. decided to create a knockout mouse, which lacked the *fosB* genes, in order to gain insight into its regulatory function; their experimental protocol for creating this *fosB* gene "knockout" mutation is summarized in Figure 1.

Initially murine DNA containing an incomplete fragment (and therefore nonfunctional) of the *fosB* gene was isolated and incorporated into a vector, which permitted three important experimental tasks. First, unique information present on the vector allowed investigators to screen transformed embryonic cells for the presence of the mutant gene. Second, specific sites on the vector, which were easily recognized at a molecular level, facilitated sequencing the incorporated genetic material. Finally, when genetic elements on the vector were phenotypically expressed, they served as a type of "marker-assisted selection" (although not referred to as such). Individual offspring expressing these traits were easily recognized as carrying the vector, and selecting them for further breeding purposes was eased. Once constructed, the vector containing the mutant *fosB* gene was electroporated into embryonic mouse cells.

Transformed embryonic mouse cells, namely those containing the mutant *fosB* gene, were implanted into mouse blastocysts. The blastocysts were then implanted into pseudopregnant female mice, which gave birth to pups expressing various levels of *fosB*. Ultimately the authors derived three strains of mice that exhibited normal Mendelian inheritance of the *fosB* mutation: one group was homozygous for the mutation [*fosB* (−/−)], one group was

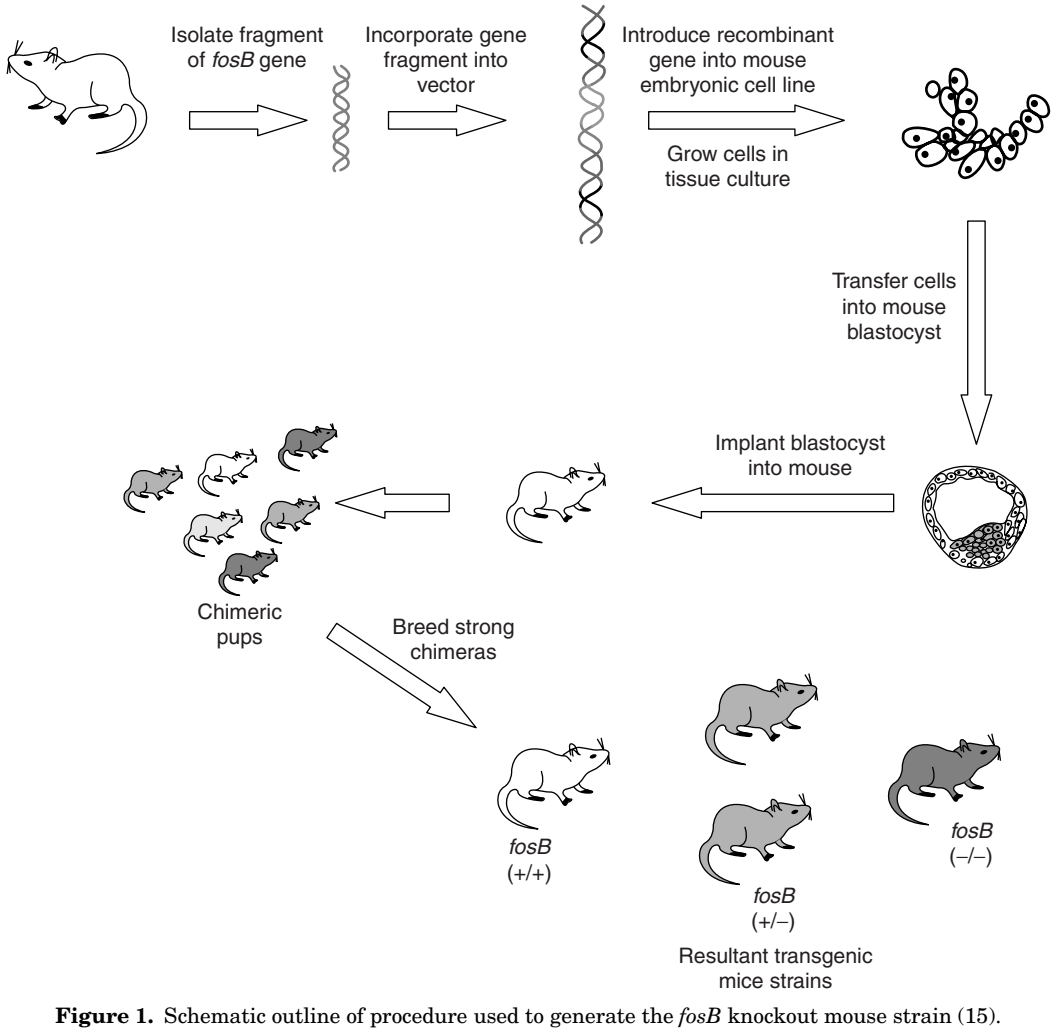

Figure 1. Schematic outline of procedure used to generate the *fosB* knockout mouse strain (15).

heterozygous for the mutation [*fosB* (+/−)], and a third group was homozygous for the wild-type state [*fosB* (+/+)].

These mice had interesting phenotypes: (*1*) the homozygous knockout mutants [*fosB* (−/−)] were health and viable; (*2*) there were no apparent histologic abnormalities, suggesting that *fosB* products are not required for normal mouse development; (*3*) the mutant strain was about 10 percent smaller than wild-type mice; and (*4*) pregnancies are normal and carried to full term. However, one phenotypic difference in the *fosB* mutant was unexpected and very significant. Pup lethality in early postnatal period was eight to ninefold higher in *fosB* knockout mutants than in wild-type mice. Serendipitously the authors discovered that lethality did not arise from a pup defect but rather arose from a failure of mothers to nurture pups. A graduate student involved in the project was concerned about put lethality. She returned to the lab one evening and discovered that the mother was in one corner of the cage and the pups isolated in another corner. When she moved the mother to the pups, they nursed normally (16). Mutant strain mothers simply abandoned the pups and failed to nurse them properly. Furthermore this was a loss of maternal "nurturing behavior" and was not due to a lactation defect in mothers.

The major significant difference in the knockout experiment discussed here and creation of a transgenic animal is the source of donor DNA. In the above knockout experiment, mouse DNA was incorporated into another mouse. In a transgenic experiment, DNA from *any* source (plant or animal) is incorporated into an organism unrelated to the source. The goal in discussing this experiment is to illustrate the relative ease of transgenic technology. Thus discussion of the ethical implications of the experiment will be deferred. However, two additional reflections are important to remember about this brief sketch. First, the technique and many potential technical complications have been greatly simplified or ignored. Furthermore there are a variety of other technical approaches to create transgenic organisms; this experiment simply serves to illustrate the basic technological concepts. Fundamentally, however, the techniques are relatively easy to perform, and new technological innovation are constantly appearing.

UTILITY OF TRANSGENIC ORGANISMS

Given this relative ease of the technology, what can it be used for? Like the recombinant DNA technology that preceded it, potential applications of transgenic technology is probably limited only by the human imagination (17). The various technological innovations that allowed development of transgenic animals have great potential utility for humankind (18,19). While there is tremendous overlap in various applications, these applications generally fall into three broad areas: commercial utility, improving human health, and furthering basic biological scientific knowledge.

Commercial Applications

We should not forget that the first type of transgenic organisms, that is, organisms intentionally engineered for specific utilitarian applications, were recombinant bacteria. Since the 1960s procaryotic organisms have been designed to carry out numerous industrial and agricultural processes, including (but not limited to) frost control on plants, increased efficiency of nitrogen fixation in soil bacteria, biodegradation and waste treatment, mineral processing, and other processes (20).

Like their procaryotic predecessors, transgenic animals can be used to produce, or are themselves, new commercial products. Utilitarian applications of transgenic organisms promise to provide humans with a variety of new capabilities not readily available by other means. The actual or potential list of commercial applications is immense, a potential briefly suggested by the following examples.

Agricultural Utility. Many early eucaryotic transgenic experiments sought agricultural benefit, and transgenic farm animals, with a variety of potential commercial applications, were created containing various growth hormones. Pigs that expressed enhanced levels of growth hormone exhibited significant daily weight gain, increased efficiency of food utilization, and a decrease in subcutaneous fat (21). While these animals had enhanced growth and improved feed efficiency, their immediate commercial utility was not feasible because of serious health complications including abnormal bone growth, enlarged internal organs, and diabetes.

Despite early enthusiasm over application of transgenic technologies to agricultural animals, that expectation has been tempered by experimental reality. In comparison with laboratory models, production of transgenic livestock is an inefficient process, and research has been hampered by lack of embryonic stem cell lines in farm animal species (22). Ward and Nancarrow (23) noted that, while it was feasible to produce transgenic animals with enhanced agricultural utility by modifying growth hormone levels, this could be done only if ways were found to tightly regulate hormone production. Nevertheless, optimism continues in areas ranging from mariculture (Cultivation of marine organisms in their natural habitats, usually for commercial purposes.) (24) to more conventional livestock (25,26). In the later area, confidence in transgenic technology is such that Murray stated: "Recent advances suggest that within the first decade of the 21st century the first transgenic animals will become available to the livestock industry, with acceptance depending upon their cost versus their potential economic benefit to the producers" (26, p. 149).

Pharmaceutical Utility. Transgenic animals have been effectively and efficiently used to produce many protein pharmaceutical products that are difficult if not impossible to produce by other means (e.g., traditional recombinant DNA techniques). As Wall et al. noted: "The objective of the emerging gene 'pharming' industry is to produce pharmaceuticals for treating human diseases. It is argued that mammary glands are an ideal site for producing complex bioactive proteins that can be cost effectively harvested and purified" (27, p. 2213).

There are two general obstacles to synthesizing these products via many traditional approaches. First, procaryotes (e.g., bacteria *E. coli*) lack the post-translational processing machinery necessary to synthesize many biologically active proteins from eucaryotic organisms. A second major difficulty arises from the difference in genetic organization between procaryotes, where genes are arranged on a single chromosome without interruption, and eucaryotes. Eucaryotic genes are often arranged in fragments so that mRNAs must be joined together before a biologically active protein can be expressed.

Large transgenic animals have been developed and successfully exploited to bypass these difficulties. Such animals are often referred to as "bioreactors" and the process called molecular or gene farming (or "pharming"). Approximately a dozen biotechnology companies now produce a variety of pharmaceuticals in larger amounts this way than can be achieved via other approaches. In addition to obviating the technical difficulties discussed previously, there are distinct advantages to using transgenic large animals to produce valuable human pharmaceutical proteins (27,28).

One advantage of using transgenic animals to produce pharmaceutical proteins is low operating cost, once the transgenic animal strain has been constructed. Furthermore, once the strain has been established, a virtually unlimited bioreactor supply becomes available by way of embryo cloning techniques (29).

However, the major advantage of using transgenic mammals in pharmaceutical protein production is that these genes can be inserted into mammary gland gene control elements so that the transgene product is expressed in milk. Complex pharmaceutical proteins, with correct post-translational modifications and full bioactivity, are correctly expressed and secreted in large amounts (in the order of grams per liter) in the milk. Since a large mammal, such as a cow, can produce 10,000 liters of milk a year, kilogram quantities of pharmaceutical protein can be synthesized per animal annually. As Smith commented, "No other production system can compete with bioreactors in production levels" (28, p. 681).

The technology now uses transgenic sheep, goats, pigs, and cattle to provide a ready supply of previously rare pharmaceutical proteins, such as Alpha 1 Antitrypsin (α-1-AT), Factor IX, and tissue plasminogen activator (t-PA). The later proteins are important therapeutic agents for a variety of human clotting disorders. Estimates of the U.S. market for pharmaceuticals transgenically produced approach $3 billion annually, and several products produced in this way are now in human clinical trials or actual therapeutic use (27).

This technological approach to pharmaceutical production is not without potential health hazard for the animal. Most of the hazards are recognizable, however, and appropriate design considerations can be developed to avoid problems (30).

Applications for Improved Human Health

Second, and of equal importance to their commercial utility, transgenic animals represent a variety of actual or potential improvements for generalized human health and welfare. Of major importance, these techniques allow researchers to develop laboratory organisms that mimic or duplicate many human diseases. Nomura noted that transgenic animals can be valuable models to follow the sequelae and treatment possibilities for these diseases and proposed criteria to evaluate objectives to develop valid animal models (31).

These animals promise exciting models to understand and conceptually intervene in human disease processes and thereby lead to alleviation of human suffering. [For an summary of the diversity of pathologies amenable to study by transgenic technology, see the text edited by Monastersky and Robl (18).] More recently, many individuals have speculated about the possible use of transgenic mammals as organ sources for human transplant. Ultimately, of course, the technology has potential for direct application to human beings. On the one hand, it holds out the promise of correcting debilitating genetic diseases. As our understanding of the interaction of genetics and human personality increases, however, the technology also has the potential to radically alter human nature.

Human Disease Models. Transgenic animals can provide insight into the development and progression of many human diseases, and cystic fibrosis (CF) is a classical example of a disease amenable to such study (1). Cystic fibrosis is a recessive, autosomal disorder and is inherited in classical Mendelian fashion. Disease symptoms originate from abnormal function of epithelial cells in the respiratory, digestive, and reproductive tracts. The abnormal function of these cells is due to a defective chloride ion channel protein called the cystic fibrosis transmembrane conductance regulator (CFTR). Defective chloride transport causes secretory tissues (e.g., lung, pancreas, intestine) to accumulate a thick mucus, characteristic of the disease. The pathology of CF arises from mucus accumulation in affected organ systems, such as lung infections and inadequate intestinal function.

The human CFTR gene has been identified and characterized. Although numerous mutations in this gene have been shown to lead to CF symptoms, defects in a very small region of the gene are responsible for 70 percent of reported cases.

While the biological basis of CF is clear, this understanding did not clarify its pathological basis or progression mechanism, nor did it lead to effective therapeutic approaches. The specific causal relationship between a defective chloride channel in a cell membrane and accumulation of mucus in the surrounding tissue remained unresolved. Consequently development of rational and effective therapies for the underlying disease has been difficult (32).

Previous work demonstrated that mice contain a gene equivalent to the human CFTR gene. Thus creation of a strain of mice with defective CFTR genes seemed to present a reasonable model to study CF in humans (32,33). Since the CF transgenic mouse apparently truly mimics human CF, it can serve as an important means to both clarify disease etiology and facilitate rational therapies. In an animal model we can follow disease sequelae in ways

that are impossible or ethically repugnant in humans. Furthermore animal models allow us to develop and explore potential treatment modalities that would also be impossible or would pose ethical questions in humans. The list of genetic diseases that can be studied by such technology is appreciable and accounts for a significant aspect of human suffering (34,35).

Transgenic animals also provide insight into other, nongenetic, diseases. For example, transgenic mice serve two roles in AIDS research (36). First, a transgenic mouse strain was developed with a complete HIV proviral transcript. Progeny mice from this strain, which develop diseaselike symptoms, can potentially serve as a model system to study the etiology of AIDS. Second, transgenic mice have been created that carry only parts of HIV. These strains can play an important role in drug development and allow testing of certain antiretroviral drugs. Both the strengths and weaknesses of using transgenic animals in AIDS research has recently been reviewed by McCune (37).

Organ Transplantation and Xenobiotic Sources. During the past quarter century organ transplantation as a relatively common clinical therapeutic approach for many disease conditions has rapidly increased. However, the procedure has been restricted by two major difficulties: (*1*) organs available for transplant are in extremely limited supply relative to the potential clinical need (38,39); and (*2*) despite use of immunosuppressive drugs, even ideally tissue-matched organs often undergo host rejection (39–41). The latter problem arises from immunological reactions elicited by the recipient's immune system to antigens present on the donor organ. Xenotransplantation initially seemed to provide a means to address the first of these difficulties, namely animals were an unlimited source of organs (39). The procedure was sufficiently promising that *Science* mentioned it as a hot research area in 1996 (42).

While xenobiotic sources seemingly promised an unending organ supply for transplantation, the procedure did nothing to address the second, and equally problematic, phenomenon of host rejection. Transgenic animals, however, have recently been considered as potential sources of organ donors to address both problems, and initial research has focused on the initial, and potentially most overwhelming, transplant immunological barrier, namely hyperacute rejection (39–41).

Hyperacute tissue rejection is triggered by complement system activation in the recipient against proteins recognized as "non-self" on the foreign tissue. Complement activation can be inhibited by a variety of drugs; however, this technique leaves the recipient with a compromised immunological defense against infectious organisms. To obviate complement system mediated host rejection, transgenic animals, such as pigs, that express human complement proteins on their cell surfaces have been created (39–41,43). Animal organs with such cell surfaces do not trigger the complement cascade reaction in an organ recipient. Because complement activation does not occur, hyperacute rejection is prevented; the host's complement defense system is also left intact (39–41). Enthusiasm for the technology has become so intense that commercial involvement has grown (29,44).

Despite initial encouragement about transgenic xenobiotic organ sources, however, serious concerns have arisen regarding potential transmission of infectious agents, especially retroviruses, from transplanted animal organs (39–41,44–49). These concerns are sufficiently pressing that while the British government gave qualified approval for continued research into the transplantation of pig organs into humans, it ruled out clinical trials until further research demonstrated safety and efficacy of the procedure (50,51). The United States will allow limited clinical trials to go forward under stringent guidelines established by the Food and Drug Administration (FDA), the Centers for Disease Control and Prevention (CDC), and the National Institutes of Health (NIH) (52,53).

Human Gene Therapy. Because the difference between human and animal is a matter of evolutionary degree, ultimately transgenic technology promises direct intervention into human genetic diseases by replacing dysfunctional genes. Although somewhat dated now, Friedmann (34) provided an excellent and very readable summary of this area of research. He noted that diseases arising from bone marrow defects would be most amenable to genetic intervention. Because bone marrow cells are readily susceptible to infection by retroviruses, a common way to introduce foreign genetic material into mammalian cells, these cells would be relatively easy to alter genetically. Furthermore, because these cells can be removed from a patient, altered and manipulated in vitro, and then reintroduced back into the patient, the potential rejection of altered gene products by a recipient's immune system is reduced.

For similar reasons, diseases resulting from hormonal or other diffusible protein deficiencies (e.g., insulin or blood-clotting factor deficiencies) may also be amenable to genetic treatment. Skin cells can be readily transformed by a variety of methods and then reintroduced back into the patient. Since the process again involves a tissue autograft, host immune rejection is again reduced. The technique has theoretical therapeutic implications for many diseases such as hemophilia and insulin-deficient forms of diabetes.

Transgenic Organisms and Basic Science

While it is difficult to draw sharp distinctions between utilitarian goals and the basic science necessary to achieve them, transgenic technology now plays a profound role in fundamental biological science. Like the recombinant DNA technologies that preceded them, undoubtedly the most significant role of the transgenic technologies will be as tools to gain new insights into nature. Reflecting on the current state of biological science, Verbeek noted, "The fast growing knowledge about the complex biology of higher eukaryotic systems demanded new experimental models. Transgenesis of mammals is one of the most fruitful techniques to create these models (54, p. 1)."

Transgenic organisms are radical new tools for fundamental research in all of the biological and medical sciences; they are powerful new instruments that are opening new windows to understand the natural

world. Since transgenic technology first began in 1980, it has grown exponentially. Transgenic animals have contributed greatly to elucidating complex biological processes at the molecular level (31). The magnitude of this new technology's impact on research inquiry is illustrated by citation analysis of Medline© references to transgenic organisms. In 1987 Medline© cited 163 references to *transgenic organisms* (including both plants and animals). Since that time the number of references has increased exponentially; by 1992 there were over 1000 Medline© citations in which the word *transgenic* was either in the title or was listed as a keyword (1). A cursory Medline© search indicates that in 1997 alone, there were almost 25,000 references to *transgenic animals*.

Organisms Used and Research Goals. Given the ubiquity of mice or rats in laboratory research, a majority of the studies noted above involved these animals. Nevertheless, these transgenic research programs also used a diversity of other animals. They span a wide evolutionary range and include relatively common laboratory organisms, such as fruit flies, zebrafish, and *Caenorhabditis elegans* (a nematode used extensively in developmental biology studies), to more exotic organisms such as the medaka fish and silkworms. Many animals reflect an ultimate commercial or agricultural interests such as goats, pigs, sheep, and cows.

Research involving transgenic organisms is as complex and diverse as the organisms themselves. Projects range from those with immediate, pragmatic goals — such as increases in agricultural products or solutions to human health concerns — to basic inquiries about the behavior of biological systems, such as antisense RNA, ribozyme action, gene expression in *Drosophila*, zebrafish developmental processes, or the role of molecules like FosB proteins described earlier.

Why Are Transgenic Organisms Important? As previously noted, development of transgenic organisms parallels the use of recombinant DNA in the early 1970s when many researchers were excited by potential commercial applications of recombinant DNA. Despite the utilitarian role, however, these same scientists saw the technology as a powerful tool to explore nature. The technique provided a means to isolate individual genes (or gene clusters) and their products away from a complex organism into a much simpler procaryotic organism. Researchers believed that such isolation would lead to a clearer understanding of gene function, regulation, and interaction (17). Indeed, the history of recent molecular biology has confirmed the validity of that belief, and our understanding of such diverse phenomena as gene action, immunology, ecological processes, or neurobiology has grown immensely during the past 20 or more years as a direct result of recombinant DNA techniques.

Furthermore it would have been difficult to create a list of new things scientists expected to discover with the early recombinant DNA research. One could point to potential societal benefits, such as new understandings of gene function and similar phenomena, expected from the work. For most scientists, however, this was a technique with tremendous investigatory power that could be used to open new vistas of biological knowledge. It was this epistemological dimension of the research that animated many scientists' interest in recombinant DNA technology (17).

In a similar fashion, it is problematic to create a tidy list of projects that transgenic organisms will solve. The dilemma arises from the nature of scientific inquiry that does not permit clear predictions about its own nature and direction, a point marvelously illustrated by the outcome of the *fosB* knockout mouse experiment described previously. Scientific inquiry fundamentally is more than a simple *accumulation of facts* about nature; rather it is a *method to understand* the natural world.

How is this digression into epistemology connected with the technology of transgenic organisms? Simply it is this: Despite their artificial creation, transgenic organisms are now a part of the natural, empirical scientific world. As such, they are important tools to gain a greater understanding of nature. Like the transgenic procaryotes that preceded them in the repertoire of scientific investigative tools, transgenic eucaryotes will ultimately provide humankind with fundamental and valuable knowledge of the natural world.

ANIMALS, BIOTECHNOLOGY, AND ETHICS

Like many technological innovations, development of transgenic organisms presents us with ethical quandaries. In addition to their wondrous utility, their creation also raises potentially troubling ethical questions. While these questions are complex, and often appear refractory to solution, they have been addressed elsewhere (1,2,55,56) and will be extensively dealt with in this volume. Only brief reflection on the issues raised in this article is appropriate here. The ethical questions seemingly fall into four broad areas: breaching species boundaries, potential for animal harm, environmental concerns, and potential human application.

Species Barriers

Transgenic organisms raise obvious concerns about "species barriers." Are *species* physical entities so inherent in the fabric of nature that we are morally culpable in breaking the barrier between them? Is there anything morally significant about being a member of a *species*? These questions might be put another way. Should sheep be allowed to be sheep without carrying burdens of non-ovine genes, some of which are intended only for human benefit?

A scientific perspective suggests negative answers for these questions. The notion of *species* as fixed natural entities is relatively new in human thought and is contrary to modern scientific views. From antiquity all species were seen as eternal and immutable, and this *essentialistic* notion of species dominated Western thinking well into the nineteenth century. From this view, the organisms we encounter in daily existence reflect an essential form that exists within created nature itself.

As noted previously, Charles Darwin began to shift the scientific notion of *species* away from this *essentialist concept*, and post-Darwinian biologists reject all *essentialist* notions of species. Thus there is nothing unique about being a member of a *species* that would seemingly command moral recognition. *Species* do have biological reality, but it is not an *essentialist* reality (14).

Species are not immutable "type-forms" woven into the fabric of nature but are defined in populational terms, a reality that is contextual and is spatiotemporally bounded (14,57). Moreover, as a scientific heuristic device, the *species* concept provides a way to organize and simplify the complex diversity of living organisms; *species* provide what Mayr refers to as a *taxon*, namely an entity in nature with taxonomic significance. Again, it is not readily apparent that taxonomic significance can mandate moral significance.

Transgenic Animal Welfare

Sentient transgenic animals raise serious concerns about potential pain or suffering that we might cause to an animal capable of such experiences. Are we morally permitted to intentionally "create" an organism that we know will ultimately suffer severe debilitation or experience great pain? CF mice exhibit many physical symptoms common to the human disease, including premature death. Creation of the CF mouse, or any animal as a model of human disease, involves potential harm or suffering to individual transgenic animals. So the moral question arises, Do we have a right to intentionally create such an animal capable of experiencing pain that we know will develop such a debilitating and painful disease?

The argument has been made elsewhere (1,2) that transgenic animals are not, in principle, significantly different from other animals. If one accepts this claim, then resolution of questions about pain and suffering are similar to questions regarding animal use in general. Commercial and laboratory use of sentient animals is controlled by regulations and principles established by the Department of Agriculture and by National Institutes of Health Guidelines (9,10). As McCarthy noted, most use of transgenic organisms will be governed by these entities. He also noted, however, that "there are gaps in oversight due in part to whether the particular methods are publicly or privately funded" (58, p. 526).

Like many questions involving research use of animals, the moral parameters for using transgenic organisms must be contextually defined (59). Potential animal suffering must be weighed against potential human suffering alleviated through knowledge gained by animal use. After due consideration we might conclude that creating a strain of mice as models of CF is morally justified because it may ultimately alleviate the acute suffering of a significant number of human children. On the other hand, using similar techniques to create a strain of dogs with some serious physical abnormality (e.g., severely shortened legs) to become novel pets would arguably be morally reprehensible.

Ecological/Environmental Concerns

A third ethical concern entails possible ecological damage arising from intentional or unintentional release of transgenic organisms. How do we ensure that these novel organisms do not unduly disrupt natural habitats and cause serious environmental damage? The potential environmental impact of transgenic organisms has been poorly studied; these questions are probably quite significant, however.

In many respects the environmental impact of organisms used in biotechnology should be minimal, as there is no intent for the animal to be released. Indeed, because of the great expense involved in their creation, extreme precautions are taken to prevent the animal from escaping the laboratory or farm environment. Serious environmental concerns arise, however, in two areas. First, care must be exercised in projects where there is a high potential for a genetically modified organism encountering naturally related organisms, for example, in mariculture. Second, one need only drive through any kudzu-covered forest in the United States to envision the potential ecological havoc that a genetically altered plant might create. (Kudzu, of course, is not a genetically altered plant.) Both of these issues have been fully addressed elsewhere (1,38,55).

Human Application

A fourth ethical issue arising from transgenic technology is perhaps the most serious and the most difficult to resolve, namely application of transgenic technology to humans. The articles by Friedmann and Jaenisch demonstrate that the clear direction of this research is toward human application. As is noted elsewhere in this volume, the ethical issues in this area are complex and often troublesome. On the one hand, for many people we are morally culpable if we have the ability to alleviate the suffering of an individual with a profound genetic defect, such as CF, Tay-Sachs, or sickle cell anemia, and we fail to use that ability. This conclusion, however, clearly places us onto the moral philosopher's slippery slope. Moral distinctions between significant genetic defects (e.g., CF or sickle cell anemia), with their associated suffering, and merely attractive traits that individuals might like for their children to possess (e.g., large body mass so that a son could become a highly paid NFL linebacker) are reasonably clear. Moral distinctions with such extremes, however, are rare. More often we face subtle and less clear choices that are thus more ethically problematic.

These moral considerations become even more problematic when we ask about our obligations to future offspring of individuals suffering from a treatable genetic disease (38). Many people readily find moral obligations to treat individuals with somatic cell deficiencies. However, obligations to treat such deficiencies at the germ-cell level are more complex. While we may be obligated to alleviate the suffering of a person with a genetic disease, are we obligated to ensure that those individuals can produce children who lack the genetic defect? Alleviation of immediate suffering, if we can do so, seems a reasonable obligation. However, does that obligation extend to some

future individual not yet conceived? These moral questions intuitively appear problematic, and answers do not appear readily obvious.

CONCLUSIONS

For many people, animal biotechnology promises a powerful new vision of general welfare and health for both humans and other animals. Domestic breeding, enhanced by modern techniques such as marker-assisted selection, and transgenic organisms (and the various technologies associated with their creation) may present a cornucopia of new wealth (both financial and abundance of valuable material possessions or resources).

Creation of transgenic animals, especially, is a new tool for scientific inquiry and has the potential to alter science itself. New questions about nature, which were impossible in the technology's absence, could be asked and radical new answers proposed. Like many new scientific tools, studies with transgenic organisms often have serendipitous turns. The apparent maternal behavior pattern linked to the *fosB* gene discussed in this article is a good example of the unexpected paths this technology can reveal.

Concomitantly the technology's power also creates profound possibilities for moral abuse and environmental chaos. Our most deeply felt sense of human values can be seriously distorted and corrupted by even moderate abuse and potentially could lead to distortions of fundamental and essential aspects of both human and animate nature. From some perspectives, transgenic technologies represent a Frankenstein-like bargain with nature and are the realization of Chargaff's prediction of the "Devil's Doctrine," that what can be done (technologically) will be done [regardless of broader social concerns (60)].

Nevertheless, transgenic organisms are not fictional or creatures of ancient mythology; they are a reality of nature that we humans must deal with. Like most scientific inventions, transgenic animals pose complex moral issues. And, as with any truly moral issue, the fashioning of transgenic organisms present us with moral ambiguities and treacherous slippery slopes. They pose questions about our moral obligations to both our fellow humans as well as the other living beings with whom we share this planet. The moral problems are made more complex as we attempt to discover our obligations to others members of the biological web within which all living beings are intertwined (1).

Science may help enlighten and focus the moral landscape of these questions; however, it does not provide us with adequate tools to derive answers. This difficulty is inherent in the limitations of scientific inquiry, for these answers "are as many as there are different cultural, religious, and philosophic perspectives" (59, p. 518). Because these pluralistic perspectives are not subject to empirical boundaries or testable propositions, science, as a mode of inquiry, is ill prepared to deal with them. Thus answers to questions on our moral obligations must lie in other aspects of the broader human condition. Wrestling to find anchors on the slippery slopes of our moral landscape nevertheless is a natural aspect of our humanity.

ACKNOWLEDGMENT

Part of the work on this article was completed during a sabbatical year at Case Western Reserve University, and the author thanks his many colleagues there for their helpful comments and criticisms. Drs. Melinda Duncan and Milton Stetson, at the University of Delaware, also provided help defining technical terms. Finally, the author thanks the National Science Foundation for financial support (Grant number SBR 9602023).

BIBLIOGRAPHY

1. R. Singleton, Jr., *Hastings Center Rep. (special suppl.)* **24**, S4–S14 (1994).
2. R. Singleton, Jr., in J.C. Gonder and E.D. Prentice, eds., *Genetic Engineering and Animal Welfare: Preparing for the 21st Century*, Scientists Center for Animal Welfare, Washington, DC, 1999.
3. V. Morell, *Science* **276**, 1647–1648 (1997).
4. C. Vilà et al., *Science* **276**, 1687–1689 (1997).
5. C. Vilà et al., *Science* **278**, 208–209 (1997).
6. N.E. Federoff and R.M. Nowak, *Science* **278**, 206 (1997).
7. J.P. Scott, O.S. Elliott, and B.E. Ginsburg, *Science* **278**, 205 (1997).
8. S. Budiansky, *Atlantic Monthly* **284**, 39–49, 52–53 (1999).
9. R. Singleton, Jr., *Persp. Biol. Med.* **37**, 576–594 (1994).
10. R. Singleton, Jr., *Persp. Biol. Med.* **38**, 41–57 (1994).
11. *The Concise Columbia Encyclopedia*, Columbia University Press, 1995.
12. Committee on Biotechnology Research Subcommittee, Fundamental Science, *Biotechnology for the 21st Century: New Horizons*, U.S. Government Printing Office, Washington, DC, 1995.
13. N. Maclean, in N. Maclean, ed., *Animals with Novel Genes*, Cambridge University Press, Cambridge, UK, 1994, pp. 1–20.
14. E. Mayr, *The Growth of Biological Thought*, Harvard University Press, Cambridge, MA, 1982.
15. J.R. Brown et al., *Cell* **86**, 297–309 (1996).
16. J. Cohen, *Science* **273**, 577–578 (1996).
17. R. Singleton, Jr., in D.A. Robins and A.R. Dyer, eds, *Ethical Dimensions of Biomedicine*, Charles C. Thomas, Springfield, IL, 1981.
18. G.M. Monastersky and J.M. Robl, eds., *Strategies in Transgenic Animal Science*, ASM Press, Washington, DC, 1995.
19. N. Maclean, ed., *Animals with Novel Genes*, Cambridge University Press, Cambridge, UK, 1994.
20. S.E. Lindow, N.J. Panopoulos, and B.L. McFarland, *Science* **244**, 1300–1307 (1989).
21. V.G. Pursel et al., *Science* **244**(4910), 1281–1288 (1989).
22. E.R. Cameron, *Mole. Biotechnol.* **7**, 253–265 (1997).
23. K.A. Ward and C.D. Nancarrow, *Mole. Biotechnol.* **4**, 167–178 (1995).
24. M. Gomez-Chiarri, V.L. Kirby, and D.A. Powers, *Mar. Biotechnol.* **1**, 269–278 (1999).
25. E.P. Cunningham, *Livest. Prod. Sci.* **58**, 1–24 (1999).
26. J.D. Murray, *Theriogenol.* **51**, 149–159 (1999).
27. R.J. Wall, D.E. Kerr, and K.R. Bondioli, *J. Dairy Sci.* **80**, 2213–2224 (1997).
28. T.J. Smith, *Biotechnol. Adv.* **12**, 679–686 (1994).

29. E. Pennisi, *Science* **279**, 646–648 (1998).
30. G. Brem and M. Müller, in N. Maclean, ed., *Animals with Novel Genes*, Cambridge University Press, Cambridge, UK, 1994.
31. T. Nomura, *Lab. Animal Sci.* **47**, 113–117 (1997).
32. M. Barinaga, *Science* **257**, 1046–1047 (1992).
33. J.N. Snouwaert et al., *Science* **257**, 1083–1088 (1992).
34. T. Friedmann, *Science* **244**, 1275–1281 (1989).
35. S.M. Weissman, *Proc. Nat. Acad. Sci. (USA)* **89**, 111–112 (1993).
36. A.D. Lewis and P.R. Johnson, *Trends in Biotechnology (Ref. ed.)* **13**, 142–150 (1995).
37. J.M. McCune, *Science* **278**, 2141–2142 (1997).
38. A.L. Caplan, *Am I My Brother's Keeper: The Ethical Frontiers of Biomedicine*, Indiana University Press, Bloomington, IN, 1997.
39. Committee on Xenograft Transplantation: Ethical Issues and Public Policy, Institute of Medicine, *Xenotransplantation: Science, Ethics, and Public Policy*, National Academy Press, Washington, DC, 1996.
40. D.K.C. Cooper, *Frontiers in Bioscience* **1**, 248–265 (1996).
41. J.L. Platt, in J.C. Gonder and E.D. Prentice, eds., *Genetic Engineering and Animal Welfare: Preparing for the 21st Century*, Scientists Center for Animal Welfare, Washington, DC, 1999, pp. 53–75.
42. Anonymous, *Science* **270**, 1902 (1995).
43. M. Schmoeckel et al., *Transplantation* **62**, 729–734 (1996).
44. F. Hoke, *The Scientist* **9**, 1 (1995).
45. X.-J. Meng et al., *Proc. Nat. Acad. Sci. (USA)* **94**, 9860–9865 (1997).
46. C. Patience, Y. Takeuchi, and R.A. Weiss, *Nature Medicine* **3**(3), 282–286 (1997).
47. R. Sikorski and R. Peters, *Science* **276**, 1893 (1997).
48. R.E. Michler, *Emerging Infect. Dis.* **2**, 64–70 (1996).
49. M.J. Hanson, L.-M. Russow, and C.R. McCarthy, *Hastings Center Rep.* **29**, 22–25 (1999).
50. A.S. Daar, *World J. Surg.* **21**, 975–982 (1997).
51. N. Williams, *Science* **275**, 473 (1997).
52. C. O'Brien, *Science* **271**, 1357 (1996).
53. G. Vogel, *Science* **279**, 30 (1998).
54. J.S. Verbeek, in L.F.M. van Zutphen and M. van der Meer, eds., *Welfare Aspects of Transgenic Animals*, Springer Verlag, New York, 1997, pp. 1–17.
55. B.E. Rollin, *The Frankenstein Syndrome: Ethical and Social Issues in the Genetic Engineering of Animals*, Cambridge University Press, Cambridge, England, 1995.
56. L.F.M. van Zutphen and M. van der Meer, eds., *Welfare Aspects of Transgenic Animals*, Springer Verlag, New York, 1997.
57. D.L. Hull, *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*, University of Chicago Press, Chicago, IL, 1988.
58. C.R. McCarthy, *Hastings Center Rep. (special suppl.)* **24**, S24–S29 (1994).
59. S. Donnelley, *Hastings Center Rep. (special suppl.)* **24**, S14–S24 (1994).
60. E. Chargaff, *Persp. Biol. Med.* **16**, 486–502 (1973).

See other entries ANIMAL, MEDICAL BIOTECHNOLOGY, LEGAL, LAWS AND REGULATIONS GOVERNING ANIMALS AS SOURCES OF HUMAN ORGANS; ANIMAL, MEDICAL BIOTECHNOLOGY, POLICY, WOULD TRANSGENIC ANIMALS SOLVE THE ORGAN SHORTAGE PROBLEM?; FDA REGULATION OF BIOTECHNOLOGY PRODUCTS FOR HUMAN USE; see also RESEARCH ON ANIMALS entries.

U

UNIVERSITY–INDUSTRY RESEARCH RELATIONSHIPS, ETHICS, CONFLICT OF INTEREST

ROBERT P. LAWRY
THOMAS W. ANDERSON
Case Western Reserve University
Cleveland, Ohio

OUTLINE

INTRODUCTION

Since World War II, and increasingly since the early 1980s, there has been a widespread public policy initiative to increase and deepen research and development relationships between industry and academic institutions. The hope is that new technologies will be more rapidly and effectively produced both to promote the general welfare as well as to increase U.S. dominance in the global marketplace. Ethical concerns, however, have arisen simultaneously. Scientific and educational integrity are threatened in a variety of ways by these new academic-industry partnerships. As questions of conflicts of interests increase and deepen, good public policy answers lag behind.

HISTORICAL CONTEXT

Until the late nineteenth century, colleges and universities in the United States were primarily teaching institutions. Their role was largely confined to the transmission of knowledge. The discovery or enlargement of knowledge was secondary to the application of practical subjects that would have utility for students and for the larger society. In the public sector, the Morrill Act of 1862 allocated federal lands to states for the founding of "at least one college where the leading object shall be, without excluding other scientific and classical studies ... to teach such branches of learning as are related to agriculture and the mechanic arts" (1). The Act led to the founding of new state institutions and to the support of recently established state colleges, embodying a vocational, practical educational focus.

Nothing changed the educational landscape in the United States more than the founding of Johns Hopkins University in 1876. It has been described extravagantly but not inaccurately as, "perhaps the single, most decisive event in the history of learning in the Western Hemisphere" (2). Prior to the founding of Johns Hopkins, graduate education and the research associated with it were left to the leading European universities, the German universities in particular. The Hopkins model gave primacy to graduate education over undergraduate instruction and brought together the German concepts of advanced training and the generation of new knowledge, particularly in the natural sciences, into the teaching environment of American universities. During the next 50 years, the uniquely American research university took root and began to flourish.

From 1900 to 1920, private and public universities were participants in the general economic well-being of the era. For private institutions, endowments were established and began to grow. The philanthropies of John D. Rockefeller and of Andrew Carnegie set examples for decades to come. Public institutions secured their places as important and useful state resources requiring more than tuition income to accomplish their increasingly diverse objectives.

It was not yet clear, however, how much of the nation's scientific research would be done in universities. Government agencies and independent research institutions competed with universities for the resources to hire researchers, build facilities, and support research activities. Public universities looked with modest success to their state legislatures for research support and both public and private universities began to appreciate the power of private philanthropy. Early in the century, Harvard President Charles Eliot noted, "it is clear that men of means, who reflect on the uses and results of educational endowments, are more and more inclined to endow research" (3). While the American research university hallmark of combined teaching and research was being firmly established in the first quarter of the twentieth century, the funding structures to adequately support these no longer discretionary activities were not yet in place.

World War I effectively nationalized the research universities, focusing all faculty and student efforts on winning the war. The role of science was enhanced, and the bonding of applied and basic research was seen as important to the war effort. Following the war, the major philanthropic foundations, established in the late nineteenth and early twentieth centuries, began to take notice of scientific research as a means for the "amelioration of the human condition through the advancement of knowledge" (3). Foundation grants enhanced and expanded university research in the years between the two world wars. Fears of foundation control over university educational and research efforts did not materialize as it became clear that the needs of

research universities would far surpass the resources of the foundation community.

Following World War I, the advancement of scientific knowledge through university-based research began to require a partnership of universities, private philanthropists, foundations, and now corporations. Major corporate laboratories had been established early in the century for applied research purposes. Interactions between applied industrial researchers and university basic scientists became common. University graduates were recruited to industrial laboratories and faculty members consulted with corporations. Corporate financial support for university research followed naturally from these relationships. Even early on, the differences between universities (dedicated to the advancement of knowledge) and corporations (dedicated to financial gain) raised the potential for misunderstandings. Corporations were not convinced of the importance of basic research and faculty members found that corporate interests were often too narrow to be of educational or scientific interest. But, interactions continued in a variety of forms (graduate fellowships, research contracts, consulting relationships, etc.) in generally ad hoc institutional arrangements.

World War II again found universities deeply involved and effected by the war effort. The Manhattan Project and the Radiation Laboratory at the Massachusetts Institute of Technology were major scientific collaborations between university scientists and the federal government. While it may be overstating the case to say that science won the war, it nevertheless played such a decisive role that neither academic research nor federal scientific interests would ever be the same again. The war not only highlighted the need for a federal science policy, it brought the federal government into a permanent funding relationship with university research. World War II marked the shift from primarily private to primarily public funding for major research projects. Postwar foundation support became focused on three broad objectives: medical and health fields, strengthening the system of university research, and social and behavioral sciences (4). Foundation funding increasingly nudged research universities in the direction of academic excellence rather than the targeted defense related research necessary for a war effort.

Federal Government Involvement

In 1950 the National Science Foundation (NSF) was created to fund basic academic science with public funds. Until this time the postwar research agenda had been driven largely by programmatic funding from the armed services. University administrators and researchers were increasingly concerned about the source and direction of military research. California Institute of Technology President Lee A. DuBridge called the prevailing military authority over the nation's research program "an anomalous and precarious situation to have the future of basic research hang by the thread of continued appropriations to the military agencies, or of their continued interests" (4). In response to this concern, the National Science Foundation was created to fund basic research, by "greasing the wheels of science," funding scientific research, and developing "a national science policy" (4). Although initially inchoate in its mission, both the role of the NSF and the nation's interest in basic research changed on October 4, 1957, with the Soviet Union's launch of Sputnik, the first space ship to orbit the earth. Because of the threat of an attack from space during the height of the cold war, the days immediately following *Sputnik* were consumed with much national soul-searching. Initial responses included the establishment of the National Aeronautics and Space Administration (NASA) and increased federal appropriations to existing agencies for basic scientific research.

While funding levels would vary over the next 40 years, commitments to basic versus applied research would wax and wane, and the interests of social justice and economic development would often compete, the foundational commitment of public funds for academic research would not be seriously threatened. Basic academic research became a growth industry, largely funded by federal money. This federal commitment also included the beginnings of a federal scientific establishment, initially presided over by the Presidential Science Advisor but eventually permeating all branches and levels of government. As the NSF gained its footing, it was joined (in federal priority) by its sister institution the National Institutes of Health (NIH), the research arm of the Public Health Service. The NIH was to biomedical science and scientists what the NSF was to natural science and scientists.

The 1960s placed significant strains on research universities — student unrest, sluggish economic conditions, governmental oversight, and changing values — leading to a loss of confidence in these institutions and their primary missions of teaching and research. Egalitarian federal programs and the proliferation of institutions seeking research funding led to changes in both the recipients of public funds and the role of private funding sources. In the 1970s one response to these conditions found the leading research university faculty and administrations beginning to seek closer relationships with corporations. While driven partially by funding considerations, these initial efforts by research universities also "implied a break with the cloistered mentality that had flourished in the 1960s" (4).

In the early years of the twentieth century, corporations purchased research capability from universities either through direct support of specific research activities or through joint institutes for applied research. By the beginning of World War II, however, many corporations had established their own laboratories, and thus looked to universities for the theoretical work that would underlie their corporate research interests. But corporate and university interests were inevitably different since universities seek to advance knowledge while industry must apply it. Universities were organized horizontally and corporations were organized vertically. Individual faculty members determine whether they will work with corporations no matter how friendly their university policies may be to corporate interactions. But despite these differences in culture, by the 1980s corporate

support of university scientific research was growing at unprecedented rates.

The mid-1970s saw fledgling efforts by the NSF to encourage university–industry interaction for the specific purpose of technology transfer. At about the same time the first large ($23 million) university–corporate contract (Harvard Medical School and the Monsanto Corporation) based on mutual research interests was announced. Unprecedented on both sides, "Monsanto provided funds for endowment, research support, and facilities for a pair of Harvard scientists. In return the company was promised the patent rights to any discoveries that resulted from their research" (4). Viewed with both interest and suspicion, the agreement was only the beginning of similar arrangements at other institutions and corporations, particularly in microelectronics and biotechnology. This highly structured form of technology transfer was supplemented by small start-up companies seeking to commercialize scientific discoveries. Such companies often were founded by scientists, engineers, and graduate students from research universities with funding from venture capital firms.

The blueprint for biotechnology start-up companies was drawn by Genentech, a firm founded in 1976 by a venture capitalist who struck an agreement with a molecular biologist who was interested in commercializing his new technology that led to synthesizing the human gene for insulin. Genentech licensed the discovery to the Eli Lilly Corporation and "thus validated the idea that genetic engineering could produce valuable commercial products..." (4). Venture capitalists, scientists, and universities immediately understood the potential for significant wealth in similar relationships. Scientists could continue their research — either within the academy or at the new firms — with private funding that could lead to both important scientific breakthroughs and the prospect of enormous wealth. Over 200 biotechnology firms were founded between 1980 and 1984, and half of all biotechnology venture capital raised by 1988 was raised during the two years following the Genentech announcement (5). The potential for conflicts of interest and commitment were recognized as enormous, although not always immediately.

The explosion of biotechnology research was coupled with the highly charged business climate of the 1980s. Government deregulation, junk bond financing, and declining federal support for research universities all enhanced the environment for business–university collaborations and partnerships of almost infinite variety. Both state and federal governments enacted legislation to encourage these partnerships in the hope of transferring technology, increasing corporate competitiveness, and retaining business and industry within state boundaries. Universities scrambled to compensate for reduced funding levels and to retain key faculty members who were increasingly being lured outside the academy by entrepreneurial opportunities.

At the federal level, the NSF had two programs (the Industry–University Cooperative Research Projects Program and the Industry–University Cooperative Research Centers Program) which were established in the 1970s to develop and sustain corporate–university research partnerships. By 1989, 41 Research Centers around the country were operational and 22 were self-sustaining. A second part of the Research Centers program was founded in 1985 to support Engineering Research Centers with 18 established by 1989 (6). All potential economic benefits from the collaborative arrangements encouraged by these centers remain with the centers despite significant funding by the NSF.

Significant Legislation

Perhaps more important were changes by federal agencies allowing universities to retain patent rights from inventions and technologies discovered in federally funded research projects. The Bayh–Dole Act of 1980 specifically granted such rights. The Federal Technology Transfer Act of 1986 permitted federal scientists and university scientists to collaborate with industry to develop commercially patentable ideas. The Act specifically authorized "private companies to gain the exclusive rights to patents, while universities and scientists could receive royalties" (6). These federal efforts, along with complementary state legislation, nudged, if not pushed, universities into the technology transfer business and into increasing collaborations with industry with the objective of commercializing the results of research.

As the Harvard Medical School–Monsanto relationship was followed by equally large and potentially controversial business-university partnerships in the 1980s, issues of academic freedom, freedom to publish, conflicts of interest and commitment, and secrecy began to surface both inside and outside of the academy. Conferences were held with academics, business, and government participants to consider and address such matters. The Government–University–Industry Research Roundtable was founded by the National Academy of Science, the National Academy of Engineering, and the Institute of Medicine in 1984 to "provide a forum where scientists, engineers, administrators, and policy makers from government, university and industry can come together on an ongoing basis to explore ways to improve the productivity of the nation's research enterprise" (6). Although decidedly "pro" business–university relationships, the Roundtable developed a model agreement for business–university research partnerships which set standards for publication, intellectual property ownership, and licensing and patenting procedures.

Despite a climate generally favorable to university–industry collaborations, there was also concern about abuse. Particularly troublesome were ethical issues arising from conflict of interest questions. In 1989 the NIH proposed guidelines that required individual decision makers involved in funded research to disclose "all financial interests and outside professional activities." Any perceived conflict of interest uncovered by these reporting requirements was to be especially noted and resolved prior to funding. Moreover, researchers were prohibited from holding equity or options in any company affected by the outcome of their research. Record keeping was extensive on the part of universities, which bore much of the

burden of administering the guidelines. Harsh and wide-ranging criticism of the proposed guidelines resulted in their withdrawal within months of publication. NIH then proceeded to develop formal regulations on the subject, which, together with similar regulations proposed by the NSF, became binding in 1995. Less restrictive, these rules required researchers funded by NSF or NIH to notify their home institution if "they, their spouses, or their dependent children have financial interests — exceeding $10,000 or 5% ownership — in companies that might be affected by their research" (7). But once the researcher has complied with this threshold requirement, it is up to the institution to decide whether it is a conflict of interest and what to do about it. The rules do not cover the situation when an institution has a financial interest in the outcome of federally funded research. Thus it became important to understand both what the concept of conflict of interest entails and to examine proposed remedies for conflict situations.

CONFLICT OF INTEREST

One widely approved definition of conflict of interest is that it is "a set of conditions in which professional judgment concerning a primary interest (such as a patient's welfare or the validity of research) tends to be unduly influenced by a secondary interest (e.g., financial gain")) (8). Although there is some scholarly debate about what is the key concept in conflict of interest analysis (9), a growing consensus seems to find in the risk of impaired professional judgment the locus of the problem. Thus, as one modern legal scholar has put it, "[T]he common feature which brings each of ... (a variety of) questions within the doctrinal niche labeled 'conflict of interest' is concern with the existence of some particular incentive which threatens the effective and ethical functioning of a 'person acting in the role of a fiduciary for the benefit of another or others'" (10). The widespread concern is that an individual, employed by the university to perform one or more of a multiplicity of tasks in its behalf, may be compromised in his or her judgment by an incentive that should be subordinate to what he or she is employed by the university to do. Under the NSF and NIH rules, only a "significant financial interest" counts as a problematic secondary incentive; however, since the university is responsible for determining initially what counts as a conflict of interest and how to deal with it, the university may also enlarge the parameters of their internal policies to deal with incentives beyond the financial. A further complication exists because of institutional conflicts of interest, that is, when the university has a financial stake in the research conducted by those employed by the university. Much less attention has been paid to this topic, and discussion of it will be postponed until later.

Conflict of interest describes a situation of risk, not actual impairment of function. If there is actual impairment, then there is true blameworthiness. Simply being in a conflict situation, however, is usually benign in and of itself (11). Although scholars realistically warn that the phrase as often used is "accusatory" (12), this should not be the case. Since federal public policy strongly supports university–industry partnerships, conflicts of interest are inevitable, at least at some times and in some ways, for a great many researchers in university settings. It is necessary first to recognize a conflict situation, and then to determine what to do about it. The identification question is logically and realistically prior to the remedy question. For some conflicts, nothing needs to be done. Others, of course, need, in the words of the federal regulation, to be "managed, reduced or eliminated."

Judgment

Although conflict of interest problems are not new, serious academic consideration of them is relatively recent. Using the legal professional literature extensively, Davis was the first professional ethics philosopher to isolate the importance of the risk to judgment impairment as of central importance (13). Previously Margolis had suggested that conflict of interest entailed "an avoidable exploiting of conflicting roles" (14). Davis argued that the issue was really the threat to judgment within a role, rather than a conflict between roles, which suggested a typical but different ethical dilemma. For Thompson, this was the key difference:

> "In ethical dilemmas, both of the competing interests have a presumptive claim to priority, and the problem is in deciding which to choose. In the case of ... conflicts of interest, only one of the interests has a claim to priority, and the problem is to ensure that the other interest does not dominate. This asymmetry between interests is a distinctive characteristic of conflicts of interests (8).

This asymmetry may not be characteristic, however, of all conflicts of interest. In the legal literature, for example, conflicts do occur when lawyers sometimes try to represent two or more clients with "conflicting interests." Nevertheless, in university–industry conflicts, Thompson's point regarding an asymmetry in interests seems well-taken because it is generally understood that the obligation to the university is primary for the researcher.

Luebke challenged Davis's analysis regarding the centrality of judgment, claiming that the issue was not the "correctness of the decision," but "the potential damage to the trust relationship existing between the bearer of the conflict and the person or entity for which the primary interest was to be maintained." (15). Trust is important to the maintenance of the relationship, but as Pritchard noted in defense of Davis, "... The maintenance of trust is what is under threat in conflict-of-interest situations, but precisely, as Davis says, because the reliability of professional judgment is thrown into doubt" (11). So "judgment" is the more precise term. Nevertheless, since the issue of trust is essentially one of trust of the professional judgment, there may be no substantive divergence between the two positions. The appearance issue will be discussed later.

Judgment for Davis "implies discretion" (13). It is not something mechanical or routine, something a clerk could do. Still it is important to stress that the idea must be generalized; it is judgment within a role, not a particular judgment in a role. To suggest otherwise is to fall into the

trap Davis himself fell into. In addressing the following hypothetical, Davis declared the answer to depend on whether discretion was called for in the case at hand. The issue was: "Is it a conflict of interest to recommend to one's own company a contract with another firm in which one holds substantial stock?" The question of whether there is a risk of impaired judgment cannot depend on whether the judgment involved discretion in the particular case but only on whether, in general, this role demands judgment-as-discretion at least some of the time. Surely we are as concerned in the above hypothetical with the avaricious person who deliberately seeks to line his pocket at the expense of the person or entity that has a fiduciary claim upon him just as we are concerned with the person who subconsciously or confusedly makes a bad judgment to the detriment of his institution. We want to recognize both kinds of problems as conflict of interest problems because both evidence a risk that there will be impaired performance. In establishing conflict of interest rules or guidelines, there is a need to have everyone treated alike before the fact. This cannot be done if we have to know the subjective answer concerning how discretionary the judgment was before we can determine whether the matter demanded some review. The risk must first be identified category by category before a proper remedy can be determined. Thus, the word judgment might better be replaced by the word "decision" to capture all that we want, especially because the former may imply discretion while the latter may not. Thus McMunigal says it better when he suggests the real concern is with an incentive which threatens "the effective and ethical functioning" of the conflict holder in his fiduciary role (10).

Interests

Thompson's formulation stresses the need for one or more "primary" interests being at risk of subordination to a secondary interest. For an academic working as a researcher in a university setting, there seem to be three primary interests: (*1*) research integrity, (*2*) the well-being and education of students, and (*3*) if in a clinical setting, the welfare of patients (8). Although the NIH and NSF rules focus on financial interests alone, there are surely other secondary interests that university watchdogs will want to be on the lookout for. Pritchard suggests some of these secondary interests: tenure, promotion, satisfaction from supporting one's graduate assistants, or colleagues or institutional connections, or even one's reputation (11). The problem here is that the list can be extended indefinitely. Davis would include "all those influences, loyalties, concerns, emotions, or the like that can make (competent) judgment less reliable than it might otherwise be" (13). Davis even includes "moral constraints." Surely this goes too far. Pritchard finds in Feinberg a more modest and more objective definition: "something one might attempt to advance, protect, or even modify" (11). Although this excludes things like moods and emotions, it may still be too broad to be manageable because it may include rather personal predilections concerning, say, working hours and conditions, irrelevant to real conflicts of interest concerns. But surely financial interests are too narrow a concentration, though Thompson is right in asserting that financial gain is more "pernicious and more objective and more fungible, and easier to regulate by impartial rules" (8). Since the federal government's regulations focus only on money, universities may lose sight of other secondary interests — some hardly trivial for the individual — that require enumeration and care in determining whether they threaten to impede a good, independent professional judgment/decision which damages a primary interest.

REMEDIES

There are a number of different ways to assess how problematic a conflict of interest might be. Thompson proposes two standards for assessing the severity of a conflict. First, there is the "likelihood" that the judgment will be affected. Rules of thumb under this standard include (*1*) the greater the value of the secondary gain; the greater the likelihood the judgment will be affected, (*2*) the longer and closer the association with those connected to the secondary interest, the greater the likelihood the judgment will be affected, and (*3*) the greater the degree of discretion in judgment, the greater the chance for judgment to be improperly affected. Thompson's second standard is cast in terms of the "seriousness" of the conflict. Crucial concerns here are (*1*) the value of the primary interest, meaning the potential effects on patient care or on the integrity of the research, (*2*) the scope of the effects on the project itself, but also on others, including the indirect harm that comes from loss of confidence in the researcher or in his or her institution, and (*3*) the relative accountability of the researcher. There is presumably less concern if there is reliable review of the work (8).

Another approach to standards can be drawn from the legal literature, which distinguishes at least three kinds of conflict of interest: actual, latent and potential. These categories are distinguished by the closeness of the conflict to the actual impairment in professional judgment that is the underlying concern. An actual conflict therefore is one that is certain to affect the judgment. A latent conflict is one for which there is a reasonable probability that the judgment will be impaired. A potential conflict is one that is reasonably foreseeable (13). Although there is confusion in the legal literature concerning the proper use of these categories (10), if they are simply standards to make the remedies chosen for a given situation more amenable to sorting out, perhaps they can be helpful. Use of these terms in any substantive way is simply confusing.

In any event, it is to standards like the ones articulated by Thompson and in the legal literature that those responsible for determining what to do about conflicts of interest instinctively turn. For example, Blumenthal suggests that the seriousness of the potential harm to patients and to the integrity of research in a clinical research setting require the most restrictive rules. Generally he would ban any conflicts of interest in these settings, except if the financial gain through the secondary interest is *de minimus*. Blumenthal is equally concerned with situations where students and trainees may be affected by the conflict of interest. Because restrictions on scientific communications are frequent

in university–industry partnerships, and the students' careers may be hampered by their lack of publishing results of the work, conflicts affecting educational decisions should be allowed but rarely, and then only under the closest of supervision. More empirical research needs to be done, Blumenthal argues, before it is clear what ought to be the approach to other conflict of interest categories.

In nonclinical settings not involving students or in clinical settings where the secondary interests are "nonexistent or attenuated," it is not clear how restrictive the rules ought to be. Here Blumenthal is thinking of things like "straight-forward academic–industry research relationships" and "patenting and licensing arrangements, in which clinical research may affect whether and how much royalties are received on patented products of research." Preliminary empirical data suggests, on the negative side, that industry sponsorship tends to affect the choice of research topics and also results in scientific information being held back longer, even beyond the time necessary to file a patent. On the positive side, technology transfer activities have increased in a variety of ways, and so far, there have not been reports of actual research misconduct attributable to academic-industry relationships.

Finally, there is a mixed result regarding scientific publications. Generally the relationships seem to spawn more publications, except among researchers who "receive more than two-thirds of their total research budgets from companies or add more than 20 percent to their total salaries from consulting to industries" (16). This last finding seems to confirm Thompson's notion that more intense relationships may have negative effects on researchers. This is based on the supposition that more publications are part of the primary interests on the part of the researcher in that role within the university. While this may be true as a general proposition, it may not be true in any individual case. Quality is, of course, often more important than quantity. This points up the need for careful scrutiny of actual situations, rather than blanket rules, except in the most serious kinds of cases, namely the ones Blumenthal identified as involving risks to patient care, to research integrity, and to students' careers.

Categories of remedies include disclosure, oversight, or some form of escape. Disclosure is usually warranted in all cases where a conflict of interest exists. Obviously no one can investigate and determine how serious is a given risk if they have no knowledge of the conflict to begin with. Even in the more controversial area of disclosure of research funding accompanying publications, one study showed that an overwhelming percentage of researchers who came out positively in favor of a certain type of drug received funding from companies that make the drug, while a much lower percentage of those critical of the class of drugs received such support. The authors of the study did not suggest that the researchers favoring the drugs were dishonest; however, they did recommend that disclosure occur "to avoid suspicion" (17). Critics of disclosure of funding for research that results in a publication claim that the reading public does not know what to do with the information (12). Whatever the merits of the two sides to that debate, there is no application to the question whether disclosure to university officials ought to be made. The better analogy here is to the lawyer in a conflict situation who must disclose the conflict to the client potentially affected. The client, after being informed, usually has the option to determine whether or not to continue to be represented by the lawyer or to ask for some change to be made. Analogously, the university has the right to know about a conflict situation, to determine whether further action is warranted or not.

Oversight may be by a standing committee within the university or by a person or group outside the university. Here the relationship between the researcher and the company may be welcomed, but the size and nature of the financial arrangement may cause sufficient concern that some additional regulation seems necessary. Since the 1995 federal guidelines require researchers to notify their own institutions if they or their close relatives have a financial interest in companies affected by the research exceeding $10,000 or 5 percent ownership, it is likely that the institutions will want to have some oversight of research that meet the federal criteria. Presumably whether that oversight is to be conducted by a committee within the university or outside it may depend on such variables as the expertise required to perform the oversight function or the manner in which the oversight must be conducted.

The last category of remedies, "escape," is a catchall to gather all those situations where it is deemed necessary or wise to prohibit the arrangement either through divestiture, abstention from decision making or some other mechanism to insulate the researcher from the work or the potential financial gain. This is obviously the most costly and serious remedy with which to handle a conflict situation; but, at times, it will be the only reasonable one.

INSTITUTIONAL CONFLICTS

Most of the literature concerning academic–industry relations has focused on the individual researcher and his or her own conflicts of interest, with the primary interest being to the academic institution and only secondary interests obliging the researcher to industry. Little attention has been paid to the problem of conflicts of interest that may affect the academic institution itself. Universities may have equity interests in companies affiliated with their institutions. Universities may also own patent rights that they license to companies and investigators employed by the institution. Institutional practices may thwart individual investigators or conspire with them to enhance the value of their equity interest or stock holdings to the detriment of their true primary interests. As the university is being asked to develop internal rules and procedures to guard against abuse in conflict situations, who will be the watchdog, guiding the university officials on the firing line from succumbing to the temptation to seek a secondary interest of the entity in preference to one or more of the institution's primary interests (18)?

Clearly, there are different conflict problems when the focus shifts from the individual researcher to the institution. First, the individual may not be at all compensated by a company, but the institution may be benefiting directly from the fruits of the research by increased value to its equity holdings or its licensing agreements. Again, no matter the situation with respect to the researcher, the institution may put subtle or not-so-subtle pressures to prefer a secondary interest in some way, which the individual researcher may be hard-pressed to avoid. The secondary interests of the institution may not be limited to economic gain either. Increased reputation may also be a secondary interest, which complicates the pursuit of the primary missions of teaching and research, and in the case of university clinical matters, the patients' welfare. There have already been suggestions made that any institutional conflict should prima facie be grounds for avoiding the conflict altogether. Of course similar remedies to those put forth for individuals have also been offered, with disclosure — even to individual patients — mandated, while internal or external monitoring providing supplementary remedies (18).

APPEARANCE OF A CONFLICT

A theme that constantly appears in all discussions of conflicts of interests is the problem of "appearance." Since conflicts are simply questions of risk, and what to do to prevent risks from ripening into actual breaches of duty, the question of appearances may arise more often in the remedy area than in determining whether or not a conflict exits. Once a conflict exits, there is, de facto, a risk that some impairment will follow. Even reasonable people may be skeptical that the remedy chosen *will truly prevent* the impairment from taking place. Since the university depends so much on its reputation for integrity in the pursuit of knowledge, these concerns are real and potentially problematic to deal with. However, it is not at all clear what should be done with appearance questions. In the absence of solid empirical evidence, it is hard to know whether any particular conflict category is seriously "risky." Nevertheless, there are those who believe the value of keeping the university's reputation clean requires curtailment of much of the activities that are now underway to foster academic–industry relationships (18).

OBSERVATIONS

Foundational to one's view of the risks and the necessary remedies for conflict of interest in corporate–university relationships is the public policy issue of the value of commercializing scientific research. During the past three decades the federal government has encouraged public–private collaborations through legislation and agency rule making. Scientists and universities, for whom the stakes may be very high economically, have encouraged and facilitated technology transfer. However, even for the strongest advocates of technology transfer and university/corporate relationships, 25 years of experience suggests that government, universities, and corporations need to remain diligent if they are going to preserve the benefits of commercialization of scientific ideas without sacrificing either scientific rigor or sound educational policy. Corporate–university agreements and federal rules need "to require disclosure of conflicts, limit the most troublesome forms of conflict, and create uniformity in ethical standards across the country" (19). The alternative seems to be an unacceptable slippery slope that threatens both the essence of the scientific process as well as the integrity of educational institutions.

On its face, technology transfer is relatively uncomplicated. Scientific researchers in universities and research institutes pursue new scientific knowledge with financial resources provided largely by the federal government. Commercially viable results are licensed to corporations who support the scientist and the scientist's institution in exchange for the opportunity to develop the idea. Scientists are often given personal consulting contracts or an equity interest or a board seat or a financial interest in future sales or all of these financial incentives by the company developing the scientific idea. The scientist's institution may also receive significant long-term financial incentives for its role in the process.

Inherently the problem is not the financial incentives for scientists or institutions to transfer technology. Within the free-enterprise system, scientists deserve the same opportunity for financial rewards from their work as do other professionals. Nor is the problem the personal gain for individual scientists and their institutions from work financed by public funds. As a matter of public policy, the federal government has determined that the benefits to society of the technology transferred outweigh the costs to taxpayers of allowing financial incentives to scientific researchers and their institutions. It is not an irrational or immoral trade-off.

The problem is the risks that standards of scientific inquiry will be compromised, diminished, or sacrificed for financial gain, or that the primary interests of educational institutions are subordinated to secondary interests. If the scientist's financial enrichment becomes tied to the success of the scientific outcome, then "society runs the risks that researchers will knowingly influence the outcome of 'neutral' scientific inquiries" (19). If the pursuit of economic gain harms students or patients, then the price is too high to pay.

Conflicts of interest threaten the integrity not only of individual scientists and their educational institutions but of the scientific process itself. Scientists caught by the potential for enormous financial gain between research for the public good and research for corporate interests, between their duties as teachers and mentors to graduate students and their duties as corporate consultants or officers, between their employer university or research institute and their corporate sponsor, or between the health interests of their research subjects and patients and the marketing interests of their corporate funders are on precariously thin ice when they record and report scientific results from an alleged position of scientific objectivity.

The classic remedies for conflict of interest of disclosure, oversight, and escape take on new meaning, and the path through the thicket of conflicting claims on the scientist's

objectivity and rigor is obscured as the potential for financial gain increases. Many would argue that scientific researchers, "particularly those conducting clinical trials, carry a fiduciary duty to the public. Like fiduciaries, there is a presumption against conflicts of interest" (19).

Current NIH and NSF rules rely on universities to decide whether there is a conflict of interest and what to do about it. Recent experiences, particularly in biomedical research, suggest that the university record to date leaves much to be desired. As pressure mounts on institutions and thus on scientists to seek corporate support and relationships, many argue that "the conflict of interest policies and standards of disclosure that universities rely upon don't do enough to protect academic freedom or the integrity of research in an environment where corporate interests are playing a growing role" (20).

Almost all current university and research institute policies are based on a principle of disclosure. But such disclosure principles are interpreted broadly, and often disclosure documents are not made public. Only a limited number of states require that disclosure documents be available under public record laws.

Very few scientific or lay publications inquire about conflicts or require disclosure by scientists writing on scientific and public policy topics. In 1992 a study of 14 journals showed that one out of three authors of nearly 800 scientific articles had a financial interest in the results of their research. Few if any of these conflicts were disclosed in the articles (20). In a 1997 study of scientists funded by drug companies it was found that 96 percent of the authors of favorable articles on a particular class of drugs had financial ties to the makers of the drugs. These conflicts were reported in only 2 of the 70 articles surveyed. Of those authors who published articles critical of the class of drugs, only 37 percent had financial conflicts of interest (17). While the mechanics of disclosure are neither easy nor obvious, it seems certain that institutions must find ways to make their disclosure requirements more visible and apparent. Moreover disclosure is only the first line of defense in conflict of interest situations when independence of judgment may be compromised.

It is not possible, or even desirable, to remove all conflicts of interest in scientific research. The benefits of collaboration between universities and corporations and the interactions of nonprofit and for profit scientists and investigators are already documented, particularly in biomedical research fields. But, as corporate relationships with universities, research institutes, and research scientists increase in number, size, and complexity, public confidence in the objectivity and rigor of the scientific process will erode rapidly in the face of undisclosed and unresolved conflicts of interest, real or perceived. Educational values and, in clinical settings, patient care may also be compromised by unattended conflicts of interests. The antidote is vigorous and persistent pursuit of institutional, governmental, corporate, and agency policies and practices that provide disclosure, oversight, and escape from conflicts of interest.

BIBLIOGRAPHY

1. R. Hofstader and W. Smith, eds., *American Higher Education: A Documentary History*, vol. 2, University of Chicago Press, Chicago, IL, 1961, pp. 568–569.
2. E. Shils, in A. Oleson and J. Voss, eds., *The Organization of Knowledge in Modern America*, 1860–1920, Johns Hopkins University Press, Baltimore, MD, 1979, p. 28.
3. R. Geiger, *To Advance Knowledge: The Growth of American Research Universities, 1900–1940*, Oxford University Press, Oxford, UK, 1986.
4. R. Geiger, *Research and Relevant Knowledge: American Research Universities Since World War II*, Oxford University Press, Oxford, UK, 1993.
5. Office of Technology Assessment (OTA), *New Developments in Biotechnology: U.S. Investment in Biotechnology*, OTA, Washington, DC, 1988.
6. N. Bowie, *University-Business Partnerships: An Assessment*, Rowman & Littlefield, Lankan, NJ, 1994.
7. J. Mervis, *Science* **269**, 294 (1995).
8. D. Thompson, *N. Engl. J. Med.* **329**(8), 573–576 (1993).
9. M. Davis, *Bus. Prof. Ethics J.* **12**(4), 21–41 (1993).
10. K. McMunigal, *Geo. J. Legal Ethics* **5**(4), 823–877 (1992).
11. M. Pritchard, *Acad. Med.* **71**(12), 1305–1313 (1996).
12. K. Rothman, *JAMA, J. Am. Med. Assoc.* **269**(21), 2782–2784 (1993).
13. M. Davis, *Bus. Prof. Ethics Jr.* **1**(1), 17–27 (1982).
14. J. Margolis, in T. Beauchamp and N. Bovice, eds., *Ethical Theory and Business*, Prentice-Hall, Englewood Cliffs, NJ, 1996, pp. 361–373.
15. N. Luebke, *Bus. Prof. Ethics J.* **6**(1), 67–81.
16. D. Blumenthal, *Acad. Med.* **71**(12), 1291–1296 (1996).
17. H. Stelfox, G. Choa, K. O'Rourke, and A. Detsky, *N. Engl. J. Med.* **338**(2), 101–106 (1998).
18. E. Emanuel and D. Steiner, *N. Engl. J. Med.* **332**(4), 262–267 (1995).
19. M. Witt and L. Gostin, *JAMA, J. Am. Med. Assoc.* **271**(7), 547–551 (1994).
20. G. Blumenstyk, *Chron. Higher Educ.* **44**(37), May 22, A41 (1998).

See other entries MEDICAL BIOTECHNOLOGY, UNITED STATES POLICIES INFLUENCING ITS DEVELOPMENT; TRANSFERRING INNOVATIONS FROM ACADEMIC RESEARCH INSTITUTIONS TO INDUSTRY: OVERVIEW.

SUPPLEMENT

Every effort was made to obtain articles for this Encyclopedia on all key organizations, government offices, industry groups, interest groups, and so on. However, it was impossible to obtain some of these entries in time for inclusion in this work. The following brief summaries are provided to call the readers' attention to important groups and organizations that treat aspects of the subjects covered by this Encyclopedia, as well as to provide suitable contact information. It is hoped that this will be helpful for the reader seeking additional information.

BIOTECHNOLOGY INDUSTRY ORGANIZATION (BIO)

BIO is a trade association for the biotechnology industry.

1625 K Street NW
Suite 1100
Washington, DC 20006
Tel: 202-857-0244
Fax: 202-857-0357
www.bio.org

The BIO Website includes information on the following topics: the biotechnology industry, the biotechnology record on ethics, legislative issues, biological warfare, biotechnology in agriculture, agricultural biotech products on the market, biotechnology in health care, approved biotechnology drugs, industrial uses of biotechnology, applications of industrial biotechnology, biotechnology for the environment, applications of environmental biotechnology, biotechnology in animal health, and marine biotechnology.

THE FOUNDATION ON ECONOMIC TRENDS (FET), headed by Jeremy Rifkin, is a nonprofit organization whose mission is to examine emerging trends in science and technology and their impacts on the environment, the economy, culture, and society.

The Foundation on Economic Trends (FET)
1660 L Street, NW, Suite 216
Washington, DC 20036
Tel: 202-466-2823
Fax: 202-429-9602
E-mail: office@biotechcentury.org
www.biotechcentury.org

The FET Website *www.biotechcentury.org* addresses environmental, social, economic, and ethical issues related to biotechnology and provides links to other organizations engaged in biotechnology issues.

HUMAN GENOME ORGANIZATION (HUGO)

HUGO is an international scientific organization.
Contact information (U.S. and Canada):

HUGO Americas
Laboratory of Genetics
National Institute on Aging
NIH/NIA-IRP. GRC, Box 31
5600 Nathan Shock Drive
Baltimore, MD 21224-6825, USA
Tel: 410-558-8337
Fax: 410-558-8331
E-mail: schlessingerd@grc.nia.nih.gov
www.gene.ucl.ac.uk/hugo

Contact information (international):

HUGO
142-144 Harley Street
London W1N 1AH
United Kingdom
Tel: (44) 171 935 8085
Fax: (44) 171 935 8341
E-mail: hugo@hugo-international.org

The Human Genome Organization (HUGO) is the international organization of scientists involved in the Human Genome Project (HGP), the global initiative to map and sequence the human genome. HUGO was established in 1989 by a group of the world's leading genome scientists to promote international collaboration within the project.

HUGO carries out a complex coordinating role within the Human Genome Project.

HUGO activities range from support of data collation for constructing genetic and physical maps of the human genome to the organization of workshops to promote the consideration of a wide range of ethical, legal, social, and intellectual property issues. HUGO fosters the exchange of data and biomaterials, encourages the spreading and sharing of technologies, provides information and advice on aspects of human genome programs, and serves as a coordinating agency for building relationships between various governmental funding agencies and the genome community. HUGO provides an interface between the Human Genome Project and the many groups and organizations interested or involved in the human genome initiative.

PHARMACEUTICAL RESEARCH AND MANUFACTURERS OF AMERICA (PhRMA)

1100 Fifteenth St. NW
Washington, DC 20005
www.phrma.org
President: Alan F. Holmer

PhRMA membership consists of approximately 100 U.S. companies that have a primary commitment to pharmaceutical research. The mission of the pharmaceutical Research and Manufacturers of America is to help the research-based pharmaceutical industry successfully meet its goal of discovering, developing, and bringing to market medicines to improve human health, patient satisfaction, and the quality of life around the world, as well as to reduce the overall cost of health care.

To achieve its goal, the industry aspires to foster a favorable environment that encourages innovative drug research; swift development and approval of safe and effective drugs; consumer and patient access to medicines in an open and competitive marketplace; support and understanding from the public and other key constituents regarding the critical role and value of the pharmaceutical industry in improving human health and quality of life and in reducing overall health care costs; public policies that allow sufficient returns to foster continued innovation.

U.S. DEPARTMENT OF COMMERCE, TECHNOLOGY ADMINISTRATION

The Technology Administration (TA) (*www.ta.doc.gov*) is a bureau of the U.S. Department of Commerce (*www.doc.gov*). The Technology Administration leads civilian technology for the Department of Commerce and works with U.S. industries to promote U.S. economic competitiveness and growth.

The Undersecretary for Technology supported by the Deputy Undersecretary for Technology, manages the Technology Administration's (TA) three agencies:

(1) The Office of Technology Policy (OTP) is an office of the federal government with the explicit mission of developing and advocating national policies that use technology to build America's economic strength.

(2) The National Institute of Standards and Technology (NIST) promotes economic growth and an improved quality of life by working with industry to develop and apply technology, measurements, and standards.

(3) The National Technical Information Services (NTIS) collects and disseminates scientific, technical, engineering, and related business information produced by the U.S. government and foreign sources.

WHITE HOUSE OFFICE OF SCIENCE AND TECHNOLOGY POLICY (OSTP)

1600 Pennsylvania Ave N.W.
Washington, DC 20502
Tel: 202-395-7347
E-mail: information@ostp.eop.gov
www.whitehouse.gov, then link White House Offices and Agencies

OSTP was established in 1976 to provide the President with policy advice and to coordinate the science and technology investment.

OSTP Divisions: Environmental Division; National Security and International Affairs (NSIA) Division; Science Division; and Technology Division.