

RESEARCH

Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel



OPEN ACCESS

Nicholas Graves *professor of health economics*¹, Adrian G Barnett *associate professor*¹, Philip Clarke *associate professor*²

¹School of Public Health and Institute for Health and Biomedical Innovation, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, Qld 4059, Australia; ²School of Public Health, University of Sydney, NSW, Australia

Abstract

Objective To quantify randomness and cost when choosing health and medical research projects for funding.

Design Retrospective analysis.

Setting Grant review panels of the National Health and Medical Research Council of Australia.

Participants Panel members' scores for grant proposals submitted in 2009.

Main outcome measures The proportion of grant proposals that were always, sometimes, and never funded after accounting for random variability arising from differences in panel members' scores, and the cost effectiveness of different size assessment panels.

Results 59% of 620 funded grants were sometimes not funded when random variability was taken into account. Only 9% (n=255) of grant proposals were always funded, 61% (n=1662) never funded, and 29% (n=788) sometimes funded. The extra cost per grant effectively funded from the most effective system was \$A18 541 (£11 848; €13 482; \$19 343).

Conclusions Allocating funding for scientific research in health and medicine is costly and somewhat random. There are many useful research questions to be addressed that could improve current processes.

Introduction

Grant funding agencies strive to support the best scientific research. Peer review is used to decide who gets funded by including the opinions of experts, but problems with peer review mean decisions might not be reliable.

Considerable research shows poor quality in peer review for scientific journals. The authors of a Cochrane review found little empirical evidence to support the use of editorial peer review as a mechanism to ensure quality of biomedical research.¹ Reviewers at the *Annals of Internal Medicine* failed to pick two thirds of deliberate errors.² Blinding reviewers to the authors and the origin of the manuscript or requiring them to sign the peer review report had no effect on the detection rate for errors.³ Using reviewers suggested by authors over those selected by editors failed to improve the quality of peer review for journals.⁴ A short training programme to improve peer review had a slight effect, which disappeared after six months.⁵ A randomised trial showed that the performance of reviewers was improved only slightly with different types of training intervention.⁶

Surprisingly, given the impact that decisions on research funding have on academic careers, relatively little research has been done into the peer review of grant applications.⁷ Thirty two proposals for funding from the McGill University Health Center Research Institute were assessed by two competing processes: a traditional review committee of 11 members and independent reviews from a committee member and content expert.⁸ Agreement was poor and chance associated with funding decisions was considerable. The authors of a study of 248 grant proposals compared two similar peer review processes for ranking proposals.⁹ They found that agreement beyond chance was only fair (Cohen's $\kappa=0.29$) and that when proposals were organised into one of two categories, clearly fundable and not clearly fundable, agreement improved (Cohen's $\kappa=0.44$). One study¹⁰ estimated that 38 416 reviews per grant would be required to get a high precision for the peer review and selection

Correspondence to: N Graves n.graves@qut.edu.au

Extra material supplied by the author (see <http://www.bmj.com/content/343/bmj.d4797/suppl/DC1>)

Project grant funding process of National Health and Medical Research Council of Australia
Explanation of non-parametric bootstrap procedure

of grants from the National Institutes of Health. Without high precision some proposals will inevitably be incorrectly ranked and may undeservedly miss out on funding. An analysis of reviewers' scores for grant proposals to the National Institutes of Health showed that adjusting for uncertainties and biases among reviewers would lead to a 25% change in the pool of funded proposals.¹¹ Similarly, a study of proposals submitted to the National Science Foundation re-reviewed 75 funded and 75 not funded grant proposals, and for 25% of proposals the funding decision changed.¹²

Lack of reliability in funding decisions might arise from variation in peer review among external assessors and members of grant review panels. A journalist observed a grant review panel for the American Cancer Society and suggested that decisions were strongly influenced by individual preference about whether the research was important; the pressure to find tiny flaws in a grant proposal, so that it might be excluded; and the ability of one or two reviewers to change the preferences of others, a "cheerleader" effect.¹³ One researcher suggested that using review panels for grants was unreliable.¹⁴ He proposed that panel members focused on reasons to reject proposals, some competed with each other to show their intellectual prowess, and some had limited knowledge about the proposals' methods. The membership of a grant review panel is somewhat random and depends on who is invited and is available. This will affect funding decisions because the preferences, personalities, and knowledge of members will vary. Regardless of the panel's membership, strong and weak grant proposals should be identified consistently, but most proposals are likely to occupy a tightly packed middle ground. These proposals are the most difficult to separate and a slight change in score can push a proposal below or above a funding line.

Although previous studies have established that the assessment of grants is often subject to relatively low inter-rater agreement, the impact this variability has on actual decisions about grant funding has received less attention. Our focus is on the variability in decisions and not on their validity. We addressed this by using scores of grant panel members from de-identified data supplied by the National Health and Medical Research Council of Australia. We also estimated the cost effectiveness of changing the size of the grant review panel. The impetus for this research was that success rates for grant proposals are falling worldwide.¹⁵ Success rates for the UK Engineering and Physical Sciences Research Council fell from 43% in 2000 to 26% in 2008.¹⁵ In Australia success rates for the National Health and Medical Research Council dropped from 30% in 2000 to 23% in 2010.¹⁶ Unsuccessful applicants experience the double jeopardy of a blow to their career and the cost of participating in a lengthy application process. Our findings should be useful for researchers considering whether to apply for a grant and funding agencies looking to improve their review processes. This research is timely, as funding agencies are expecting 5-10 years of flat or contracting research budgets as large national debts are reduced by spending cuts.¹⁷

Methods

The National Health and Medical Research Council of Australia committed 50.3% of its annual \$A714m (£452m; €514m; \$745m) budget to the project grants scheme in 2009. Proposals are unsolicited and cover most health and medical issues. Applications are between 70 and 120 pages, including a nine page research plan. The web extra shows the process.

Reliability in funding

We obtained the category and summary scores for all project grants considered in the 2009 funding round. Proposals had been assessed by one of 45 discipline specific review panels of between seven and 13 members. Each panel scored between 42 and 92 proposals. The average score across all panel members is calculated and applicants are successful if their score is above a funding line (see web extra).

We estimated the variability in panel members' scores and examined how this variability translated into the variability in ranks, and hence variability in decisions on funding. For each grant we estimated the 90% confidence interval for the rank and its minimum and maximum rank. We did this using a non-parametric bootstrap procedure because we could not assume that the panel scores followed a normal distribution. The variability in a proposal's score is estimated by re-sampling original scores from panel members with replacement to generate a slightly different panel (see web extra). After estimating the range in ranks we grouped the proposals into three categories: never funded, if the maximum rank was below the funding line; always funded, if the minimum rank was on or above the funding line; and sometimes funded, if the range in ranks straddled the funding line. We used each proposal's 90% confidence interval to determine whether it was effectively funded. If the lower limit of the 90% confidence interval was above the funding line then the proposal was correctly funded (effectively funded). If the upper limit of the 90% confidence interval was below the funding line then a proposal was correctly rejected (also effectively funded); the 90% value is arbitrary. We also assessed the effect of changing the size of review panels by re-sampling different numbers of scores to represent panels of seven, nine, and 11 members.

Cost and effectiveness

The costs of the grant allocation processes are shared by researchers who prepare proposals, peer reviewers who are either external reviewers or members of grant review panels, and the National Health and Medical Research Council which administers the scheme. To understand preparation costs, one week after the 2009 closing date we surveyed 42 chief investigators from two medical research institutes in Brisbane, Australia who led 54 grant proposals. They were asked to estimate how many working days they spent preparing their grant proposal. Their salary grade and the full costs of employment were available from the institutions' websites. To estimate peer review costs we assumed average reading times per grant proposal of four hours for external peer reviewers and primary spokespeople and two hours for secondary spokespeople. We assumed that review panel members spent 20 minutes reading each of the other grant proposals and 46 hours working during the week the review panel met. These estimates are based on our experiences as grant review panel members and external reviewers and the opinions of our colleagues. The National Health and Medical Research Council of Australia provided the information on the cost of administering the scheme. It reported the costs of extra staffing, booking the hotel and conference facilities for the review panel, travel, and sitting fees.

Costs were expressed in Australian dollars and the effectiveness outcome by the proportion of grants effectively funded. We estimated both for a panel of seven, nine, and 11 members, and we calculated the incremental cost per extra grant effectively funded.

Results

Observed scores and range in ranks

Overall, 2983 grant proposals were submitted in 2009. Of these, 278 were under a special initiative and were excluded from analysis because they were assessed using different criteria. Of the 2705 that remained, 620 (23%) were funded, ranging from 9% to 38% across the 45 panels. Table 1 shows the impact of the variability in panel members' scores on funding decisions. Overall, 59% of the 620 originally funded proposals were sometimes not funded when the effect of random variability was included. Eighty per cent (n=1662) of the 2085 grant proposals not originally awarded were never funded. In total only 9% (n=255) of grants were always funded, 61% (n=1662) were never funded, and 29% (n=788) were sometimes funded. Figure 1 shows how these proportions changed with different panel sizes. Adding panel members increased reliability, as shown by the shrinking percentage of proposals in the sometimes funded category.

The range in ranks was plotted for the panel with the largest proportion of proposals that were sometimes funded (fig 2) and the panel with the smallest proportion of proposals that were sometimes funded (fig 3). Figure 2 shows a wide variability in ranks, and only one proposal was always funded. The rank for the proposal marked with an asterisk ranged from the fifth best to the worst. In contrast the ranges in ranks for figure 3 are much tighter, meaning that the reliability in decisions in the panel was greater and hence the proportion of always funded and never funded proposals was higher. A cluster of six grants was scored the same by every panel member and were all always funded.

Costs and effectiveness

Most researchers spent between 20 and 30 days preparing their grant proposal, with a median 22 (interquartile range 20–32) days per grant proposal. On two of 54 occasions the lead investigator spent more than 65 days preparing the application and on five of 54 occasions the chief investigator spent less than 15 days (although these researchers indicated that they had shifted preparation costs on to junior researchers). The total costs of the funding exercise were \$A47.87m with 85% (\$A40.85m) incurred by applicants, 9% (\$A4.44m) for peer review by external assessors and review panel members, and 5% (\$A2.59m) to administer the scheme. Both effectiveness and costs increased with larger panels (table 2).

Discussion

The assessment of grant proposals is costly and subject to a high degree of randomness owing to variation in panel members' assessments. The relatively poor reliability in scoring by panels might be expected given the complexity of the task and the subjective nature of the assessment process. The degree of reliability varied greatly between panels (figs 2 and 3) suggesting that in some disciplines review panels find it more difficult than in others to agree on a proposal's quality. The total cost per proposal was \$A17 744, with around 85% incurred by applicants. The median estimate of 22 days preparing a grant multiplied by the 2983 grant proposals submitted shows that in 2009 180 years of researcher time was used up. The costs per grant proposal were similar in the United Kingdom, at £196m a year, or £9797 (\$A15 676) per proposal.¹⁸

The benefits of participating vary for the applicants: those who score in the top 9% are always funded (table 1), the next 29% face uncertainty and may or may not be rewarded for their

efforts, and the remaining 61% face certain outcomes of zero as the variation among assessors was insufficient for them to ever score above the funding line. For this last group the time invested in the process is likely to be a deadweight loss other than some process utility from writing the grant and from participating in peer review. Applicants in this group might have benefited more from doing something else with their time.

Reliability can be increased by using the most effective system of 11 panel members. This is probably worth while as the extra cost per extra grant effectively funded is \$A18 541, only 3% of the average grant value awarded in 2009.

Limitations of the study

The scores of panel members are unlikely to be independent. Each grant is assigned a primary and secondary spokesperson who would tend to lead the panel discussion (see web extra). It would be unusual if every panel member reviewed every grant in detail, and we think most panel members are grateful for the opportunity to listen to the views of the assigned spokesperson. This does not mean that panel members will not hijack or rescue a grant they believe is being inappropriately judged. The direction and magnitude of any dependence, however, could not be tested from the available data. Importantly, this study has only considered one aspect of variation: variation owing to panel members' scores in relation to the funding line. Also, the costs of preparing grants were elicited from a small sample of researchers based at two institutions, and a nationally representative sample would more accurately reflect costs.

Supply and demand side regulation

Given the high costs associated with research funding schemes it is worth considering other approaches by implementing supply and demand side regulations to change the rules of funding. A supply side measure is to impose a production quota. Each applicant might be limited to one grant proposal per round; at present applicants are limited to holding a maximum of six project grants. A concern from using a quota system is that creativity is stifled and only safe proposals are submitted.¹⁸ However, the high level of competition for funding and the tendency for review panels to be risk averse compared with individuals suggests pioneering and so potentially risky proposals are unlikely to be funded anyway.¹⁹ This type of production quota is likely to thwart researchers who are particularly skilled at winning grants. An alternative is a targeted production quota that excludes unsuccessful applicants for a cooling off period as implemented by the UK Engineering and Physical Sciences Research Council. Disgruntled applicants complained of unfairness and of being singled out.¹⁵ This system would reduce total proposals but not shut out the researchers skilled at winning grants. Support for this approach came from an editorial in *Nature* titled "Tough Love" whose author recognised that falling rates of success mean some researchers will incur large deadweight costs from which they might be protected.²⁰ A demand side intervention is to inform applicants about the probability their grant was funded using the results generated by the bootstrap procedure, described in this paper and the web extra. If failed applicants are shown they always ranked below the funding line then they might decide it is not worth putting any more time into re-submitting the proposal. The application procedure could also be simplified. Currently, proposals are between 70 and 120 pages long, with only nine used for the research plan. A reduction in paperwork will save the costs of preparation and peer review and should make it easier to recruit assessors.²¹ A limit may, however, apply to which there is scope for changing application processes as much

of the information requested is mandated by government, meaning the National Health and Medical Research Council of Australia would have to negotiate for its removal.

Further research

Further research might examine other sources of variation. An interesting experiment would be to give the same set of grants to two independent review panels whose members did not know whether their scores were going to be used to inform the funding decision. Our hypothesis is that the variability in funding decisions between these two panels would be greater than we found from our analysis. This experiment might be repeated for different disciplines, such as cell biology or public health, given the large differences between panels we observed. The same design could be used to assess a shortened application and panel process to see whether similar levels of reliability could be achieved at lower cost. A short application process would be easier to apply for and easier to review and may take less administration time. Reliability may even increase as external reviewers and panel members have less information to synthesise. Other systems could be tested for reliability and cost such as a journal style approach where grant proposals are submitted to a subeditor who makes an initial cull. Survivors are reviewed externally and are then considered by expert editors and a recommendation made to fund or not. An initial cull is already used by some grant agencies, by only asking for detailed proposals from those proposals that make it through the first round. Although these experiments will take time and effort, the costs of a research programme to look at ways of improving the funding process is likely to represent only a small fraction of the total research money allocated each year through competitive funding mechanisms.

Another avenue for investigation would be to assess the formal inclusion of randomness. There may be merit in allowing panels to classify grants into three categories: certain funding, certain rejection, or funding based on a random draw for proposals that are difficult to discriminate. Random allocation has been used in the assignment of medical places in the Netherlands to increase diversity in student backgrounds,²² but we are not aware of it ever being used in research funding. It may save costs too by reducing the duration of panels' discussion. In 1998 Greenberg²³ called for 15-20% of funding by the National Institutes of Health to be allocated by lottery saying "instead of dodging the fact that chance plays a big part in awarding money, the system will sanctify chance as the determining factor."

Conclusions

Prospectively funding individuals for their research is fraught with problems that can be reduced but not eliminated. Retrospective assessment of actual performance may be a better system.²⁴ It could be based on research productivity and broader health impacts. Funding could be allocated using quantifiable evidence rather than promises, including published papers, new policies, or observable improvements in health. Anecdotal evidence suggests researchers skilled at winning funding already use this approach by completing most of the research activity before applying for funding. There would of course have to be a reasonable programme of seed funding to get talented researchers started.

Given the impact research funding decisions can have on academic careers,²⁵ it is surprising that relatively little research has been carried out into its processes.⁷ This study tackled the degree of random variation due to differences in the ratings of grant review panel members, and the costs of the process. We

found that random variation affected many proposals that were assessed by an Australian medical research council in a single funding year. This information is useful and represents a starting point for additional research to understand the degree of variability in funding decisions, its causes and consequences, and how funding processes might be improved.

We thank the National Health and Medical Research Council of Australia for allowing us to access their peer review data as part of their commitment to improving their reviewing processes, and the 42 researchers who completed our survey on preparation times.

Contributors: All authors conceived and designed the study, analysed and interpreted the data, drafted the article or revised it critically for important intellectual content, approved the version to be published, and are the guarantors.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: This research has not been approved by an ethics committee. We believe the data are non-threatening, are reported in summary, do not identify any individual, and have been used to increase knowledge of how health and medical research is funded.

Data sharing: No additional data available.

- 1 Jefferson T, Rudin M, Brodney Folse S, Davidoff F. Editorial peer review for improving the quality of reports of biomedical studies. *Cochrane Database Syst Rev* 2007;(2):MR000016.
- 2 Baxt WG, Waeckerle JF, Berlin JA, Callahan ML. Who reviews the reviewers? Feasibility of using a fictitious manuscript to evaluate peer reviewer performance. *Ann Emerg Med* 1998;32:310-7.
- 3 Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA* 1998;280:237-40.
- 4 Schroter S, Tite L, Hutchings A, Black N. Differences in review quality and recommendations for publication between peer reviewers suggested by authors or by editors. *JAMA* 2006;295:314-7.
- 5 Schroter S, Black N, Evans S, Carpenter J, Godlee F, Smith R. Effects of training on quality of peer review: randomised controlled trial. *BMJ* 2004;328:673.
- 6 Schroter S, Black N, Evans S, Godlee F, Osorio L, Smith R. What errors do peer reviewers detect, and does training improve their ability to detect them? *J R Soc Med* 2008;101:507-14.
- 7 Demicheli V, Di Pietrantonj C. Peer review for improving the quality of grant applications. *Cochrane Database Syst Rev* 2011;2:MR000003.
- 8 Mayo NE, Brophy J, Goldberg MS, Klein MB, Miller S, Platt RW, et al. Peering at peer review revealed high degree of chance associated with funding of grant applications. *J Clin Epidemiol* 2006;59:842-8.
- 9 Hodgson C. How reliable is peer review? An examination of operating grant proposals simultaneously submitted to two similar peer review systems. *J Clin Epidemiol* 1997;50:1189-95.
- 10 Kaplan D, Lacetera N, Kaplan C. Sample size and precision in NIH peer review. *PLoS One* 2008;3:e2761.
- 11 Johnson VE. Statistical analysis of the National Institutes of Health peer review system. *Proc Natl Acad Sci USA* 2008;105:11076-80.
- 12 Cole S, Cole JR, Simon GA. Chance and consensus in peer review. *Science* 1981;214:881-6.
- 13 Powell K. Research funding: making the cut. *Nature* 2010;467:383-5.
- 14 Birkett NJ. The review process for applied-research grant proposals: suggestions for revision. *CMAJ* 1994;150:1227-9.
- 15 Van Noorden R, Brumfiel G. Fixing a grant system in crisis. *Nature* 2010;464:474-5.
- 16 Australian Government. NHMRC research funding facts book. 2010. www.nhmrc.gov.au/_files_nhmrc/file/publications/synopses/nh138.pdf.
- 17 Van Noorden R. UK government warned over "catastrophic" cuts. *Nature* 2010;466:420-1.
- 18 Research Councils UK. Report of the Research Councils UK Efficiency and Effectiveness of Peer Review Project. 2006. www.rcuk.ac.uk/documents/documents/rcukrpreport.pdf.
- 19 Gong M, Baron J, Kunreuther H. Why do groups cooperate more than individuals to reduce risks? 2010. http://opim.wharton.upenn.edu/risk/library/WP201002-MG_JB_HK_WhyCooperate.pdf.
- 20 Anon. Tough love. *Nature* 2010;464:465.
- 21 Schroter S, Groves T, Hojgaard L. Surveys of current status in biomedical science grant review: funding organisations' and grant reviewers' perspectives. *BMC Med* 2010;8:62.
- 22 Bennett M, Wakeford R. Health policy, student selection and curriculum reform. *Health Policy Educ* 1982;3:173-81.
- 23 Greenberg DS. Chance and grants. *Lancet* 1998;351:425.
- 24 Abrams PA. The predictive ability of peer review of grant proposals: the case of ecology and the US National Science Foundation. *Soc Stud Sci* 1991;21:111-32.
- 25 Melin G, Danell R. The top eight percent: development of approved and rejected applicants for a prestigious grant in Sweden. *Science and Public Policy* 2006;33:702-12.

What is already known on this topic

Health and medical research aims to progress evidence based medicine, but decisions about which proposals to fund are not grounded in evidence

There is a shortage of research in this area

The best research proposals should be chosen for funding

What this study adds

Decisions about funding health and medical research are somewhat random

Applicants bear the most costs because of the length of time needed to prepare a proposal

Larger panels are better than smaller ones

Accepted: 29 June 2011

Cite this as: *BMJ* 2011;343:d4797

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: <http://creativecommons.org/licenses/by-nc/2.0/> and <http://creativecommons.org/licenses/by-nc/2.0/legalcode>.

Tables

Table 1 | Impact of variability in review panel members' scores on funding decisions

Alternatively awarded	No (%) not originally awarded	No (%) originally awarded	Total No (%)
Never	1662 (80)	0	1662 (61)
Sometimes	423 (20)	365 (59)	788 (29)
Always	0	255 (41)	255 (9)
Total	2085	620	2705

Table 2| Costs and effectiveness of different sized panels

Funding process	Total costs (\$A)	Grants effectively funded	Increase in cost (\$A)	Extra grants effectively funded	Incremental cost effectiveness ratio* (\$A)
7 panel member	46 276 513	2200			
			797 258	53	15 043
9 panel members	47 073 771	2253			
			797 258	43	18 541
11 panel members	47 871 029	2296			

\$A1.00 (£0.66; €0.75; \$1.10).

*Extra cost per grant effectively funded from choosing next most effective funding process.

Figures

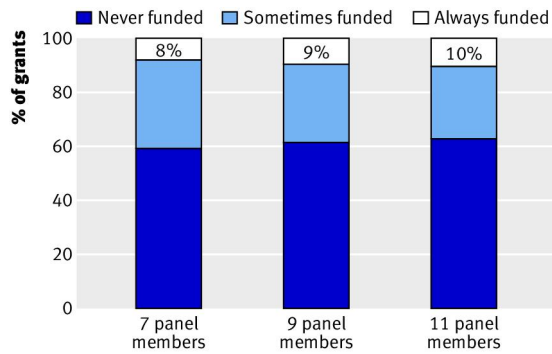


Fig 1 Proportion of grants that were never funded, sometimes funded, and always funded if review panels had seven, nine, or 11 members

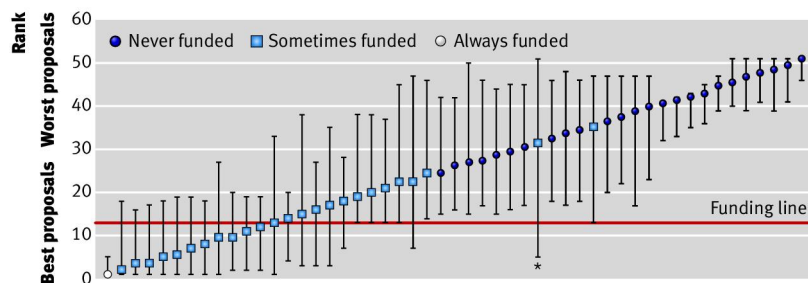


Fig 2 Range in ranks for grants assessed by review panel with largest proportion of sometimes funded proposals

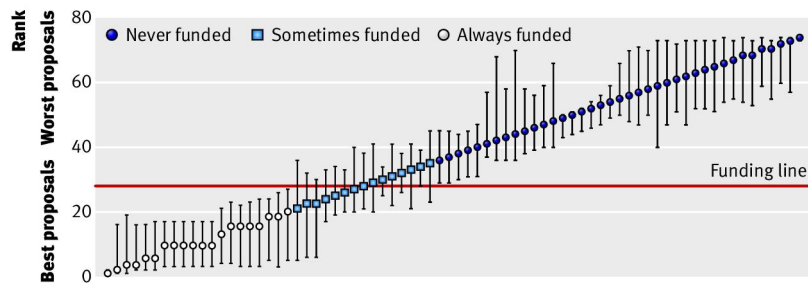


Fig 3 Range in ranks for grants assessed by review panel with smallest proportion of sometimes funded proposals