Mendelian randomisation and vitamin D: the importance of model assumptions

There are issues of concern raised by analyses presented by the Emerging Risk Factors Collaboration, EPIC-CVD, and Vitamin D Studies Collaboration in "Estimating dose-response relationships for vitamin D with cardiovascular mortality, stroke, and all-cause mortality: observational and Mendelian randomisation analyses", published in *The Lancet Diabetes & Endocrinology*.[1] The bottom panel of figure 2 of the Article (appendix) presents the estimated causal effect of 10nmol/L higher vitamin D on all-cause mortality. There is a precisely estimated null overall effect, yet in every strata of residual vitamin D concentration the effects are in a protective direction, and substantially so in the "deficient" strata.

See **Online** for appendix

Such a pattern of findings, with the overall central effect estimate lying outside of all four subgroup central effect estimates, is simply not plausible. When combining a causal interpretation of Simpson's paradox[2] with the intention of Mendelian randomisation[3] to estimate causal effects, it is not possible to envisage situations in which the key assumptions of the Mendelian randomisation analysis[3] are not violated, the assumptions regarding other causal effects are consistent with what is known, and the data are distributed as presented. The central finding of the paper—a detrimental effect of low concentrations of vitamin D—was said to have "important public health and clinical consequences" in an accompanying commentary in *The Lancet Diabetes & Endocrinology*.[4] Similar probably spurious findings have been reported in two other papers[5,6] using the same analytical approach and overlapping datasets from which impossible estimates were generated in the paper from EPIC-CVD.[1] The method is also being widely applied in many other contexts. When evaluating findings from such analyses, considerable caution should be applied.

I declare no competing interests.

George Davey Smith
kz.davey-smith@bristol.ac.uk

MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

1 Emerging Risk Factors Collaboration/EPIC-CVD/Vitamin D Studies Collaboration. Estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: observational and Mendelian randomisation analyses. *Lancet Diabetes Endocrinol* 2021; **9:** 837–46.
2 Hernán MA, Clayton D, Keiding N. The Simpson's paradox unraveled. *Int J Epidemiol* 2011; **40:** 780–85.
3 Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nat Rev Methods Primers* 2022; **2:** 6.
4 Butler-Laporte G, Richards JB. Targeting of vitamin D supplementation to individuals with deficiency. *Lancet Diabetes Endocrinol* 2021; **9:** 803–04.
5 Zhou A, Selvanayagam JB, Hyppönen E. Non-linear Mendelian randomization analyses support a role for vitamin D deficiency in cardiovascular disease risk. *Eur Heart J* 2022; **43:** 1731–39.
6 Sutherland JP, Zhou A, Hyppönen E. Vitamin D deficiency increases mortality risk in the UK Biobank : A Nonlinear Mendelian Randomization Study. *Ann Intern Med* 2022; **175:** 1552–59.

In a previous issue of the journal, we had commented[1] on the use of a novel Mendelian randomisation method that tested for non-linear effects of vitamin D and its measured metabolite 25-hydroxyvitamin D (25[OH]D), on diseases, as reported by the Emerging Risk Factors Collaboration, EPIC-CVD, and Vitamin D Studies Collaboration.[2] One of their findings was an association between the genetic score and mortality outcomes in a strata of the study population with low concentrations of 25(OH)D. Mendelian randomisation is a genetic epidemiology tool that evaluates the causal effect of an exposure (here, 25[OH]D) on a disease by limiting confounding and reverse causation bias. This is important, since the role of 25(OH)D on human disease remains controversial, especially in people with low concentrations of 25(OH)D, who were under-represented in large vitamin D supplementation trials such as the VITAL study.[3] Thus, the effects of increasing 25(OH)D concentrations in people with low concentrations is not well known and could be explored using novel Mendelian randomisation methods.

Although Mendelian randomisation has generally been able to replicate, or anticipate, results from randomised trials, in this study,[2] the method makes important assumptions about the relationship between the genetic variants it uses as instrument variables, 25(OH)D concentrations, and the disease outcome of interest. In assessing the non-linear effects of 25(OH)D on disease, the method used in this study[2] also made the additional assumption that the effect of genetic variants on concentration of 25(OH)D was constant across strata of 25(OH)D concentrations.

When using a different, more robust non-linear Mendelian randomisation method, it appears that this additional assumption was not fulfilled and there was no evidence of a beneficial effect of vitamin D supplementation in people with low concentrations of 25(OH)D. Hence, our previous interpretation of these results was wrong.

Mendelian randomisation has revolutionised epidemiology, allowing for genetics to provide causal insights in human disease, with novel therapeutic development translational opportunities. However, research in Mendelian randomisation is a rapidly evolving field, with constant progress being made. It is therefore important to remember that Mendelian randomisation relies on assumptions of which consequences can take time to fully understand. We are happy to see that the authors were rapid to rectify their previous observations