Meta-analysis of All-Cause Mortality According to Serum 25-Hydroxyvitamin D

We examined the relationship between serum 25-hydroxyvitamin D (25(OH)D) and all-cause mortality. We searched biomedical databases for articles that assessed 2 or more categories of 25(OH)D from January 1, 1966, to January 15, 2013. We identified 32 studies and pooled the data.

The hazard ratio for all-cause mortality comparing the lowest (0–9 nanograms per milliliter [ng/mL]) to the highest (> 30 ng/mL) category of 25(OH)D was 1.9 (95% confidence interval = 1.6, 2.2; \(P < .01\)). Serum 25(OH)D concentrations less than or equal to 30 ng/mL were associated with higher all-cause mortality than concentrations greater than 30 ng/mL (\(P < .01\)).

Our findings agree with a National Academy of Sciences report, except the cutoff point for all-cause mortality reduction in this analysis was greater than 30 ng/mL rather than greater than 20 ng/mL. (Am J Public Health. 2014;104:e43–e50. doi:10.2105/AJPH.2014.302034)

AN INVERSE ASSOCIATION was proposed between solar irradiance and incidence of colon and breast cancer, based on a mechanism involving insufficient vitamin D. Individuals with lower serum 25-hydroxyvitamin D (25(OH)D) have higher risk of breast and colon cancer, specific cancers, and coronary heart disease. Physiological mechanisms for the inverse association of 25(OH)D with cancer have been reported.

Despite research on the association between low vitamin D status and many diseases, no consensus has emerged on the optimal serum 25(OH)D concentration. The concern is whether it is safe to maintain serum 25(OH)D concentrations in the range high enough to prevent some types of cancers and coronary heart disease.

We decided to analyze the strength and consistency of the inverse association between levels of serum 25(OH)D and age-adjusted mortality hazard ratios in a rapidly expanding field of public health. A previous meta-analysis summarized 12 studies, another summarized 14, and another summarized a broader range.

We hypothesized that lower serum 25(OH)D was associated with higher all-cause mortality hazard ratios, and defined the age-adjusted hazard ratio for death from any cause as the outcome addressed by the meta-analysis. This analysis includes all studies of all-cause mortality hazard ratios by categories of serum 25(OH)D in healthy or general medical clinic cohorts that met the eligibility criteria. Twenty new studies of serum 25(OH)D and all-cause mortality entered the literature since the Zittermann et al. review, for a total of 32 in this review.

Two investigators (J. J. K. and S. B. M.) searched the biomedical literature for observational studies of serum 25(OH)D concentrations in association with age-adjusted all-cause mortality, published and indexed in databases between January 1, 1966, and January 15, 2013. The searches were an experienced epidemiologist with a PhD in public health and a research scientist (J. K.). They found 88 studies (Figure 1). The inclusion criterion was any study reporting a measure of association according to 2 or more categories of serum 25(OH)D concentration. Of these, 32 studies met the inclusion criterion. The remaining 56 studies were ineligible for inclusion because they did not report an association by categories of serum 25(OH)D.

We calculated age-adjusted hazard ratios comparing the
lowest with the highest quantile of serum 25(OH)D concentration by using the O-E and variance method for combining studies, an application of the Peto method.\textsuperscript{52} We calculated the overall O-E and variance by using the individual hazard ratios from each study and their 95% confidence intervals.\textsuperscript{53}

We calculated the \( P \) value for the overall summary odds ratio by using a \( z \)-score, where the numerator was the natural logarithm of the pooled hazard ratio and the denominator was the standard error of the natural logarithm of the pooled hazard ratio. This is a standard method for calculating the \( P \) value with Peto’s assumption-free method.\textsuperscript{54} Hazard ratios comparing the highest with the lowest quantities for each study were displayed in a forest plot.\textsuperscript{55,56} We used the DerSimonian–Laird statistic with a random effects model to assess heterogeneity among studies.\textsuperscript{57}

To provide a pooled estimate of the dose–response relationship between serum 25(OH)D and all-cause mortality, we determined the ratio of the age-adjusted hazard rates for each 10 nanograms per milliliter (ng/mL) stratum of 25(OH)D, compared with the lowest stratum of 25(OH)D in the study. We considered the highest stratum, typically 30 ng/mL or higher, the reference stratum. We plotted the mean of the hazard ratios for all studies for each stratum of 25(OH)D. We calculated 95% confidence intervals (CIs) for the pooled hazards ratios with standard procedures.\textsuperscript{58–60} The hazard ratios that were plotted were the age-adjusted hazard rate in each stratum divided by the age-adjusted hazard rate in individuals in the top stratum. We also calculated means weighted by the inverse of the variance of each study. This estimate was dominated by one study\textsuperscript{49} that was much larger than the others. Therefore, we used unweighted means of the hazard ratios, because they provided a combined dose–response curve not determined mainly by a single large study.\textsuperscript{19} We also fitted a standard decreasing exponential curve to the results, using the function \( f = y_0 + a \cdot e^{b \cdot x} \). We determined the values of \( a \) and \( b \) by using multiple regression.

The rationale for selecting age-adjusted mortality hazard ratios as the data for this analysis was that these directly address the hypothesis of the study. It is routine to use age-adjusted mortality hazard ratios to take into account differences that may exist in the age distributions of the study populations. Two investigators independently extracted data from the articles, and there was a 100% cross-check between investigators of the accuracy of data extraction. All hazard ratios were age adjusted, reducing the chances of confounding by age.

This review was conducted according to recommendations of the Meta-Analysis of Observational Studies in Epidemiology Group,\textsuperscript{51} including providing subjective comments regarding the quality of each study (Table A, final column). Inclusion criteria were reported according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.\textsuperscript{62} Several journal articles became available after this study was completed. These are summarized in Table B (available as a supplement to this article at http://www.ajph.org). Their findings are consistent with the association found in the 32 studies that were analyzed. Their inclusion would not have substantially changed the conclusions of this meta-analysis.

**RESULTS**

Twenty-five studies identified a significant inverse relationship between 25(OH)D concentration and age-adjusted all-cause mortality hazard rates (Figure 2). In 5 studies,\textsuperscript{21,34,35,37,47} an inverse trend was present, but was not statistically significant. In 2 studies, 1 in the United States\textsuperscript{48} and 1 in Linxian, China,\textsuperscript{50} no association was seen. The studies are described and evaluated in Table A.

The overall age-adjusted hazard ratio for all-cause mortality comparing the lowest (0–9 ng/mL) to the highest (>50 ng/mL) categories of 25(OH)D concentration was 1.9 (95% CI = 1.6, 2.2; \( P < .001 \); Figure 2). Serum concentrations less than or equal to 30 ng/mL also were associated with higher all-cause mortality when compared with those greater than 30 ng/mL (\( P < .01 \)). The studies were not homogeneous (DerSimonian–Laird \( \chi^2 = 183.66; df = 28; P < .001 \)). A funnel plot analysis (not shown) revealed no indication of publication bias.

A pooled dose–response curve showed that the hazard ratios declined steeply between 0 ng/mL and 30 to 39 ng/mL, then appeared to plateau at serum 25(OH)D concentrations above 50 ng/mL (Figure 3). The curve was steep at lower concentrations through 30 to 39 ng/mL, with a slight trend toward lower risk at 40 to 49 ng/mL compared with 30 to 39 ng/mL. There was a reduction in the hazard ratio of 0.4 units from 0 to 9 ng/mL through 10 to 19 ng/mL, 0.3 units from 10 to 19 ng/mL through 20 to 29 ng/mL, and approximately 0.1 units from 20 to 29 ng/mL through 30 to 39 ng/mL. Overall, there was a mean drop of 0.1 units in the hazard ratio per 10 ng/mL of 25(OH)D.

According to a decreasing exponential curve fit to the data from all studies combined, the point at which the estimated hazard ratio
was no longer significantly different from 1.0 was a serum 25(OH)D level of 36 ng/mL. The relative hazard at that point was 1.09, and its 95% confidence limits were 1.0 and 1.2. The dose–response curve for each study is shown in Figure 4.

Because one possible source of heterogeneity among studies was differences in the length of follow-up, we analyzed whether length of follow-up was related to the association between 25(OH)D and all-cause mortality. We sorted the studies into 2 categories of length of follow-up with respect to the median length of follow up of 7 years. The mean hazard ratio for studies of less than 7 years duration was 2.2, whereas it was 1.6 for studies of greater than or equal to 7 years duration ($P < .05$).

**DISCUSSION**

This study confirmed an inverse association between serum 25(OH)D concentrations and age-adjusted all-cause mortality rates. Overall, individuals whose 25(OH)D concentrations were in the lowest quantile (0–9 ng/mL) had nearly twice the age-adjusted death rate as those in the highest quantile (>35 ng/mL).

Ingestion of some minerals, such as boron, may reduce the rate of catabolism of 25(OH)D, potentially resulting in accumulation of 25(OH)D that could create an appearance of higher mortality rates at higher 25(OH)D concentrations, because of the presence in the population of sick individuals who may be unintentionally consuming such minerals in supplements. However, in healthy populations, such as the Rancho Bernardo cohort, higher hazard ratios were not present at higher 25(OH)D concentrations. Overall, there was no harm associated with being in the top quantile of 25(OH)D, which was generally 35 to 40 ng/mL or higher. Some degree of caution may be reasonable at higher 25(OH)D concentrations until more data pertinent to higher concentrations are available.

**Adjustments**

All studies but 1 adjusted for age, and that study stratified by age. Many studies adjusted for other covariates. These adjustments revealed that the hazard ratios were not much affected by adjustment. For example, 7 of the 32 studies adjusted for body mass index, but body...
mass index did not appear to play a major role in the association. Six studies adjusted for physical activity,19,24,33,35,46,47 but these adjustments also did not play a major role in the association. Seven studies adjusted for race,26,30,40,42-44,47 without major changes in their conclusions. Of the 23 studies that did not adjust for race, 20 were in areas with racially and ethnically homogenous populations. A total of 8 studies adjusted for smoking,19,21,33,35,43,46,47,50 but smoking also did not influence the association between serum 25(OH)D and all-cause mortality. One study adjusted for 17 covariates, yet the association persisted.47

Fifteen studies used volunteers from the general population enrolled in cohort studies,20,23,27,28,32,35-38,42,44-46,48,50 or registered as patients of a medical practice.22,31 Two others used electronic medical record systems of health maintenance or insurance organizations,25,30 and 3 used data from the Third National Health and Nutrition Examination Survey population26,40,43 or other National Health and Nutrition Examination Survey data.29,34 Two studies21,26 analyzed coronary angiography patients, but any patient who was acutely ill was excluded. Several other studies used clinic patients of various types.19,41,49 Zittermann et al. used a mixture of ill and healthy individuals.21 One study used patients from a randomized controlled trial of lifestyle intervention on cardiovascular disease36 and another47 used data from the Women’s Health Initiative. The association that was found between high serum 25(OH)D and lower mortality rates persisted despite the diversity of types of studies and geographic sites.

There was substantial heterogeneity among the studies. One possible reason is that the hazard ratio was higher for studies with shorter length of follow-up, specifically less than 7 years compared with those with greater than or equal to 7 years. This may have occurred because less time on the average had elapsed between measurement of 25(OH)D and death in the studies whose follow-up was shorter. If a participant’s 25(OH)D changed since study measurement, that would cause misclassification of the exposure. If the change were more or less random, it would tend to make the observed hazard ratio come out closer to 1.0 than is true. If it were not random, it could influence the hazard ratio in either direction. In this analysis, it appeared that longer follow-up was associated with hazard ratios closer to 1.0. This source of heterogeneity could be addressed in future studies by taking multiple measurements of 25(OH)D over time.

Differences in the amount of adjustment may account for some of the heterogeneity that was present among studies. All studies adjusted for age, and most adjusted for other covariates. Even with extensive adjustment, residual confounding is always possible. Confounding is also possible by factors that were not measured, such as genetic predisposition. It is inherently difficult to completely exclude any possibility of confounding in observational studies, so caution in interpretation of the findings is appropriate.

Supplements of 2000 to 4000 international units (IU) per day of vitamin D3 would produce an approximately 20 to 40 ng/mL increase in serum 25(OH)D.65,66 Two randomized controlled clinical trials using 2000 IU per day of vitamin D3 are under way in the United States, a study designed to replicate the study by Lappe et al.8 and a new randomized controlled trial by Manson et al.57 These studies will not have results available for several years.

Limitations

The validity of a meta-analysis depends upon the validity of the studies that were analyzed. Meta-analyses may include studies with different designs and varied populations. An argument could be made against combining studies with heterogeneous results. However, the benefits of a meta-analysis that is inclusive probably outweigh concerns about heterogeneity. Some meta-analyses may exclude studies that increase the heterogeneity of findings, and make estimates of combined results from an analysis limited to studies with similar findings. In the present study, this would have required limiting the studies with similar findings, and excluding those that were dissimilar. A case can be made for performing a meta-analysis limited to studies that are homogeneous, and there is reasonable logic to including only studies with similar results. Although excluding studies to ensure homogeneity is a reasonable concept, it has a countervailing disadvantage of losing some of the information available from existing eligible studies when the combined estimate is calculated. This analysis included all studies that met the inclusion criteria. This choice makes the findings as comprehensive as possible.
FIGURE 4—Dose–response relationships in each study between serum 25-hydroxyvitamin D and all-cause mortality, in 32 studies of serum 25-hydroxyvitamin D in association with all-cause mortality: 1966–2013


because all relevant studies are included. Including all studies is arguably a conservative approach, because if studies that caused heterogeneity had been dropped, the results would more strongly support the existence of an association of higher serum 25(OH)D with lower all-cause mortality rates. We obtained the pooled dose–response gradient for the present study (Figure 3) by calculating hazard ratios for each of the 32 studies for each of 8 strata of 25(OH)D and calculating the mean hazard ratio in all studies for each stratum. We also calculated weighted means of hazard ratios, using the inverse of the variance of the hazard ratio of each study as its weight. The weighted approach resulted in the findings of one study largely determining the shape of the dose–response gradient, with only minor contributions from other studies. A few studies did not report hazard ratios for the highest few strata, so they did not contribute to the mean hazard ratios for those strata. The curve fit that most closely approximated the observed findings was a decreasing exponential.

A weakness inherent in observational studies is that they may be subject to confounding. For example, people who were ill may not have received as much solar exposure outdoors as those who were healthy. As a result, serum 25(OH)D could be lower to begin with in ill people, who were likely to be at higher risk of death than healthy people. This could have created a spurious association between low serum 25(OH)D and high risk of death. This is known as reverse causation. Therefore, any study detecting an inverse association of serum 25(OH)D with mortality rates must be regarded as tentative because of the possibility that healthier people may have spent more time outdoors than those who had illnesses that may have predisposed them to premature death. We observed an inverse association even in studies with longer follow-up periods (≥ 7 years), which partly mitigates concern about reverse causation. We did not include studies consisting of solely urgently, critically ill patients. This is because a critical illness may have influenced 25(OH)D concentration. Excluding studies of such patients may have reduced the risk of reverse causation, although not necessarily definitively so, and it would have reduced the generalizability of the findings with regard to critically ill patients.

Studies that did not report mortality by categories of serum 25(OH)D were also not eligible for this study, because such data are needed to calculate dose–response or to compare risk of death. All studies were published in biomedical journals and met the usual standard for design of epidemiological studies. Almost all (30 of 32) were cohort studies, which is the strongest observational study design; the other 2 were nested case–control studies, which are of similar capability. The endpoint of mortality hazard ratios does not require a complicated procedure for diagnosis, and it unlikely that there were substantial differences among the studies in accounting for mortality. Determination of vital status is a strong point of virtually all cohort studies. The only substantial differences in quality among studies would be attributable to laboratory procedures used for determining the concentration of serum 25(OH)D. Regarding laboratory methods, there are some data showing that high-performance liquid chromatography (either with ultraviolet detection or coupled with liquid chromatography–mass spectrometry) is usually considered as a gold standard. The problem of determining circulating 25(OH)D is not confined just to the method itself as the precision of different commercial kits could be very different. The validity of this analysis is dependent on the validity of the laboratory procedures that were used in each study. All studies used standard, well-recognized testing procedures. A brief assessment of the quality of the design and analysis of each study is provided in the last column of Table A.

The reader should be cautioned that age-adjusted hazard ratios do not automatically guarantee fair comparisons of mortality risks between study populations that are as different as those analyzed with regard to the distribution of risk.
Factors for death other than the serum 25(OH)D level. Also, the objective of this study was to determine the overall long-term appropriateness and safety of various concentrations of serum 25(OH)D, rather than its associations with specific causes of death, such as cardiovascular disease. It is not possible to draw conclusions regarding specific causes of death from the data provided here.

**Strengths**

A meta-analysis has the advantage of larger numbers and, therefore, higher precision than any individual study. It also provides an analysis procedure for assessing the possibility of publication bias, the funnel plot. In the present study, the plot was funnel-shaped and symmetrical around the overall relative risk of 1.9 for lowest compared with highest category of serum 25(OH)D. In general, studies with the highest precision tended to be closest in magnitude to the overall point estimate of the relative risk, with some exceptions. The overall shape of the funnel plot argues against the possibility of substantial publication bias.

Another strength of this study is that 30 of the 32 studies were favorable with respect to the inverse association of 25(OH)D with all-cause mortality. These 30 studies differed only in degree of association, not direction, increasing confidence that there is an inverse association between 25(OH)D and all-cause mortality rates. Randomized controlled trials would provide a higher level of assurance that this association was not attributable to confounding or reverse causation. It would not take a large number of randomized controlled trials to provide this assurance, but it would take 5 to 10 years to conduct them.

The findings of this study may be generalized to populations similar to persons who volunteer to participate in medical research, enroll in clinical programs, or are Medicare recipients. Generalization to other demographic groups may not be appropriate. For example, almost all studies in this analysis were based on predominately White populations, so the findings cannot necessarily be generalized to persons of other races. Studies of the type we encountered overrepresent older individuals, so the results may not apply to younger people. Similar studies should be performed of younger and non-White populations.

**Conclusions**

There was a downward slope in hazard ratios of mortality according to the serum 25(OH)D concentration. This confirms observations from the National Academy of Sciences-Institute of Medicine Committee to Review Dietary Reference Intakes for Vitamin D and Calcium that concentrations less than 20 ng/mL are too low for safety. The present analysis also suggests that serum 25(OH)D concentrations of less than 30 ng/mL may be too low for safety.

A target range of 25(OH)D of greater than 30 ng/mL could be achieved in most individuals by intake of approximately 1000 IU per day of vitamin D₃, which is one quarter the National Academy of Sciences-Institute of Medicine tolerable upper level of intake of 4000 IU per day at ages 9 years and older. Although it is above the National Academy of Sciences-Institute of Medicine-recommended daily allowance of 600 to 800 IU per day, intake of 1000 IU per day has been reported as safe for daily use for almost all adults, according to the recent Endocrine Society clinical guidelines.

Still, some authors have expressed concern about the efficacy and absolute safety of doses greater than 1000 IU per day, so caution is reasonable.

The Endocrine Society has established a tolerable upper-limit intake of 10,000 IU per day at ages 19 years and older. Doses of vitamin D₃ below 10,000 IU per day in adults have not been associated with toxicity, and serum 25(OH)D concentrations less than 200 ng/mL are generally not considered toxic. This leaves a considerable margin of safety for efforts to raise the population concentration of 25(OH)D to 40 ng/mL.

Well-monitored randomized controlled clinical trials using larger doses of vitamin D₃ would be desirable. While such studies are being considered, it would be reasonable to try to improve the vitamin D₃ status of the population, because of its known associations with lower risk of several diseases. The tolerability of vitamin D₃ intake of 1000 IU per day at ages 9 years and older has been confirmed by a committee report monograph of the National Academy of Sciences-Institute of Medicine.

**About the Authors**

Cedric F. Garland and Edward D. Gorham are with the Division of Epidemiology, Department of Family and Preventive Medicine, University of California San Diego, La Jolla, CA. June J. Kim, Sharif B. Mohr, and Kenneth Zeng are with Leidos Biomedical Research, Inc., San Diego. William B. Grant is with Sunlight and Nutrition Research Center, San Francisco, CA. Edward L. Giovannucci is with departments of Epidemiology and Nutrition, Harvard School of Public Health, Boston, MA. Les Baggerly is with Research Department, GrassrootsHealth, San Diego. Heather Hofflich and Joe W. Ramsdell are with Internal Medicine Group, Department of Medicine, University of California San Diego. Robert P. Heaney is with Department of Medicine, Creighton University, Omaha, NE.

Correspondence should be sent to Dr. Cedric F. Garland, Professor, Department of Family and Preventive Medicine, Division of Epidemiology 0620, University of California San Diego, La Jolla, CA 92039-0620 (e-mail: cgarland@ucsd.edu). Reprints can be ordered at http://www.aph.org by clicking the “Reprints” link. This article was accepted on April 13, 2014.

**Contributors**

C. F. Garland, J. J. Kim, and S. B. Mohr participated in conceptualizing and designing the analysis and drafted sections of the article. J. J. Kim and S. B. Mohr also reviewed the literature and performed the analyses. E. D. Gorham developed plans for performing the study and its design, suggested wording, and made valuable comments and additions to the text. W. B. Grant expanded the literature review, offered many suggestions regarding the interpretation of the results, and reviewed and commented on multiple drafts of the article. E. L. Giovannucci read and provided substantial scholarly changes and edits to 2 major drafts of the article, and provided critiques and wording that were incorporated into the final draft. R. P. Heaney also performed an independent analysis of the dose–response relationship that was incorporated into the article. C. F. Garland had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analyses.

**Acknowledgments**

The authors would like to express their deep gratitude and appreciation to CAPT Gregory Utz, MD, US Navy Medical Corps, for his many contributions and steadfast support of this research program and his outstanding leadership as the commanding officer of the Naval Health Research Center from 2011 to 2013.

Note. The positions expressed in this article are solely those of the authors, and do not represent official positions of the Bureau of Medicine and Surgery, Department of the Navy, Department of Defense, or the US Government. The authors have no conflicts of interest except for W. B. Grant, who receives funding from the Vitamin D Council (San Luis Obispo, CA), UV Foundation (McLean, VA), Bio-Tech Pharmacal (Fayetteville, AR), and the Vitamin D Society (Canada).
Human Participant Protection
This article was a systematic review and analysis of published studies of aggregated data. No personally identifiable information was used. This type of research is exempt from human participants review (45 CFR 46.101(b)).

References


