# **Title: Genetic evidence for vitamin D and cardiovascular disease: choice of variants is critical**

***Stephen Burgess PhD1,2, Dipender Gill BMBCh PhD3,4,5,6***

## **Affiliations**

1Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK.

2Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

3Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.

4Clinical Pharmacology and Therapeutics Section, Institute of Medical and Biomedical Education and Institute for Infection and Immunity, St George’s, University of London, London, UK.

5Clinical Pharmacology Group, Pharmacy and Medicines Directorate, St George’s University Hospitals NHS Foundation Trust, London, UK.

6Genetics Department, Novo Nordisk Research Centre Oxford, Old Road Campus, Oxford, UK.

## **Address for correspondence**

Dr Stephen Burgess

MRC Biostatistics Unit, East Forvie Building, University of Cambridge, Cambridge, CB2 0SR, UK.

Telephone: +44 (0)1223 768259

E-mail: sb452@medschl.cam.ac.uk.

## **Conflicts of interest**

DG is employed part-time by Novo Nordisk. SB has no conflicts of interest to declare.

**Word count:** 1290 (+ 1 Table)

Mendelian randomization is an epidemiological technique that compares disease risk in groups of individuals defined based on their genetic variants to make causal inferences (1). Rather than comparing individuals with high levels of an exposure versus those with low levels of an exposure as in a conventional epidemiological analysis, the approach compares those with genetic variants predisposing them to increased versus decreased levels of the exposure. This is analogous to the analysis of a randomized controlled trial for a treatment that increases levels of the exposure, which does not compare those with high versus low levels of the exposure after treatment, but rather compares those who were randomly assigned to receive the treatment versus those who were randomly assigned to the control group. The logic is that randomization should be independent of all competing risk factors, meaning that the randomly-assigned groups only differ systematically with respect to their average levels of the treatment and any downstream consequences of the treatment. Hence, an association between randomization and the trial outcome is indicative of a causal effect of the treatment. In the same way, if genetic variants act analogously to randomization by dividing the population into groups that differ systematically only with respect to the exposure and its consequences, then an association between the genetic variants and the outcome is indicative of a causal effect of the exposure (2).

Non-linear Mendelian randomization is an extension to standard Mendelian randomization that first stratifies the population based on levels of the exposure, and then conducts separate Mendelian randomization analyses within each stratum (3, 4). One important methodological point is that stratifying on the exposure directly would break randomization and lead to biased estimates in the strata (5). This is because the distribution of the genetic variants would no longer be the same within each stratum, as genetic variants predisposing individuals to higher levels of the exposure would be more common in strata with high levels of the exposure and less common in strata with low levels of the exposure. For this reason, it is recommended to first regress the exposure on the genetic variants, and stratify on residual values of the exposure, as this “residual exposure” is independent of the genetic variants, and hence randomization still holds within strata of the residual exposure (3, 4).

In their paper, Zhao, Selvanayagam, and Hyppönen considered genetic predictors of 25-hydroxyvitamin D [25(OH)D], a clinical biomarker used to measure vitamin D status, and assessed whether these variants were associated with cardiovascular disease risk in individuals of European genetic ancestry from the UK Biobank dataset (6). They observed inverse associations in overall analyses, but particularly strong inverse associations in non-linear analyses for strata of the population with low levels of 25(OH)D. The non-linear analyses indicated a threshold-like relationship, suggesting greater benefit of vitamin D supplementation for those who are vitamin D deficient, and less benefit for those with adequate vitamin D levels.

Results are similar to those recently published by Sofianopoulou *et al*, which showed threshold relationships for all-cause and cardiovascular mortality in individuals of European genetic ancestry from the UK Biobank dataset (as well as 3 smaller datasets), with evidence supporting a beneficial effect of higher 25(OH)D in individuals that are vitamin D deficient up to a threshold of around 40 nmol/L, but with no evidence of benefit when considering the population as a whole (7). Although they used broadly the same data, there are two major methodological differences between these studies. First, the Zhao paper performed a polygenic analysis considering 35 genetic variants from multiple gene regions (referred to as a “polygenic score”), whereas the Sofianopoulou paper considered variants from 4 gene regions that have biological links to vitamin D synthesis and metabolism (referred to as a “focused score”). Secondly, the Zhao paper considered a wider outcome definition, incorporating coronary artery disease, stroke, and peripheral vascular disease, whereas the Sofianopoulou paper considered coronary heart disease and stroke as separate outcomes, as well as cardiovascular mortality. Analyses in the Sofianopoulou paper for coronary heart disease and stroke provided null associations overall, but inverse associations in the deficient stratum, but with less statistical precision arising from the lower number of events due to considering the outcomes separately. Table 1 summarizes differences in the methodology and results between the studies.

One concern in the Zhao paper is potential pleiotropy via low-density lipoprotein (LDL)-cholesterol, which is implicated in cardiovascular disease pathophysiology. While the authors conducted analyses to assess pleiotropic associations with several covariates, they did not specifically consider LDL-cholesterol. Taking genetic associations in White British participants from UK Biobank reported by Neale *et al* (8), three of the 35 variants in the Zhao paper were associated with LDL-cholesterol at p<5×10-8 (rs7528419, rs78151190, and rs261291), and a further four variants were associated at p<0.005 (rs6671730, rs6123359, rs17216707, rs2585442). A polygenic score was also considered in the Sofianopoulou paper as a secondary analysis; however, this analysis was also affected by pleiotropy via LDL-cholesterol. The polygenic analysis of Sofianopoulou found an overall association with coronary artery disease that attenuated in a multivariable Mendelian randomization model adjusting for genetically predicted LDL-cholesterol and triglyceride levels. However, evidence for an inverse association with stroke in vitamin D deficient individuals in the polygenic analysis persisted despite the same adjustment; while the estimate in the deficient stratum attenuated on adjustment, evidence for an inverse association remained. This suggests that associations indicated by polygenic analyses in the deficient stratum may be robust to pleiotropy, but associations in the overall population may not be robust to pleiotropy via LDL-cholesterol.

A further potential concern with the Zhao paper is the presentation of population impact factors. Genetic estimates in Mendelian randomization represent differences in disease risk between groups having lifelong differences in the distribution of the exposure (9). Replicating such estimates in practice would require lifelong interventions to increase 25(OH)D levels. Hence the achievable population benefit for cardiovascular risk reduction through vitamin D supplementation is likely overestimated (10). There are also potential harms of high levels of vitamin D supplementation for those with already adequate levels of 25(OH)D (11). These factors suggest that any clinical trial to demonstrate benefits of vitamin D supplementation should be large in size, focused on a vitamin D deficient population, and have a long follow-up period. Contamination by vitamin D supplementation in the control group would further potentially limit the effectiveness of the trial. Such a trial may therefore be difficult to implement in practice.

Overall, the finding of the Zhao paper supporting that vitamin D supplementation reduces risk of cardiovascular disease for vitamin D deficient individuals is plausible, and is supported by the existing literature (7). The association of genetically-predicted levels of 25(OH)D with cardiovascular disease risk estimated by Zhao in deficient individuals is so large in magnitude that it is unlikely to be wholly attributable to pleiotropic associations of the genetic variants in the analysis. In contrast, the notion that vitamin D reduces risk of cardiovascular disease in individuals with adequate levels of 25(OH)D is more questionable, as findings in the genetic analyses could be fully explained by pleiotropic associations of variants with LDL-cholesterol. Generally speaking, it is recommended that Mendelian randomization studies consider a biologically-informed choice of genetic variants where possible, and if not possible, consider pleiotropic associations with key determinants of disease risk (12). For cardiovascular disease, this should typically include LDL-cholesterol. For this reason, we would favour Mendelian randomization analyses with vitamin D as an exposure that use a more focused set of genetic variants, rather than ones that include all genome-wide significant predictors of 25(OH)D.

In conclusion, this investigation strengthens the evidence base for vitamin D supplementation as a beneficial intervention to reduce cardiovascular disease risk in individuals with low vitamin D levels. Future research should aim to consider other disease outcomes and population groups, particularly those with low average levels of vitamin D.

FUNDING

SB is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (204623/Z/16/Z). This research was supported by the National Institute for Health Research Cambridge Biomedical Research Centre (BRC-1215-20014). The views expressed are those of the authors and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care.

ACKNOWLEDGEMENT

The authors would like to thank Lucy Grabiner for help with creating the graphical abstract.

**Table.** Differences between the Zhao and Sofianopoulou paper in methodology and results.

|  |  |  |
| --- | --- | --- |
| **Factor** | **Zhao paper** | **Sofianopoulou paper** |
| Choice of variants | 35 variants associated with 25(OH)D levels at genome-wide significant level | Variants in 4 gene regions chosen due to biological link with vitamin D synthesis or metabolism (“focused score”); a secondary analysis using 71 genome-wide significant variants was also performed |
| Pleiotropic associations | Several genetic variants had pleiotropic associations with LDL-cholesterol | Minimal evidence of pleiotropy for primary analysis; substantial pleiotropy via LDL-cholesterol and triglycerides for secondary analysis using genome-wide variants |
| Cardiovascular outcomes | Cardiovascular disease (combined outcome of coronary artery disease, stroke, and peripheral vascular disease) | Separate analyses for coronary heart disease, stroke, and cardiovascular mortality |
| Results | Strong associations of genetically-predicted 25(OH)D with outcomes in vitamin D deficient strata; weak (but non-null) associations overall | **Primary analysis:** Associations of genetically-predicted 25(OH)D with outcomes in vitamin D deficient strata (particularly for cardiovascular mortality); null associations overall  **Secondary analysis:** Weak association with coronary heart disease in overall analysis that attenuated toward null on adjustment for LDL-cholesterol and triglycerides; stronger association with stroke in deficient stratum that persisted despite adjustment for LDL-cholesterol and triglycerides |

**References:**

1. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? Brit Med J. 2005;330(7499):1076-9.

2. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. Statistical Methods in Medical Research. 2007;16(4):309-30.

3. Burgess S, Davies NM, Thompson SG. Instrumental variable analysis with a nonlinear exposure–outcome relationship. Epidemiology. 2014;25(6):877.

4. Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure‐outcome relationship using instrumental variables with application to Mendelian randomization. Genetic Epidemiology. 2017;41(4):341-52.

5. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. International Journal of Epidemiology. 2010;39(2):417-20.

6. Zhao A, Selvanayagam J, Hyppönen E. Non-linear Mendelian randomization analyses support a role for vitamin D deficiency in cardiovascular disease risk. Eur Heart J. 2021.

7. Sofianopoulou E, Kaptoge S, Afzal S, Jiang T, Gill D, Gundersen TE, et al. Estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: observational and Mendelian randomisation analyses. Lancet Diabetes & Endocrinology. 2021.

8. Neale B, Abbott L, Bryant S, Churchhouse C, Ganna A, Howrigan D, et al. UK Biobank GWAS results Round 2 2018 [Available from: <http://www.nealelab.is/uk-biobank>].

9. Ference BA. How to use Mendelian randomization to anticipate the results of randomized trials. European Heart Journal. 2018;39(5):360-2.

10. Burgess S, Butterworth A, Malarstig A, Thompson SG. Use of Mendelian randomisation to assess potential benefit of clinical intervention. British Medical Journal. 2012;345:e7325.

11. Marcinowska-Suchowierska E, Kupisz-Urbańska M, Łukaszkiewicz J, Płudowski P, Jones G. Vitamin D Toxicity–A Clinical Perspective. Frontiers in Endocrinology. 2018;9:550.

12. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations. Wellcome Open Research. 2019;4:186.

GRAPHICAL ABSTRACT

Graphical user interface, text, application

Description automatically generated