Effect of Daily Vitamin D_3 Supplementation on Human Health and Performance: Development of an Online Systematic Review and Meta-Analysis Tool for Health Education

Dissertation Proposal

Tom Weishaar
Spring, 2017

Abstract

Background: A best practice for health research is the systematic review and meta-analysis. Health educators would benefit from an open platform for continuous online updating and replication of systematic reviews and meta-analyses. Like Wikipedia, such a tool would crowdsource the discovery and results-entry of health-related randomized controlled trials. It would provide health educators with a database of randomized controlled trials (that is, with an already-completed systematic review) and tools for instant meta-analysis. A similar online tool, called AidGrade, is available now for those who work in international development (www.aidgrade.org). Such a tool could be helpful to health educators in several ways. For example, there are many discordant systematic reviews and meta-analyses of vitamin D supplementation. The resulting controversies reduce the confidence that physicians, nutritionists, and public health professionals have when they educate their patients about vitamin D. In general, these systematic reviews and meta-analyses accept trials of any form of vitamin D (D_2, D_3, their metabolites and analogues, with and without calcium supplementation); trials of both daily supplementation and supplementation with less-frequent but larger bolus doses; and trials with participants having any baseline 25(OH)D status. Standard practice also restricts each meta-analysis to a single disease or condition, even though vitamin D status correlates with a wide range of beneficial effects on human health and performance. This restriction limits the number trials available for analysis. The literature's discordant results may be largely due to these issues. Thus the proposed tool could be used to perform a systematic review and meta-analysis that includes only randomized controlled trials of human subjects with comparison groups that differ only by the amount of daily vitamin D_3 supplementation and with outcomes of health or performance.

Methods: I propose to develop a working model of this tool at the web address *open-meta.org*. The model's technological framework will be based on the statistical programming language R and a variety of R packages that facilitate meta-analysis and provide a web-based interface for R. I will populate the tool with a sample systematic review and meta-analysis on vitamin D supplementation. The rest of the methodology presented here is based on the requirements of the PRISMA-P 2015 guidelines for systematic reviews and meta-analyses (Shamseer et al., 2015) and is structured following the format used by PROSPERO, the International Prospective Register of Systematic Reviews (Booth et al., 2012). After final approval and before beginning the study, I will obtain ethics approval from the institutional review board of Teachers College, Columbia University and will register this protocol with PROSPERO.

Support: The proposed study is unfunded.

Introduction

Background

Vitamin D status has been associated with childhood rickets and adult bone disease since vitamin D's discovery about 100 years ago. Since then, however, researchers have realized that bone disease is only the beginning of vitamin D's story. Inside the nucleus of cells, genetic processes use the blueprints embedded in DNA to make the molecules of life. These processes are controlled by *nuclear receptors*. Signaling molecules interact with these receptors, which then up- or down-regulate the genome's production of specific proteins. Different organisms have different numbers of these receptors, which respond to different substances. Human cells have 48 nuclear receptors, one of which responds specifically to at least three vitamin D metabolites, each regulating, for the most part, different genes (Tuohimaa et al., 2013). Vitamin D's signaling function, which depends upon adequate vitamin D, appears to control at least two percent of the human genome and has been associated with a wide range of human diseases and conditions (Holick, 2007).

Although there are many factors that impact vitamin D status, skin color and body weight account for much of the variation in specific populations (Weishaar, Rajan, & Keller, 2016). Darker skin provides protection from the intense sunlight found near the equator, but at other latitudes, dark skin requires more sun exposure to create as much vitamin D as lighter skin (Whitney & Rolfes, 2008, p. 380). Moreover, vitamin D is measured as a concentration (weight/volume); consequently individuals who weigh more but receive the same amount of vitamin D as smaller individuals will have lower vitamin D concentrations. The health disparities we see in the U.S. between those with dark and light skin and between heavier and lighter individuals may be related to these discrepancies in vitamin D status (Weishaar & Vergili, 2013).

These effects are profound enough to be seen in human evolution. Following an idea first proposed in 1934 (Murray, 1934) and resurrected in the 1960s (Blois, Blum, & Loomis, 1968; Loomis, 1967), Nina Jablonski, an anthropologist, suggests that as human populations moved away from the equator, individuals in those populations with lighter skin had vitamin D levels closer to those of the ancestral population and consequently better health. These disparities, over many generations, led to the evolution of the spectrum of human skin colors we see in human populations today (Chaplin & Jablonski, 2009; Jablonski, 2004; Jablonski & Chaplin, 2000; Yuen & Jablonski, 2010).

Jablonski's theory is supported by U.S. population-based health surveys, which find that vitamin D concentrations average 17.4 ng/mL in non-Hispanic blacks, 21.9 ng/mL (26% higher) in Mexican-Americans, and 28.3 ng/mL (62% higher) in non-Hispanic whites (Weishaar et al., 2016). In terms of the benefits of dark skin, the most dangerous type of skin cancer, known as invasive melanoma, is 24.7 times more likely in U.S. whites than in African-Americans (U. S. Department of Health and Human Services, 2014). Rural residents of Ghana living near the equator (6° N) have an average vitamin D concentration of 30 ng/mL, which is higher than any group in the U.S. (Durazo-Arvizu et al., 2014). Traditionally-living people in east Africa (4° S) have an average concentration of 46 ng/mL (Luxwolda, Kuipers, Kema, Dijck-Brouwer, & Muskiet, 2012). This is our best available estimate of vitamin D concentrations over the course of human evolution.

Another anthropologist, Kathleen Fuller, was the first to turn the theory of the evolution of human skin color around and look at it from a health perspective (Fuller, 2003). From that perspective, the theory implies that health disparities will be found in all populations with diverse skin colors. Where sunlight is intense, those with lighter skin will be at a disadvantage;

where sunlight is less intense, those with darker skin will have poorer health. However, these disparities can easily be treated with sun protection and vitamin D supplementation.

There are dozens of other articles in the literature supporting these relationships, which are slowly gaining some acceptance among medical and nutritional professionals, but have yet to be recognized or taken seriously either by those who study health disparities or by those who set public health policy in the United States. The biggest barrier to complete acceptance of these theories in health policy appears to be discordance among the conclusions of systematic reviews and meta-analyses of randomized controlled trials of vitamin D supplementation.

A *systematic review* is a special type of literature review. Systematic reviews, which can be done with or without meta-analysis, are designed to control researcher bias (Chalmers, Hedges, & Cooper, 2002). Instead of allowing a reviewer to cherry-pick literature that the reviewer agrees with, a systematic review requires the reviewer to explicitly state the strategy that will be used to discover studies (with the goal of including all relevant studies in the analysis). The reviewer must also explicitly state the criteria that will be used for including and excluding discovered studies from the review. Ideally, before the actual review process begins, the reviewer must detail this strategy and these criteria, along with additional information about the proposed review process, in an online registry. If any changes are made to the proposed process, they must be explained as part of the review itself. Again, the purpose of this process is to control reviewer bias.

A *meta-analysis*, which can be done with or without a systematic review, brings statistical precision to the process of examining a group of studies and determining what they mean (Chalmers et al., 2002). Before meta-analysis, reviewers often evaluated the literature by counting the number of studies with significant and non-significant results and declared the truth

resided in the group with the larger count. But statistical significance is highly dependent on the number of subjects in a study; any study with enough subjects will be statistically significant. The statisticians who developed meta-analysis realized that what was important was not so much the statistical significance of the study's effect but the size of the effect (Glass, 2015). One advantage of using effect sizes is that the results of different studies can be combined with statistical rigor to produce an overall effect size. Moreover, the overall effect size can be tested for significance and other statistical characteristics.

Currently the best practice for determining whether an intervention has an effect is a combined systematic review and meta-analysis of randomized controlled trials. There have been hundreds of randomized controlled trials looking at various health and performance effects of vitamin D supplementation. There is also already an abundance of systematic reviews and meta-analyses combining groups of these trials. But their results do not agree.

Conceptual Framework

I propose that the discordant results of these reviews are due to three unappreciated sources of heterogeneity – the form of the vitamin D given as a supplement; daily versus less-frequent, larger bolus dosing; and differences from study to study in the baseline vitamin D status of the participants.

Existing reviews typically assume that vitamin D_2 and vitamin D_3 are equally effective, but this view has been challenged**.** Vitamin D_2, also known as *ergocalciferol*, is produced by fungi, including yeast. Vitamin D_3, also known as *cholecalciferol*, is produced by plants and animals (Japelt & Jakobsen, 2013). Vitamin D has a long evolutionary history going back to the development of cells with a nucleus.

Long before vitamin D was discovered, it was known that certain foods – particularly cod liver oil – would cure rickets, a disease known since antiquity. Vitamin D was discovered in 1922 after a process for destroying the vitamin A in cod liver oil was developed and the resulting oil continued to cure rickets in rats. About the same time, another traditional cure for rickets, fresh air and sunshine, was confirmed by curing rickets in children using exposure to ultra-violet light (UV) from quartz-mercury lamps. Soon researchers discovered that UV radiation would also give antirachitic properties to many foods. Then they discovered that a fungal steroid derived from ergot and called ergosterol was the substance that picked up antirachitic properties when exposed to UV light. By 1931, several research groups had purified and crystallized the resulting product, which they named ergocalciferol (Wolf, 2004).

But plants and animals, unlike fungi, don't produce ergosterol. Even today many experts think that plants, like fungi, produce vitamin D_2; this misconception is a result of fungal contamination of plants and high concentrations of ergosterols in fungi (Japelt & Jakobsen, 2013). It took until 1937 to discover that in animals the precursor that gains antirachitic properties when exposed to UV light is a form of cholesterol. The resulting substance was named cholecalciferol or vitamin D_3.

Another 50 years passed before reports appeared in the literature suggesting that vitamin D_2 was not as effective in humans as vitamin D_3 (Tjellesen, Hummer, Christiansen, & Rodbro, 1986). By 2006 Houghton and Vieth argued in the *American Journal of Clinical Nutrition* that "vitamin D_2 should no longer be regarded as a nutrient appropriate for supplementation or fortification of foods" (Houghton & Vieth, 2006). However, now, over a decade later, most medical, nutritional, and public health professionals continue to assume that vitamin D_2 and D_3 are equivalent. For example, in the U.S., vitamin D supplements available by prescription (rather

than over-the-counter) continue to be compounded with vitamin D_2 rather than D_3. Meta-analyses of vitamin D trials typically make no distinction between the two.

The existing reviews also tend to assume that vitamin D supplements will be equally effective in either small daily doses or in larger, less-frequent bolus doses. The human liver rapidly converts the parent form of vitamin D to 25-hydroxyvitamin D (25(OH)D). Consequently, assays to determine vitamin D status are based on serum concentrations of 25(OH)D, which will rise no matter how vitamin D supplement doses are split, leading to the assumption that daily and bolus dosing are equally effective. However, after the liver converts the bolus dose to 25(OH)D, there is little remaining parent vitamin D in the serum. The liver-made 25(OH)D primarily ends up attached to vitamin D binding protein, an albumin-like blood protein. This 25(OH)D, which can be removed from vitamin D binding protein and further metabolized by the kidney, has a major role in calcium balance. However, the primary form of vitamin D absorbed by cells throughout the body for DNA signaling is likely the parent form (Hollis & Wagner, 2013). Cells outside the calcium-balance system may not depend upon the 25(OH)D bound to vitamin D binding protein at all. If this theory is correct, daily dosing, which maximizes the daily levels of parent vitamin D, should have a larger impact on health than bolus dosing.

The third source of heterogeneity is differences in the baseline 25(OH)D status of trial participants. Vitamin D supplementation studies are different from drug studies in that drug studies can assume that participants have not received the drug from any source other than the intervention. This is simply not the case with vitamin D. In addition to the intervention dose, both the control and intervention groups are exposed to vitamin D from sunlight and dietary sources. There can be additional differences in 25(OH)D status related to body weight and other

personal, cultural, and geographic characteristics of the study's population. Randomization, in expectation, accounts for these differences within a single study, but *differences between studies* in the baseline vitamin D status of the subjects is typically an unanalyzed source of heterogeneity in vitamin D meta-analyses. A study with subjects having very low baseline vitamin D status may show a larger effect than a study using the same dose, but with subjects having a very high baseline vitamin D status. As a measure of total exposure to vitamin D, 25(OH)D status should be a better predictor of effect than dose.

Moreover, if serum availability of the parent form of vitamin D is the actual determinant of any beneficial effects, serum 25(OH)D status should be understood as a biomarker for vitamin D *exposure* rather than a biomarker for *effect*. Nonetheless, as a biomarker of *total* exposure, 25(OH)D status should be a better indicator of effect than *dose*, which accounts for only a part of total exposure. Study-to-study differences in vitamin D exposure, as measured by the mean 25(OH)D status of the control group at outcome, is likely to be a third source of important heterogeneity in meta-analyses of vitamin D supplementation.

In addition to issues with heterogeneity, many of the vitamin D systematic reviews cited in the next section show a trend toward effectiveness, but that trend is not statistically significant because of the limited number of trials for any particular outcome. However, the health effects of vitamin D may be homogenous enough to combine trials with different health and performance outcomes in a single meta-analysis. This would address the statistical problem posed by the limited number of trials for any single outcome. Generalizing the outcome has a long history in meta-analysis; the first meta-analysis ever done mixed trials with various outcomes of psychotherapy (Smith & Glass, 1977). I propose to improve the statistical qualities of the meta-

analysis by generalizing this mixing at scale and including any study with an outcome related to human health or performance.

Specific aims. In the context of these specific aims, a *beneficial effect* is an effect size favorable to daily vitamin D_3 supplementation with a 95% confidence interval that does not include the no-effect value. This study has three specific aims:

1) To develop an online model for health educators of an open platform for continuous online updating and replication of systematic reviews and meta-analyses.
2) To seed the online platform with studies from a systematic review of daily vitamin D_3 supplementation.
3) To use the online site to answer three research questions:
    a. Does daily vitamin D_3 supplementation have a beneficial effect on human health and performance outcomes overall?
    b. Does daily vitamin D_3 supplementation have a beneficial effect on human health and performance outcomes for which there are known racial health disparities?
    c. Does lower control-group 25(OH)D status at outcome measurement have a larger beneficial effect than high status?

Study innovations and significance. Systematic reviews and meta-analyses of vitamin D supplementation typically generalize the allowable intervention (e.g., by accepting trials with different supplementation interventions – vitamin D_2, D_3, their metabolites, and their analogues with either daily or bolus dosing) while specifying a single health outcome. The unique feature of this study is that it will specify the vitamin D supplementation intervention exactly (daily D_3 only) while generalizing the outcome to any effect on human physical or mental health or performance.

Literature Review

This literature review includes some of the features of a systematic review. A systematic review of systematic reviews is sometimes called an *umbrella review*. What is systematic about this review is that I searched for all of the relevant literature using specific criteria. Those criteria are that this review centers on systematic reviews and meta-analyses of randomized controlled trials in which the intervention was vitamin D_2 or D_3 supplementation. I accepted systematic reviews of trials written in English with any human participants, any comparison group, and any outcome related to human health or performance.

Search strategy. On December 6, 2016 I completed four literature searches for systematic reviews and meta-analyses of vitamin D supplementation trials. Overall results of these four searches are shown in Table 1. A search on PubMed using the terms *(((((("vitamin d") OR "vitamin d2") OR "vitamin d3") OR ergocalciferol) OR cholecalciferol) AND ("losubjtsystematic reviews" [Filter] AND "english and humans"[Filter])* returned 53 hits. Two searches on Cochrane Central using the terms *Vitamin D OR ergocalciferol OR cholecalciferol* returned 58 Cochrane reviews and 152 reviews done by others. Finally a page on the Vitamin D wiki (VitaminDWiki, 2016) lists a menu of 230 vitamin-D-related meta-analyses. I wrote a program that searched those 230 wiki pages for a PubMed id (PMID) or a document object identifier (DOI) and developed an additional list of 131 reviews (the remaining wiki pages didn't have a PMID or DOI).

Of the 394 total reviews, 66 (17%) were duplicates. I did a stage 1 review on the remaining 328 reviews and discarded 75% of them, as shown in Table 1. The primary reasons were that some or all of the trials included in the review were observational (124 reviews, 38%) or that the intervention wasn't vitamin D_2 or D_3 (75 reviews, 23% – some trials in these reviews

used vitamin D analogues, which are versions of vitamin D that have been chemically altered so that they can be patented; some included or were primarily about calcium supplementation; yet others studied metabolites of vitamin D as the intervention). Some of the reviews (26, 8%) examined genetic techniques rather than supplementation interventions. Finally, a very small subset of reviews (5, 2%) used trials with the outcome measure of 25(OH)D concentration, which isn't a measure of health or performance.

As shown in the lower part of Table 1, I also did a duplication analysis by database. Surprisingly, the Cochrane reviews appeared only in Cochrane CENTRAL. Although these reviews are considered the gold standard for systematic reviews and are listed in PubMed, the PubMed search didn't find them. If they are listed at all on the Vitamin D Wiki, they don't include a PMID or DOI. Moreover, only 30% of the PubMed hits were unique to PubMed. The Vitamin D Wiki provided the most unduplicated reviews, 103, and had the second highest percentage of unique reviews (79%) after CENTRAL's 100% performance on its own reviews. Returning to the Stage 1 Pass line of the table, the CENTRAL search on reviews from non-Cochrane sources had the highest hit rate for reviews I was actually interested in (36%), followed by PubMed (26%), the Vitamin D Wiki (23%) and CENTRAL's Cochrane reviews (17%). Of the original 328 unduplicated reviews, 83 (25%) passed the Stage 1 review.

Of these 83, an additional 16 (19% of the group of 83; 5% of the original 328) failed Stage 2 review, leaving 67 reviews. Of those that failed, eight weren't systematic reviews or meta-analyses after all, four were umbrella reviews, two studied the wrong treatments (one 25(OH)D, one calcium), one was all genetic trials, and one used 25(OH)D status as the outcome. Of the 67 remaining reviews, the oldest was published in 1998 and half (34) were published in the two-year period of 2013-2014. Before that there were 25 (37%) reviews; since then 8 (12%).

Table 1. Search results and stage 1 review												
	PubMed		Cochrane Reviews		Other Cochrane		Vitamin D Wiki		without duplicates		duplicates	Total
	n	%	n	%	n	%	n	%	n	%	n	n
Total	53		58		152		131		328		66	394
Stage 1 Pass	14	26%	10	17%	55	36%	30	23%	83	25%	26	109
Stage 1 Fail												
Not a meta-review	1	2%	0	0%	3	2%	7	5%	11	3%	0	11
All observational	2	4%	0	0%	17	11%	58	44%	68	21%	9	77
Some observational	15	28%	14	24%	30	20%	14	11%	56	17%	17	73
All genetic	0	0%	0	0%	16	11%	10	8%	26	8%	0	26
All analogues	5	9%	4	7%	6	4%	3	2%	13	4%	5	18
Treatment not D	16	30%	29	50%	18	12%	7	5%	62	19%	8	70
Only 25OHD status	0	0%	0	0%	5	3%	1	1%	5	2%	1	6
Other	0	0%	1	2%	2	1%	1	1%	4	1%	0	4
Duplicate Analysis by database												
In 1 db	16	30%	58	100%	90	59%	103	79%				
In 2 dbs	32	60%	0	0%	62	41%	28	21%				
In 3 dbs	5	9%	0	0%	0	0%	0	0%				

Characteristics of selected reviews. The characteristics of randomized controlled trials are typically described using *PICO*: participants, intervention, comparison group, and outcome.

In terms of the participants in the randomized controlled trials, 32 of the reviews (48%) used all the trials otherwise available without regard to their participants. Filters used by the other reviews were primarily based on age (e.g. neonates, less than 5, 18-40, postmenopausal, greater than or equal to 50 or 60 or 65 or 75), but also included trials of subjects with or without specific diagnoses (e.g. chronic heart failure, tuberculosis, cystic fibrosis, without bone disease), pregnancy status (both pregnant only and not pregnant only), specific risks (e.g. "at risk for depression"), specific co-treatments (e.g. "patients taking corticosteroids"), and specific living conditions (e.g. community-dwelling, living in long-term care). None of the reviews filtered on 25(OH)D status, although some trials do so at baseline.

In terms of the intervention in the randomized controlled trials, 57 (85%) of the reviews did not distinguish between trials of vitamin D_2 and vitamin D_3. Of the remaining reviews, 5 (7%) accepted only trials of vitamin D_3 supplementation; however, of these two allowed trials that also included calcium supplementation and three didn't. Two accepted trials of either D_2 or D_3, all without calcium. Three more accepted trials of D_2, D_3, and either vitamin D metabolites (2) or ultraviolet light (1).

In terms of the comparison group, 61 of the reviews (91%) did not report a comparison group in the abstract. Of the remaining six, all reported accepting trials that compared supplementation to a placebo group. Of those, two also reported allowing trials that compared supplementation to "no intervention", one to "calcium", one to "other supplement", and one to "active comparators". (At the trial level, the "comparison group" has to do with the types of control groups allowed in the trials included in the systematic review. There are also review-

level comparisons: one review made clear that it would compare daily versus bolus dosing and another, which did not mention accepted comparisons in the trials themselves, reported comparing trials with "biological flaws" to trials without such flaws. PICO refers to trial-level, not review-level, comparisons.)

Note that so far PICO hasn't done much to distinguish between these reviews. Most reviews of vitamin D supplementation accept trials with any subjects, any form of vitamin D, and any comparison group. But these reviews get specific with trial outcomes. Bone fracture was the most popular outcome for vitamin D supplementation reviews (11, 16%). Some reviews looked at all fractures, some at non-vertebral fractures, and some at hip fractures. Next most popular was a tie between falls and bone mineral density (7 each, 10%). Other reviews allowed trials with related outcomes, such as six outcomes related to heart disease or to type II diabetes. Many others looked at a single outcome: mortality, depression, blood pressure, asthma exacerbation, dental caries, childhood pneumonia, breast cancer prevention, muscle strength, pain score, and so on.

Two massive reviews looked at numerous outcomes. Both were published in 2014. *Vitamin D and Calcium: A Systematic Review of Health Outcomes (Update)* is 929-page review by the U.S. Agency for Healthcare Research and Quality (Newberry et al., 2014). *Screening for Vitamin D Deficiency: Systematic Review for the U.S. Preventive Services Task Force Recommendation* is a 223-page review (LeBlanc, Chou, Zakher, Daeges, & Pappas, 2014). Although both reviews appear to combine multiple outcomes, in fact they are aggregations of multiple meta-analyses, each of which looks at a single outcome.

Two of the 67 reviews found only two acceptable trials. On average, the reviews include about 14 trials. Twelve reviews (18%) include 20 or more trials; the most frequent outcome in

this group is mortality, data on which can be obtained from trials actually designed to investigate any other outcome as long as mortality is tracked and reported. The maximum number of trials included in one meta-analysis was 56 (Bjelakovic et al., 2014).

Findings of selected reviews. The overall conclusions of 39 (56%) of the articles were positive toward vitamin D supplementation, 8 (11%) were neutral, and 23 (33%) were negative. In general, almost all of the neutral and negative conclusions were due to a lack of sufficient data; the differential tone of these conclusions may have reflected researcher bias. Likewise, some of the positive conclusions were based on very weak results, which may reflect bias in the other direction. One researcher in particular insisted on an effect size of 15% or more, which led to consistent negative conclusions.

In general, the pooled effect sizes usually favored supplementation, although the majority weren't statistically significant because of a lack of trials. For example, 25 of the reviews pooled effect sizes as risk ratios and reported results for 53 outcomes with more than one trial. The minimum RR was .25 (for stillbirths, 3 trials); the maximum was 1.22 (for risk of vertebral fracture, 9 trials); 40 outcomes (77%) were less than 1, 12 (23%) were 1 or greater; the average was .82. Three of the reviews pooled effect sizes as hazard ratios and reported results for nine outcomes. The lowest HR was .70 (for risk of hip fracture) and the highest was 1.07 (for stroke); the average was .89. Other reviews pooled effect sizes as odds ratios, standardized mean differences, weighted mean differences, and unstandardized-unweighted mean differences.

Conclusions. Existing systematic reviews and meta-analyses of vitamin D supplementation tend to have exact inclusion criteria only for outcomes. These reviews tend to include trials with any participants, any vitamin D-related intervention, and any comparison group. About half conclude with a positive comment about vitamin D supplementation; the other

half conclude with an inconclusive or negative comment. The number of trials included in each meta-analysis tends to be small, limiting statistical significance, but effect sizes tend to favor vitamin D supplementation. There are no reviews in the literature that include a variety of health or performance outcomes that aren't closely related. Likewise, none of the existing analyses look only at outcomes related to health disparities and only a small handful look only at daily vitamin D_3 supplementation interventions.

Methodology

This proposal combines the development of a model online tool for open, crowd-sourced systematic review and meta-analysis with an initial demonstration of the tool. The initial demonstration of the tool will consist of a systematic review and meta-analysis of the effect of daily vitamin D3 supplement on health and performance.

The methodology for development of the model tool will include establishing a domain for the project at *open-meta.org*; establishing a code repository for the project at *github.com*; and developing the model tool using the statistical programming language R and associated packages that facilitate meta-analysis and provide R with a web-based interface.

The model tool at *open-meta.org* will provide the following features:

- The ability to enter information about database searches, including the results, which will be able to be uploaded into *open-meta.org.*
- The ability to automatically determine PubMed identification numbers of uploaded articles using PubMed's *Entrez* programming utilities and to use these ids to identify duplicated articles.
- The ability of multiple reviewers to do a Stage 1 review with result categories customized for the project. The Stage 1 review screen will show the title and abstract of each article.
- The ability of multiple reviewers to do a Stage 2 review with result fields customized for the project. The Stage 2 review screen will allow data entry in the format provided in the article but will display results in an effect size statistic chosen by the user.
- The ability to select subgroups of trials for analysis

- The ability to display a PRISMA flow diagram, forest plot, or a funnel plot for the selected trials.

Regarding this proposal's demonstration systematic review and meta-analysis; I will obtain ethics approval from the Institutional Review Board of Teachers College, Columbia University and register the following study protocol on PROSPERO. The following methodology is organized following the PRISMA-P 2015 guidelines for creating a protocol for systematic reviews and meta-analyses (Shamseer et al., 2015):

> Protocols of systematic reviews and meta-analyses allow for planning and documentation of review methods, act as a guard against arbitrary decision making during review conduct, enable readers to assess for the presence of selective reporting against completed reviews, and, when made publicly available, reduce duplication of efforts and potentially prompt collaboration (p. 1).

Administrative Information

Title. Effect of daily vitamin D_3 supplementation on human health and performance: protocol for a systematic review and meta-analysis

Registration. This protocol will be registered with the International Prospective Register of Systematic Reviews (PROSPERO) after committee and IRB approval.

Author. Tom Weishaar, Doctoral Candidate, Department of Health and Behavior Studies, Teachers College, Columbia University, 501 W 120th St - #8W, New York, NY 10027 jtw2117@columbia.edu

Support. The study will be submitted in partial fulfillment of the requirements for a doctorate degree in Health Education. Sponsorship and guidance was provided to the author by

Dr. Sonali Rajan, Department of Health and Behavior Studies and by Dr. Beth Tipton, Department of Human Development, Teachers College, Columbia University. The study is not funded.

Introduction

Rationale. The nucleus of each human cell has 48 nuclear receptors. Each of these receptors, after interacting with its related ligands, up- or down-regulates specific genes. One of these nuclear receptors responds specifically to at least three vitamin D metabolites, each of which controls, for the most part, different genes. Vitamin D status has been recognized as a determinant of human health for almost a century and has been associated with a wide range of human diseases and conditions. In the U.S. population, skin color accounts for much of the variation in vitamin D status. Darker skin provides protection from the intense sunlight found near the equator, but at the latitude of the U.S. requires more sun exposure to create as much vitamin D as lighter skin. Research suggests that racial health disparities may be related to these disparities in vitamin D status. These relationships, which are slowly gaining some acceptance among medical and nutritional professionals, have yet to be recognized or taken seriously either by those who study health disparities or by those who set public health policy in the United States. The biggest barrier to complete acceptance of these theories in health policy appears to be discordance among the conclusions of systematic reviews and meta-analyses of randomized controlled trials of vitamin D supplementation.

Objectives. The discordant results of these reviews may be due to three unappreciated sources of heterogeneity – the form of the vitamin D given as a supplement, daily versus bolus dosing, and the baseline 25(OH)D status of a study's participants. Existing systematic reviews typically assume that vitamin D_2 and vitamin D_3 are equally effective, but this view has been

challenged**.** Existing reviews also tend to assume that vitamin D supplements will be equally effective in either small daily doses or in larger, less-frequent bolus doses. This, too, has been challenged. Existing reviews pay close attention to the dose used in the intervention but not to total vitamin D exposure as measured by the 25(OH)D status of the participants in the study.

In addition to problems with heterogeneity, most existing vitamin D systematic reviews show a trend toward effectiveness, but that trend is not statistically significant because of the limited number of trials for any particular outcome. However, the health effects of vitamin D may be homogenous enough to combine trials with different health and performance outcomes in a single meta-analysis. This would address the statistical problem posed by the limited number of trials for any single outcome. Generalizing the outcome has a long history in meta-analysis; the first meta-analysis ever done mixed trials with various outcomes of psychotherapy. This systematic review will include any trial arm with an outcome related to human physical or mental health or performance.

Specific Aims. In this context, a *beneficial effect* is an effect size favorable to daily vitamin D_3 supplementation with a 95% confidence interval that does not include the no-effect value.

- Does daily vitamin D_3 supplementation have a beneficial effect on human health and performance outcomes overall?
- Does daily vitamin D_3 supplementation have a beneficial effect on human health and performance outcomes for which there are known racial health disparities?
- Does lower control-group 25(OH)D status at outcome measurement have a larger beneficial effect than high status?

Systematic reviews and meta-analyses of vitamin D supplementation typically generalize the allowable intervention (e.g., by accepting trials with different supplementation interventions – vitamin D_2, D_3, their metabolites, and their analogues with either daily or bolus dosing) while specifying a single health outcome. The unique feature of this study is that it will specify the vitamin D supplementation intervention exactly (daily D_3 only) while generalizing the outcome to any effect on human physical or mental health or performance.

Methods

Eligibility criteria. These criteria apply to arms of experimental trials. Many trials have multiple arms; these eligibility criteria will be applied to arms, not to trials. In other words, any given trial may have multiple arms; some arms may be eligible and some may not.

Study designs. We will accept only study arms from randomized controlled trials (RCTs) with individual randomization (not cluster randomization).

Participants. Humans. No other eligibility limitations. Where the information is available, we will record group mean age and its standard deviation, group mean weight and its standard deviation, percent female for each group, and health status of the participants.

Intervention. The intervention must be daily vitamin D_3 supplementation by any method of administration (by mouth, by injection, etc.). The method of administration will be recorded.

Comparators. The comparison group can be no-intervention, placebo, or standard-of-care. The type of control group will be recorded. Comparison groups must be identical except for the daily dose of vitamin D_3, so, for example, both groups may also take an equal amount of calcium. No-intervention control groups will be considered identical to placebo control groups receiving no supplementation. Study arms in which the control group receives a smaller dose of vitamin D_3 than the intervention group (typically a standard-of-care control group) will be

accepted. When a trial includes multiple arms with intervention groups taking different dose sizes, the control group for all arms will be the group taking the smallest dose (including none). If a control group receives vitamin D_3, the difference in dose size between the control and intervention group will be considered the intervention dose for that arm.

Outcomes. The study will accept any outcome measure related to human physical or mental health or performance. This study does not consider 25(OH)D status itself to be an eligible outcome.

Timing. This study concerns active supplementation. Arms with an outcome measured less than four weeks after supplementation begins or after supplementation ends are not eligible. Paired control and intervention groups may be measured at multiple times for a single outcome. These will be recorded as separate arms with different durations of supplementation. Duration of supplementation at outcome measurement will be recorded.

Setting. No restrictions. The setting describes living conditions of the participants, such as community-dwelling, assisted living, long-term care, or hospice. When available the setting will be recorded.

Geographic location. No restrictions. When available the latitude of the research location will be recorded.

Language. Adequate information about eligible study arms must be available in English.

Information sources. As noted earlier, this study has a single author and will fulfill a dissertation requirement for a doctoral degree. The study is not funded and is limited to the resources of the single author. To maximize the author's productivity, databases searched for this study will be restricted to the Cochrane Central Register of Controlled Trials (CENTRAL). Since 1998, Cochrane review groups have completed 11 systematic reviews of vitamin D

supplementation on pregnancy, infection prevention in children, management of asthma, treatment of chronic pain in adults, mortality (2), cancer prevention, fracture prevention, cystic fibrosis, bone mineral density in children, and corticosteroid-induced osteoporosis. The most recent study on mortality includes more trials than any other vitamin D systematic review ever published.

CENTRAL contains a record for every study examined in its own reviews. In addition, CENTRAL is updated monthly with records of new randomized controlled trials retrieved from Medline and EMBASE, from specialized registers created by Cochrane's review groups, and from Cochrane's hand search results register. Because of the comprehensive nature of CENTRAL's database of randomized controlled trials, CENTRAL by itself can provide a systematic view of all relevant randomized controlled trials in both the primary and the grey literature. This avoids the additional work of stage 1 reviews finding duplicative results in additional databases that are a mix of randomized controlled trials and other types of publications, making this project feasible for a single author. Complete information on the contents of CENTRAL and how it is updated is available at http://www.cochranelibrary.com/help/central-creation-details.html.

Any other vitamin D_3 trials not discovered by this process may be included in the study. The discovery method of each trial will be recorded. CENTRAL will be notified of any trials missing from its database.

Search strategy. The CENTRAL search strategy will use the MeSH descriptor *[Vitamin D] explode all trees*. Adding additional vitamin D-related MeSH terms does not increase the number of records returned by CENTRAL. Adding additional vitamin-D related *text terms* vastly increases the number of records returned, but the author does not have the resources to review

that many papers, particularly since almost all will fail stage 1 review anyhow (these are trials in which vitamin D is mentioned somewhere in the full text; the MeSH descriptor identifies trials that are actually about vitamin D). All trials returned by the MeSH search will be reviewed for eligibility without further limits. CENTRAL is limited to trials in humans by design. Based on preliminary testing of this strategy, CENTRAL will provide records for more than 2,750 journal articles.

Study records – Data management. The author will provide a computer application in the R statistical programming language at *open-meta.org*. The application will be used throughout the study.

The complete record set from the CENTRAL search will be downloaded and then loaded into *open-meta.org*. At a minimum, each CENTRAL record includes the article's title, authors, date of publication, journal name, and journal volume, number, and page. Most CENTRAL records also include the article's PubMed ID (PMID) and abstract. *Open-meta.org* will use the PMID (as well as other available data for records lacking a PMID) to identify duplicates. *Open-meta.org* will track the original source of a record (CENTRAL or other).

Study records – Selection process. The author will complete the Stage 1 review using *open-meta.org*, which will display each trial's record, including title and abstract, and allow the reviewer to record the result of the review, a comment about the review, or a comment about the trial. If a PMID is available for the record, the title and abstract will be obtained from PubMed as the page is displayed to the reviewer, otherwise *open-meta.org* will display the information obtained from CENTRAL. *Open-meta.org* will allow for multiple reviews (and multiple reviewers, although this demonstration study will have only one). If any review receives a stage 1 pass, the trial will be included in the stage 2 review. Reasons for failing a stage 1 review will

be: not available in English; not a randomized controlled trial; no valid participants; no valid intervention – not daily; no valid intervention – not D_3; no valid intervention – other; no valid comparison group; and no valid outcome.

At least six weeks after the initial review is completed, the author will repeat the Stage 1 review process, blinded to the initial review, for 10% of the initial study database. These ratings will be compared to the original ratings, a reliability score will be calculated and reported, and any additional studies given a Stage 1 Pass will receive a Stage 2 Review.

Study records – Data collection process. Full-text articles will be obtained for all trials that pass Stage 1 review. As the data is collected from the full-text article, the author will search for any mention that the trial is part of a larger study. If so, the name of the larger study will be recorded and the author will search PubMed for other articles containing that name to determine if there are others from the same larger study. If so, these sets of articles will receive special handling to make sure all information from the larger study is included while duplicative or illogical information is not.

For each arm reported in each study, the author will enter into *open-meta.org* both mandatory and supplemental information. If any of the mandatory information is missing for all trial arms, or if it becomes apparent during data collection that the trial should not have passed stage 1 review, then the trial will fail stage 2 review and the reason for failure will be recorded.

Data items. Mandatory information for an arm includes the outcome, how the outcome is reported (e.g., group means, risk ratio), and the duration of supplementation at the time of outcome measurement. Mandatory information for both the control group and the intervention group includes the number of participants at baseline, the number of participants at outcome measurement, the daily vitamin D_3 dose, mean 25(OH)D status at outcome measurement, and the

outcome measure in terms of both central tendency and variance. In addition to the mandatory information, supplemental information about each arm will be recorded if available. Supplemental information for the arm includes the type of control group, how the supplement was administered, research setting, latitude, participant diagnosis, pregnancy status, specific risks, and co-treatments. Supplemental information for each group includes mean and standard deviation of age and body weight, as well as percent female – preferably at baseline but acceptable at outcome measurement. Supplemental information for each group will also include the mean and standard deviation of 25(OH)D status at baseline and the equivalent standard deviation at outcome measurement.

Outcomes and prioritization. An arm's outcome in this study can be any physical or mental measure of health or performance. Outcomes measured less than four weeks after supplementation begins or after supplementation ends are not eligible. The study is particularly interested in outcomes for which there are racial health disparities. The only anticipated invalid outcome is 25(OH)D status, which is the primary outcome measure of some vitamin D dosing trials, but using that as a measure of health or performance in the context of this study would be illogical.

Risk of bias in individual studies. In addition to the mandatory and supplemental information about trial *arms*, the author will enter data into *open-meta.org* on the risk of bias in each *trial*. This data will be categorical (high risk of bias, low risk of bias, unclear risk of bias) for the following five trial characteristics: industry funding, randomization and allocation concealment, blinding of participants, blinding of research personnel and outcome assessors, and level of attrition and exclusions. These categories are adapted from Table 8.5.a in the *Cochrane*

Handbook for Systematic Reviews of Interventions. This data will be analyzed and reported in the evaluation of the overall strength of evidence of this systematic review.

Data synthesis – criteria under which study data will be quantitatively synthesized. The author will synthesize the data using meta-analysis. Because one trial may provide multiple outcomes and even multiple arms for a single outcome (trial arms could have different doses or durations or both), the author will use the robust variance estimation approach proposed by Hedges, Tipton, and Johnson (Hedges, Tipton, & Johnson, 2010) to address correlated effects within studies. Unlike random effects meta-analysis methods, robust variance meta-analysis methods allow inclusion of all outcomes from a study, but require more studies for stable results than random effect methods. Random effect methods, on the other hand, require selecting just one outcome or averaging the outcomes over each trial, reducing the number of data points for the analysis from the number of outcomes to the number of trials. With robust variance methods, the weight of a trial is proportional to the trial's variance, but is distributed across the trial's outcomes. In secondary analyses where the number of studies is too limited to use robust variance methods, the author will use random effects methods.

Data synthesis – planned summary measures, method of handling data, and methods of combining data from studies, including exploration of consistency. Trial results will be recorded in the format used in the report of the trial but will be converted to an effect size reported as a risk ratio and its 95% confidence interval by *open-meta.org*. The author will calculate and report an overall effect size for all outcomes and an effect size for all outcomes associated with racial health disparities. Each effect size will be accompanied by an estimate of trial homogeneity/heterogeneity as measured by the I^2 statistic.

If there is heterogeneity in either of the first two main analyses ($I^2 > 50\%$ or $p < .05$), the author will investigate and attempt to identify the source of the heterogeneity using subgroup and sensitivity analysis incorporating type of outcome, dose, duration of supplementation, outcome 25(OH)D status in the control and in the intervention groups, and the data recorded about trials (e.g., industry funding), about arms (e.g., setting, latitude, specific risks, co-treatments), and about groups (e.g., age, body weight, percent female, attrition).

The result for the third specific aim of this study, which relates to 25(OH)D status in the control and intervention groups at outcome, will be determined using meta-regression. In meta-regression the experimental unit is the study. In this regression, study effect sizes will be the dependent variable and the mean 25(OH)D status of the control group at outcome measurement will be the explanatory variable. The results of interest are the relative size, direction, and significance of the regression coefficient for the explanatory variable. If the coefficient for control group 25(OH)D status is negative, that would indicate that lower group baseline status leads to larger effects.

When mandatory data is missing for a trial arm, that arm will be dropped from this analysis. If all of a trial's arms are dropped because of missing mandatory data, the trial will be dropped from this analysis and recorded as a Stage 2 Review failure.

Data synthesis – additional proposed analyses. In addition to the three main analyses, the author will perform separate subgroup analyses for outcomes related to physical health, mental health, physical performance, and mental performance.

Meta-bias(es). Empirical evidence suggests that a meta-analysis itself, as differentiated from the trials it summarizes, can be biased. The two primary forms of this bias are publication bias and outcome reporting bias. Publication bias results from the likelihood that trials with

statistically significant results are more likely to be published. Outcome reporting bias results when researchers report only the significant results from a trial and leave unsignificant outcomes unreported. Because the data comes from CENTRAL, this study has already taken advantage of all of the Cochrane vitamin D research group's work to identify missing trials. The methods required to thoroughly investigate publication and reporting bias require resources beyond those available to the author. However, assuming that this study's summary analysis finds a beneficial effect, the author will compute Orwin's Fail-safe *N*, with a specified overall RR of 0.95 and a mean effect in the missing studies of zero, to determine how many missing, no-effect studies would nullify the results of the summary analysis.

Moreover, there is an additional source of meta-bias that the author will be able to examine. Cochrane and many other systematic reviewers don't include trials in which the control group received a smaller dose of supplementation than the intervention group. These reviews require a zero dose in the control group. This type of bias could be called Institutional Review Board or *IRB bias*. High-quality studies from major institutions are dropped from systematic reviews because an IRB insisted that the control group receive the standard-of-care dose of vitamin D supplementation. This is a particular problem with studies of pregnancy and infants. The author will do a subgroup analysis comparing trials by type of control group to investigate whether IRB bias can be detected.

Confidence in cumulative evidence. The strengths (specific intervention, large group of trials and outcomes) and limitations (single author, single database) of this study will be noted in published results, including author rating reliability over time and an analysis of the data collected on the risk of bias in individual studies.

Timeline

March 10, 2017 – Proposal approved

March 17, 2017 – IRB approval granted; Project registered in PROSPERO

March 31, 2017 – CENTRAL search completed, records loaded into *open-meta.org*

May 31, 2017 – Stage 1 review complete

November 30, 2017 – Data collection complete

January 31, 2018 – Analysis complete

March 31, 2018 – Dissertation write-up complete

April 2018 – Dissertation defense

References

Bjelakovic, G., Gluud, L. L., Nikolova, D., Whitfield, K., Wetterslev, J., Simonetti, R. G., . . . Gluud, C. (2014). Vitamin D supplementation for prevention of mortality in adults. *Cochrane Database of Systematic Reviews*, (1). Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007470.pub3/abstract doi:10.1002/14651858.CD007470.pub3

Blois, M. S., Blum, H. F., & Loomis, W. F. (1968). Vitamin-D sunlight and natural selection. *Science, 159*(3815), 652-&. doi:10.1126/science.159.3815.652

Booth, A., Clarke, M., Dooley, G., Ghersi, D., Moher, D., Petticrew, M., & Stewart, L. (2012). The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev, 1*, 2. doi:10.1186/2046-4053-1-2

Chalmers, I., Hedges, L. V., & Cooper, H. (2002). A brief history of research synthesis. *Eval Health Prof, 25*(1), 12-37. doi:10.1177/0163278702025001003

Chaplin, G., & Jablonski, N. G. (2009). Vitamin D and the Evolution of Human Depigmentation. *Am J Phys Anthropol, 139*(4), 451-461. doi:10.1002/ajpa.21079

Durazo-Arvizu, R. A., Camacho, P., Bovet, P., Forrester, T., Lambert, E. V., Plange-Rhule, J., . . . Luke, A. (2014). 25-Hydroxyvitamin D in African-origin populations at varying latitudes challenges the construct of a physiologic norm. *Am J Clin Nutr, 100*(3), 908-914. doi:10.3945/ajcn.113.066605

Fuller, K. E. (2003). Health disparities: reframing the problem. *Med Sci Monit, 9*(3), SR9-15.

Glass, G. V. (2015). Meta-analysis at middle age: a personal history. *Res Synth Methods, 6*(3), 221-231. doi:10.1002/jrsm.1133

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Res Synth Methods, 1*(1), 39-65. doi:10.1002/jrsm.5

Holick, M. F. (2007). Vitamin D deficiency. *N Engl J Med, 357*(3), 266-281.

Hollis, B. W., & Wagner, C. L. (2013). The Role of the Parent Compound Vitamin D with Respect to Metabolism and Function: Why Clinical Dose Intervals Can Affect Clinical Outcomes. *Journal of Clinical Endocrinology & Metabolism, 98*(12), 4619-4628. doi:10.1210/jc.2013-2653

Houghton, L. A., & Vieth, R. (2006). The case against ergocalciferol (vitamin D2) as a vitamin supplement. *Am J Clin Nutr, 84*(4), 694-697.

Jablonski, N. G. (2004). The evolution of human skin and skin color. *Annu Rev Anthropol, 33*, 585-623. doi:10.1146/annurev.anthro.33.070203.143955

Jablonski, N. G., & Chaplin, G. (2000). The evolution of human skin coloration. *Journal of Human Evolution, 39*(1), 57-106. doi:10.1006/jhev.2000.0403

Japelt, R. B., & Jakobsen, J. (2013). Vitamin D in plants: a review of occurrence, analysis, and biosynthesis. *Front Plant Sci, 4*, 136. doi:10.3389/fpls.2013.00136

LeBlanc, E., Chou, R., Zakher, B., Daeges, M., & Pappas, M. (2014). U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews *Screening for Vitamin D Deficiency: Systematic Review for the U.S. Preventive Services Task Force Recommendation*. Rockville (MD): Agency for Healthcare Research and Quality (US).

Loomis, W. F. (1967). Skin-pigment regulation of vitamin-D biosynthesis in man. *Science, 157*(3788), 501-&. doi:10.1126/science.157.3788.501

Luxwolda, M. F., Kuipers, R. S., Kema, I. P., Dijck-Brouwer, D. A. J., & Muskiet, F. A. J. (2012). Traditionally living populations in East Africa have a mean serum 25-hydroxyvitamin D concentration of 115 nmol/l. *British Journal of Nutrition, 108*(9), 1557-1561. doi:10.1017/s0007114511007161

Murray, F. G. (1934). Pigmentation, sunlight, and nutritional disease. *American Anthropologist, 36*(3), 438-445. doi:10.1525/aa.1934.36.3.02a00100

Newberry, S. J., Chung, M., Shekelle, P. G., Booth, M. S., Liu, J. L., Maher, A. R., . . . Balk, E. M. (2014). Vitamin D and calcium: a systematic review of health outcomes. *Evid Rep Technol Assess (Full Rep)*(217), 1-929.

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., . . . Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *Bmj, 349*, g7647. doi:10.1136/bmj.g7647

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *Am Psychol, 32*(9), 752-760.

Tjellesen, L., Hummer, L., Christiansen, C., & Rodbro, P. (1986). Serum concentration of vitamin-D metabolites during treatment with vitamin-D2 and vitamin-D3 in normal premenopausal women. *Bone and Mineral, 1*(5), 407-413.

Tuohimaa, P., Wang, J. H., Khan, S., Kuuslahti, M., Qian, K., Manninen, T., . . . Lou, Y. R. (2013). Gene expression profiles in human and mouse primary cells provide new insights into the differential actions of vitamin D metabolites. *Plos One, 8*(10), e75338. doi:10.1371/journal.pone.0075338

U. S. Department of Health and Human Services. (2014). Reports of the Surgeon General *The Surgeon General's call to action to prevent skin cancer*. Washington (DC): Office of the Surgeon General (US).

VitaminDWiki. (2016). Meta-analysis of Vitamin D. Retrieved from http://www.vitamindwiki.com/Meta-analysis+of+Vitamin+D

Weishaar, T., Rajan, S., & Keller, B. (2016). Probability of Vitamin D Deficiency by Body Weight and Race/Ethnicity. *J Am Board Fam Med, 29*(2), 226-232. doi:10.3122/jabfm.2016.02.150251

Weishaar, T., & Vergili, J. M. (2013). Vitamin D status is a biological determinant of health disparities. *J Acad Nutr Diet, 113*(5), 643-651. doi:10.1016/j.jand.2012.12.011

Whitney, E., & Rolfes, S. R. (2008). *Understanding nutrition* (11 ed.). Belmont, CA: Thomson Higher Education.

Wolf, G. (2004). The discovery of vitamin D: the contribution of Adolf Windaus. *J Nutr, 134*(6), 1299-1302.

Yuen, A. W. C., & Jablonski, N. G. (2010). Vitamin D: In the evolution of human skin colour. *Med Hypotheses, 74*(1), 39-44. doi:10.1016/j.mehy.2009.08.007

Appendix – Data fields required

Stage 1/2 Review categories:

- not available in English
- not a randomized controlled trial
- no valid participants
- no valid intervention – not daily
- no valid intervention – not D_3
- no valid intervention – other
- no valid comparison group
- no valid outcome
- missing mandatory data (Stage 2 only)

Data about outcomes:

- physical or mental?
- health or performance?
- benefit or harm?
- related to health disparities?
- reverse scored?

Data about trials:

- how discovered (CENTRAL or…)
- part of a larger study?
- industry funding – (risk of bias: high, low, unclear)
- randomization and allocation concealment – (risk of bias: high, low, unclear)
- blinding of participants – (risk of bias: high, low, unclear)
- blinding of research personnel and outcome assessors – (risk of bias: high, low, unclear)
- level of attrition and exclusions – (risk of bias: high, low, unclear)

Data about arms:

- *outcom*e
- *how the outcome is reported*
- *duration of supplementation*
- participant diagnosis – (healthy or…)
- participant pregnancy status – (pregnant only, non-pregnant only, either)
- participant specific risks – (e.g., “at risk for depression”)
- participant co-treatments – (e.g., “patients taking corticosteroids”)
- type of control group – (no intervention, placebo, standard-of-care)
- research setting – (community-dwelling, assisted living, long-term care, hospice)
- latitude
- how administered (by mouth, by injection, other (specify)

Data about groups:

- *daily vitamin D_3 dose*
- *outcome measure (central tendency)*
- *outcome measure (variance)*
- *number of participants – baseline*
- *number of participants – at outcome measurement*
- *25(OH)D status at outcome – mean*
- 25(OH)D status at outcome – SD
- 25(OH)D status at baseline – mean and SD
- age – mean and SD (preferably at baseline but acceptable at outcome)
- body weight – mean and SD (preferably at baseline but acceptable at outcome)
- percent female – (preferably at baseline but acceptable at outcome)