Estimating dose-response relationships for vitamin D with coronary heart disease, stroke, and all-cause mortality: observational and Mendelian randomisation analyses

Emerging Risk Factors Collaboration/EPIC-CVD/Vitamin D Studies Collaboration

Summary

Background Randomised trials of vitamin D supplementation for cardiovascular disease and all-cause mortality have generally reported null findings. However, generalisability of results to individuals with low vitamin D status is unclear. We aimed to characterise dose-response relationships between 25-hydroxyvitamin D (25[OH]D) concentrations and risk of coronary heart disease, stroke, and all-cause mortality in observational and Mendelian randomisation frameworks.

Methods Observational analyses were undertaken using data from 33 prospective studies comprising 500 962 individuals with no known history of coronary heart disease or stroke at baseline. Mendelian randomisation analyses were performed in four population-based cohort studies (UK Biobank, EPIC-CVD, and two Copenhagen population-based studies) comprising 386 406 middle-aged individuals of European ancestries, including 33 546 people who developed coronary heart disease, 18 166 people who had a stroke, and 27 885 people who died. Primary outcomes were coronary heart disease, defined as fatal ischaemic heart disease (International Classification of Diseases 10th revision code I20-I25) or non-fatal myocardial infarction (I21-I23); stroke, defined as any cerebrovascular disease (I60-I69); and all-cause mortality.

Findings Observational analyses suggested inverse associations between incident coronary heart disease, stroke, and all-cause mortality outcomes with 25(OH)D concentration at low 25(OH)D concentrations. In population-wide genetic analyses, there were no associations of genetically-predicted 25(OH)D with coronary heart disease, stroke, or all-cause mortality. However, for the participants with vitamin D deficiency (25[OH]D concentration <25 nmol/L), genetic analyses provided strong evidence for an inverse association with all-cause mortality (odds ratio [OR] per 10 nmol/L increase in genetically-predicted 25[OH]D concentration 0·69 [95% CI 0·59–0·80]; p<0·0001) and non-significant inverse associations for stroke (0·85 [0·70–1·02], p=0·09) and coronary heart disease (0·89 [0·76–1·04]; p=0·14). A finer stratification of participants found inverse associations between genetically-predicted 25(OH)D concentrations and all-cause mortality up to around 40 nmol/L.

Interpretation Stratified Mendelian randomisation analyses suggest a causal relationship between 25(OH)D concentrations and mortality for individuals with low vitamin D status. Our findings have implications for the design of vitamin D supplementation trials, and potential disease prevention strategies.

Funding British Heart Foundation, Medical Research Council, National Institute for Health Research, Health Data Research UK, Cancer Research UK, and International Agency for Research on Cancer.

Lancet Diabetes Endocrinol 2021; 9: 837–46

Published Online October 27, 2021 http://doi.org/10.1016/S2213-8587(21)00263-1

See Comment page 803

Correspondence to: Dr Stephen Burgess, MRC Biostatistics Unit, East Forvie Building, University of Cambridge, Cambridge CB2 0SR, UK
sb452@medschl.cam.ac.uk

Introduction

Vitamin D is an essential nutrient obtained from sunlight, dietary intake, and supplementation.[1] Observational epidemiological studies have consistently found that low concentrations of circulating 25-hydroxyvitamin D (25[OH]D), a metabolite used as a clinical indicator of vitamin D status, are associated with an increased risk of cardiovascular disease and all-cause mortality, as well as other chronic diseases.[2,3] However, several large randomised trials of vitamin D supplementation have reported null results,[4–6] casting doubt on the observational evidence. However, as trials have typically recruited participants irrespective of baseline 25(OH)D concentration, they have not been sufficiently powered to test supplementation effects in subgroups with low 25(OH)D concentrations.[7]

An efficient approach for assessing the potential causal effect of vitamin D supplementation is Mendelian randomisation. Mendelian randomisation uses genetic variants specifically related to a particular exposure to compare genetically-defined population subgroups with different average levels of the exposure. The independent segregation of alleles at conception means that these genetically-defined subgroups should not differ systematically with respect to confounding variables, creating a natural experiment analogous to a randomised

Research in context

Evidence before this study

We searched for randomised trials investigating the effects of vitamin D supplementation on cardiovascular disease or all-cause mortality published in any language from inception up until April 16, 2021, in PubMed, Scientific Citation Index Expanded, and Embase using the search terms listed in the appendix. An example search is ("Cardiovascular Diseases"[MeSH] OR "Cardiovascular Disease" OR "All-cause Mortality" OR "Mortality" OR "Survival"). However, no study had combined the following key features required to achieve reliable estimates of the effect of vitamin D supplementation on cardiovascular disease or mortality in vitamin D deficient individuals: availability of individual-participant data on large sample sizes (eg, >1000 individuals with 25-hydroxyvitamin D [25(OH)D] concentrations of less than 25 nmol/L) with moderate numbers of cardiovascular disease or all-cause mortality events (eg, > 100 events). In summary, large randomised controlled trials and mendelian randomisation investigations have not been able to show beneficial effects of vitamin D supplementation on risk of coronary heart disease, stroke, or mortality.

Added value of this study

We characterised dose-response relationships for 25(OH)D, a metabolite used as a clinical indicator of vitamin D status, with cardiovascular disease and mortality in the largest such epidemiological analysis done to date, using statistical methods tailored to the investigation of non-linear risk factor-disease relationships. Our observational analyses indicate threshold relationships between 25(OH)D concentrations and risk of both cardiovascular disease and all-cause mortality. Genetic analyses support a dose-dependent inverse causal effect of 25(OH)D concentrations on all-cause mortality risk up to a threshold of around 40 nmol/L. Estimates for coronary heart disease and stroke risk at low 25(OH)D concentrations were in the inverse direction, but generally non-significant.

Implications of all the available evidence

Taken together, these findings suggest that vitamin D supplementation could reduce mortality risk, but only in individuals with low 25(OH)D concentrations.

See Online for appendix

trial.[8] Therefore, Mendelian randomisation analyses provide more reliable insights into causal relationships between risk factors and disease outcomes than conventional observational analyses. Previous Mendelian randomisation analyses have reported null associations of genetically-predicted 25(OH)D concentrations with coronary heart disease[9,10] and ischaemic stroke.[11,12] An inverse association has been observed between genetically-predicted 25(OH)D and all-cause mortality.[13] Null findings have been observed for several further outcomes, including other cardiovascular diseases and cancers.[14]

Previous Mendelian randomisation analyses assumed a linear dose-response relationship between genetically-predicted 25(OH)D and cardiovascular disease. However, some observational analyses have reported non-linear associations,[15,16] suggesting methods that assume linearity might not provide an accurate picture of the dose-response relationship. In this study, we performed the largest observational analysis to date to characterise the shape of association between 25(OH)D concentrations and cardiovascular disease outcomes in an individual participant data meta-analysis of 33 prospective studies. We then did stratified Mendelian randomisation analyses to assess evidence for potential causal effects of 25(OH)D concentrations on risk of major cardiovascular disease outcomes including coronary heart disease and stroke, all-cause mortality, and cause-specific mortality for population subgroups with different 25(OH)D concentrations.

Methods

Study design and participants

We undertook observational analyses using data from UK Biobank, the European Prospective Investigation into Cancer and Nutrition Cardiovascular Disease study (EPIC-CVD), and 31 studies from the Vitamin D Studies Collaboration (VitDSC). Genetic analyses were done using data from UK Biobank, EPIC-CVD, and two Copenhagen population-based studies.

UK Biobank is a prospective cohort study of around 500 000 people aged 40 to 69 years at baseline, recruited in 2006–10 from the UK and followed up for a median of 10·9 years (IQR 10·1–11·7).[17] For observational analyses, we analysed 384 721 individuals with a valid 25(OH)D measurement and without previous known cardiovascular disease at baseline. For genetic analyses, we included data on 333 002 unrelated individuals of European ancestries with a valid 25(OH)D measurement and genetic data that passed quality control steps described previously.[18] EPIC-CVD is a case-cohort study derived from a cohort of over 500 000 individuals recruited at 23 centres across 10 European countries.[19,20] Participants were followed up for a median of 9·5 years (IQR 6·1–12·6). We analysed 26 336 individuals with a valid 25(OH)D measurement and without previous known cardiovascular disease at baseline in observational analyses, and 22 142 individuals of European ancestry in genetic analyses with a valid 25(OH)D measurement and genetic data that passed quality control steps. VitDSC comprises 89 915 participants from 31 mostly population-based, prospective studies across 11 countries. We analysed individual participant data on 25(OH)D concentrations, conventional cardiovascular risk factors, and major incident cardiovascular morbidity and mortality for 67 992 individuals without previously known cardiovascular disease. The Copenhagen City Heart Study (CCHS) and Copenhagen General Population Study

(CGPS) are prospective cohort studies in the Danish population.[21,22] CCHS was initiated in 1976 and participants were followed up periodically until 2018. Median follow-up was 21·4 years (IQR 12·3–32·6). CGPS was initiated in 2003 and has a median follow-up of 8·8 years (IQR 8·1–13·6). For genetic analyses, we analysed a total of 31 262 individuals from both studies with genetic data and a 25(OH)D measurement. For all studies, written informed consent was obtained from participants and approval was obtained from relevant ethics committees.

Procedures

Concentrations of 25(OH)D in blood were measured using the Liaison immunoassay analyser (DiaSorin; Saluggia, Italy) in the UK Biobank and Copenhagen studies, and liquid chromatography-tandem mass spectrometry in the EPIC-CVD study. In VitDSC, concentrations were measured by radioimmunoassay, direct chromatographic approaches, or other immunoassays (appendix p 9). Stratification of participants by 25(OH)D concentrations was based on guidelines set by the National Institute for Health and Care Excellence[23] (adequate [>75 nmol/L], sufficient [50–75 nmol/L], insufficient [25–49 nmol/L], and deficient [<25 nmol/L]). Measurements were seasonally adjusted in each study to correspond to a measurement taken in autumn by subtracting the study-specific mean 25(OH)D concentration for the season the measurement was taken in and then adding the study-specific mean 25(OH)D concentration for autumn measurements. In EPIC-CVD, centre-specific means were used rather than study-specific means.

To minimise potential bias due to horizontal pleiotropy, we considered genetic variants from four gene regions previously shown to be strongly associated with 25(OH)D and implicated in the transport, metabolism, and synthesis of vitamin D:[24] *GC*, *DHCR7*, *CYP2R1*, and *CYP24A1*. The *GC* gene encodes vitamin D binding protein. The *DHCR7* gene product converts 7-dehydrocholesterol to cholesterol, reducing 7-dehydrocholesterol available for conversion to previtamin D_3 by solar radiation. The *CYP2R1* gene encodes vitamin D 25-hydroxylase, a regulator of 25(OH)D synthesis through 25-hydroxylation of vitamin D in the liver. The *CYP24A1* gene product inactivates the active form of vitamin D ($1\alpha25(OH)_2D$). To maximise the variance explained by the genetic instrument, we considered available variants at each genetic locus and selected variants associated with 25(OH)D concentrations using a stepwise selection method (appendix p 4). For UK Biobank and EPIC-CVD, 21 variants were included in the analysis (appendix p 11). For the Copenhagen studies, because of limited availability of genetic measurements, analyses were restricted to three variants: two from the *CYP2R1* locus (rs12794714 and rs117913124) and one from the *DHCR7* locus (rs7944926).

We also considered a score based on 71 genetic variants from across the genome (referred to as a genome-wide score) previously shown to be associated with 25(OH)D concentrations at a genome-wide level of statistical significance.[25]

Outcomes

Outcomes were classified using International Classification of Diseases, Tenth Revision (ICD-10) codes. Primary outcomes were coronary heart disease, defined as fatal ischaemic heart disease (ICD-10 code I20–I25) or non-fatal myocardial infarction (I21–I23); stroke, defined as any cerebrovascular disease (I60–I69); and all-cause mortality. We performed secondary analyses for cause-specific mortality divided into cardiovascular mortality, cancer mortality, or non-cardiovascular non-cancer mortality using ICD-10 codes (appendix p 10). Prespecified observational analyses included incident events only. We performed supplementary genetic analyses restricted to incident coronary heart disease and stroke events, and separating ischaemic stroke (I63–I64) and haemorrhagic stroke (I60–I61).

Statistical analysis

Observational associations were assessed by inverse-variance weighted random-effects meta-analysis of study-specific hazard ratios (HRs), calculated using Cox proportional hazards regression models stratified by sex and, as appropriate, centre or trial group. Primary analyses were adjusted for conventional risk factors, namely age at blood draw for 25(OH)D measurement, calendar month of blood draw, smoking status (current *vs* other), total cholesterol, HDL cholesterol, systolic blood pressure, known history of diabetes, and BMI.

The primary dose-response analyses assessed the shape of association between 25(OH)D and outcomes by meta-analysis of fractional polynomials adjusted for the conventional risk factors. Supplementary analyses combined study-specific HRs by tenths of 25(OH)D and plotted the pooled HRs against the pooled mean 25(OH)D within each tenth.

We calculated a genetic risk score weighted by the conditional associations of the genetic variants with 25(OH)D concentration in UK Biobank (appendix p 11). Mendelian randomisation estimates were calculated using the ratio method by dividing the genetic association with the outcome by the genetic association with 25(OH)D concentration and scaling the estimate to a 10 nmol/L difference in genetically-predicted 25(OH)D concentration. Genetic associations were estimated using logistic regression for disease outcomes and using linear regression for 25(OH)D concentrations. All regression models were adjusted for age at baseline, sex, centre (for UK Biobank and EPIC-CVD), and ten genetic principal components of ancestry. We assessed specificity of the genetic risk score by testing its associations with a range of cardiovascular risk factors in the UK Biobank study. We undertook sensitivity analyses using the UK Biobank data

	UK Biobank (n=333 002)	EPIC-CVD (n=22 142)	Copenhagen studies (n=31 262)
Age at baseline, years	57·1 (8·1)	54·8 (9·4)	57·5 (12·9)
Sex			
Female	177 733 (53·4)	11 426 (51·6)	17 311 (55·4)
Male	155 269 (46·6)	10 716 (48·4)	13 951 (44·6)
25(OH)D concentration, nmol/L	54·5 (19·6)	46·9 (16·4)	53·8 (25·9)
Residual 25(OH)D strata*			
Deficient (<25 nmol/L)	12 957 (3·9%)	1522 (6·9%)	1158 (3·7%)
Insufficient (25–49 nmol/L)	133 922 (40·2%)	11 845 (53·5%)	9730 (31·1%)
Sufficient (50–74 nmol/L)	138 605 (41·6%)	7717 (34·9%)	12 375 (39·6%)
Adequate (≥75 nmol/L)	47 518 (14·3%)	1058 (4·8%)	7999 (25·6%)
Coronary heart disease events	22 363 (6·7%)	5942 (26·8%)	5241 (16·8%)
Stroke events	10 489 (3·1%)	5478 (24·7%)	2199 (7·0%)
Deaths	20 340 (6·1%)	N/A†	7545 (24·1%)
BMI (kg/m²)	27·3 (4·8)	26·7 (4·3)	25·9 (4·2)
SBP (mm Hg)	137·5 (18·6)	137·4 (21·3)	140·1 (21·0)
Smoking‡			
Current	34 085 (10·2%)	6867 (31·0%)	8387 (26·9%)
Other	298 940 (89·8%)	15 275 (69·0%)	22 784 (73·1%)
Diabetes			
Known history	15 822 (4·8%)	1234 (5·6%)	1244 (4·0%)
No known history	317 203 (95·2%)	20 908 (94·4%)	30 018 (96·0%)

Data are mean (SD) for continuous variables or N (%) for categorical variables. 25(OH)D concentrations are season-shifted to correspond to a measurement taken in autumn. 25(OH)D=25-hydroxyvitamin D. CHD=coronary heart disease. SBP=systolic blood pressure. *Stratification of participants into categories of 25(OH)D concentrations was done using season-shifted measurements based on guidelines set by the UK National Institute for Health and Care Excellence. †EPIC-CVD was specifically designed as a case-cohort study of cardiovascular disease outcomes, and therefore does not contribute to the analysis of non-cardiovascular disease or all-cause mortality outcomes. ‡91 participants had missing data in the Copenhagen studies.

Table: **Baseline characteristics**

excluding variants from each of the four gene regions in turn to check whether results were driven by a single gene region.

In addition to analyses performed in the overall study sample to estimate population-averaged causal effects, we also undertook stratified analyses in strata of the population defined according to residual 25(OH)D. Residual 25(OH)D was calculated as the residual from regression of 25(OH)D on the mean-centred genetic risk score. By stratifying on residual 25(OH)D, we compare individuals in the population who would have 25(OH)D concentrations in the same stratum if they had the same genotype. We then calculated Mendelian randomisation estimates for each stratum (adequate, sufficient, insufficient, and deficient) using the ratio method with the genetic risk score as an instrumental variable, and combined stratum-specific estimates across studies using fixed-effect meta-analysis. We also divided each study sample into finer categories at 5 nmol/L intervals and meta-analysed stratum-specific ratio estimates to investigate any threshold in the potential effect of 25(OH)D.

Although stratification on residual 25(OH)D rather than 25(OH)D directly is important to avoid collider bias, correlation between residual 25(OH)D and 25(OH)D was 0·977. Hence, regarding the interpretation of estimates, distinction between stratifying on residual 25(OH)D and 25(OH)D directly is minimal. All statistical analyses were done in R version 3.4.3, Stata/SE 15.1, or BOLT-LMM version 2.3.4.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

386 406 participants from the four studies were included in genetic analyses (table), including 33 546 people who had coronary heart disease, 18 166 people who had a stroke, and 27 885 people who died, and 500 962 participants were included in observational analyses (appendix pp 12–14). Mean age of participants included in the genetic analysis ranged from 54·8 years (SD 9·4) to 57·5 years (12·9), with similar numbers of men and women in each study, and the mean season-shifted 25(OH)D concentrations (corresponding to an autumn measurement) were 54·5 nmol/L (SD 19·6) in UK Biobank, 46·9 nmol/L (16·4) in EPIC-CVD, and 53·8 nmol/L (25·9) in the Copenhagen studies. 12 957 (3·9%) of 333 002 people in the UK Biobank study and 1158 (3·7%) of 31 262 people in the Copenhagen studies were in the deficient category, compared with 1522 (6·9%) of 22 142 people in EPIC-CVD (appendix p 15). Mean 25(OH)D estimates did not notably differ by assay type (appendix p 24). The focused genetic risk score explained 4·7% of the variance in 25(OH)D concentrations in UK Biobank study, 5·8% in EPIC-CVD, and 1·8% in the Copenhagen studies (appendix p 25). This genetic risk score was not associated with a range of cardiovascular risk factors in UK Biobank, except for BMI and HDL cholesterol, although these associations were small (appendix p 26). The genome-wide score was strongly associated with LDL cholesterol and triglycerides (appendix p 27), and so Mendelian randomisation estimates using this score are unreliable.

Observational associations had a similar non-linear shape for all outcomes (figure 1; appendix p 28): at low concentrations of 25(OH)D, there was an inverse association with all outcomes, whereas at higher concentrations of 25(OH)D, the association was null for cardiovascular mortality and weakly positive for other mortality outcomes. For coronary heart disease and stroke, there was no strong association with 25(OH)D concentrations above 50 nmol/L, but a progressively steeper association was observed below this threshold. For all-cause mortality, in comparison to other outcomes, the strength of the inverse association at lower 25(OH)D concentrations was stronger and began at a higher 25(OH)D concentration.

The shapes of the observational associations in the three primary data sources were broadly similar

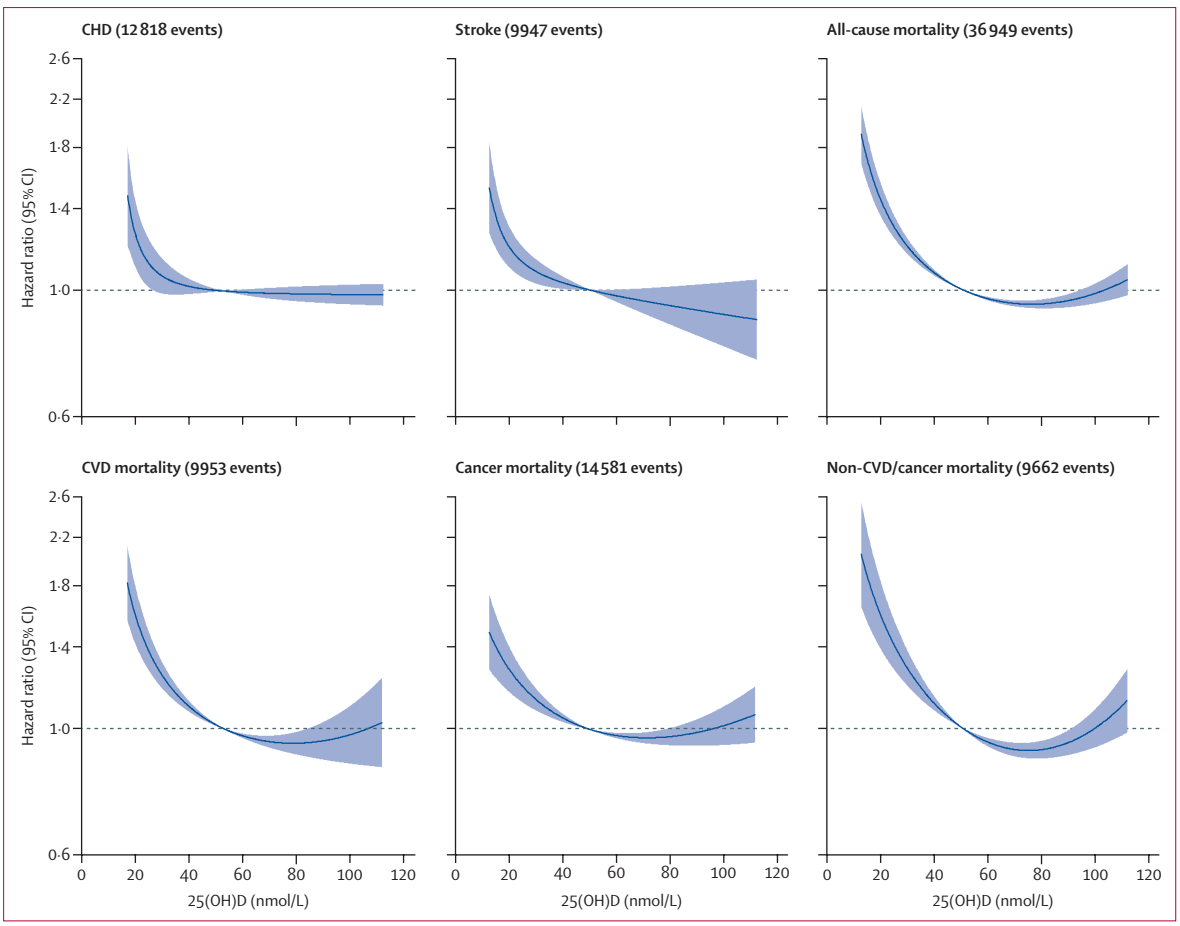

Figure 1: **Observational associations of 25(OH)D concentration with outcomes**
Reference value is 50 nmol/L. The shaded area represents the 95% CI for the dose—response curve. 25(OH)D=25-hydroxyvitamin D. CHD=coronary heart disease. CVD=cardiovascular disease.

(appendix p 29). Dose-response findings were also similar in supplementary analyses that combined study-specific HRs by deciles of 25(OH)D or according to the four 25(OH)D categories (appendix pp 16, 30).

In overall Mendelian randomisation analyses (that is, population-averaged estimates across the full range of the 25[OH]D concentration distribution), there was no association between genetically-predicted 25(OH)D and coronary heart disease (odds ratio [OR] 0·98 [95% CI 0·95–1·01]; p=0·18), stroke (OR 1·01 [0·97–1·05]; p=0·61), or all-cause mortality (OR 0·99 [0·95–1·02]; p=0·39; figure 2; appendix pp 17, 31). However, there was some evidence of an overall inverse association with all-cause mortality in the Copenhagen studies (OR 0·89 [0·80–0·99]; p=0·030; appendix p 17). In combined analyses within strata, there was strong evidence for an inverse association with all-cause mortality in the deficient stratum (OR 0·69 [95% CI 0·59–0·80]; p<0·0001), and non-significant inverse associations in the deficient stratum with coronary heart disease (OR 0·89, 95% CI 0·76–1·04, p=0·14) and stroke (OR 0·85, 95% CI 0·70–1·02, p=0·09). By contrast, estimates for other strata were much closer to the null (figure 2). Similar results were observed for supplementary analyses that considered incident stroke outcomes only and ischaemic stroke only (appendix p 18), and incident coronary heart disease outcomes only (appendix p 19). The association in the deficient stratum for haemorrhagic stroke was much closer to the null than for ischaemic stroke (appendix p 18). In sensitivity analyses excluding variants from each gene region in turn (appendix p 20), the association with all-cause mortality in the deficient stratum was attenuated, but remained nominally significant when excluding variants in the *CYP2R1* gene region from the analysis (OR 0·80 [95% CI 0·66–0·98]; p=0·027). No attenuation occurred when excluding variants from other gene regions. Estimates using the pleiotropic genome-wide score are presented in the appendix (p 21).

In Mendelian randomisation analyses for cause-specific mortality in UK Biobank and the Copenhagen studies, inverse associations in the deficient stratum were evident for cardiovascular mortality (OR 0·69 [95% CI 0·52–0·92]; p=0·011) and non-cardiovascular

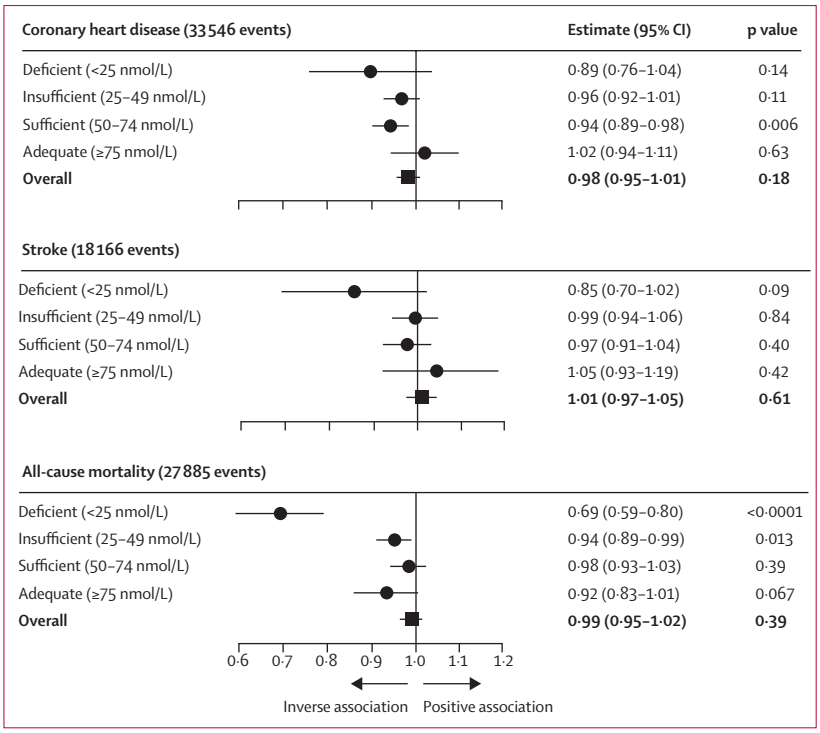

Figure 2: **Mendelian randomisation estimates for primary outcomes in overall population and strata of residual 25(OH)D concentrations**
Estimates (95% CIs) represent odds ratios per 10 nmol/L increase in genetically-predicted concentration of 25(OH)D in strata of the population defined by residual concentration of 25(OH)D.

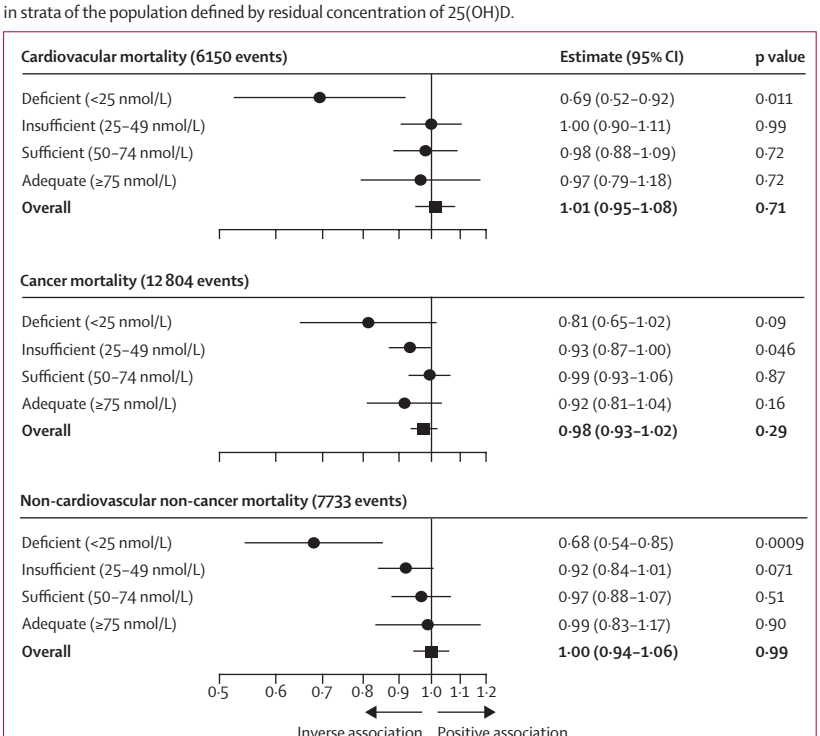

Figure 3: **Mendelian randomisation estimates for cause-specific mortality in overall population and strata of residual 25(OH)D concentrations**
Estimates (95% CIs) represent odds ratios per 10 nmol/L increase in genetically-predicted concentration of 25(OH)D in strata of the population defined by residual concentration of 25(OH)D. 25(OH)D=25-hydroxyvitamin D.

non-cancer mortality (OR 0·68 [0·54–0·85]; p=0·0009; figure 3; appendix pp 22, 32). A non-significant inverse association was observed for cancer mortality (OR 0·81, 95% CI 0·65–1·02, p=0·09). Again, overall estimates and estimates within other strata were generally null. When dividing non-cardiovascular non-cancer mortality into more specific mortality categories, the strongest associations in the deficient stratum were for diseases of the nervous system and the digestive system (appendix p 23).

In Mendelian randomisation analyses for a finer stratification of 25(OH)D concentrations, the dose-response curve for all-cause mortality has a clear threshold shape with evidence of an inverse association at 25(OH)D concentrations below 40 nmol/L, and null associations above 40 nmol/L (figure 4; appendix p 33); there was a notable trend in estimates below 40 nmol/L, with stronger inverse associations observed at lower 25(OH)D concentrations. The shape of dose-response curves for coronary heart disease and stroke is not as clear, although larger inverse estimates were generally observed for strata with low 25(OH)D concentrations.

Discussion

In observational analyses, we found evidence for non-linear dose-response relationships of 25(OH)D concentrations with coronary heart disease, stroke, and mortality outcomes. Population-averaged estimates from our genetic analyses suggest that increases in 25(OH)D concentrations are unlikely to translate into substantial risk reductions for cardiovascular disease or all-cause mortality in the population overall. However, genetic analyses restricted to individuals with low concentrations of 25(OH)D provided strong evidence supporting a causal relationship between 25(OH)D concentrations and all-cause mortality risk up to a threshold of around 40 nmol/L. Our results have potential implications for aetiological understanding and disease prevention.

Our results challenge the interpretation of null findings from previous randomised trials and Mendelian randomisation analyses. Most previous trials for cardiovascular disease and mortality were done in broadly selected groups of the population, so could not reliably assess evidence for causality in individuals with low concentrations of 25(OH)D. Furthermore, previous Mendelian randomisation analyses have not considered estimates for strata of the population defined according to baseline 25(OH)D concentrations, and hence have not considered the shape of the causal relationship between 25(OH)D concentrations and cardiovascular disease or all-cause mortality.

By contrast, our investigation showed inverse associations between genetically-predicted 25(OH)D concentrations and mortality outcomes that were limited to the stratum of the population with low 25(OH)D concentrations. This finding suggests that any potential benefit of vitamin D supplementation could be restricted to those with low concentrations of 25(OH)D. This concept is

supported by a previous meta-analysis of vitamin D supplementation trials for acute respiratory tract infections, which reported stronger evidence of risk reduction in those with 25(OH)D concentrations below 25 nmol/L.[26] A general implication of our approach is to encourage the use of data and methods that allow identification of causal non-linear relationships that might otherwise go undetected.

Although our analyses suggest that a 10 nmol/L increase in genetically-predicted concentration of 25(OH)D is associated with about 30% lower risk of all-cause mortality in the deficient stratum (<25 nmol/L), previous Mendelian randomisation analyses for other risk factors have generally over-estimated the potential quantitative benefit compared with real-world interventions. For example, Mendelian randomisation estimates for low-density lipoprotein cholesterol on coronary heart disease risk are around 3-times larger than estimates from statin trials with a median follow-up of 5 years.[27] This finding might be because differences in risk factors investigated in Mendelian randomisation represent lifelong differences in usual levels of the risk factor. In contrast, randomised trials typically investigate short-term to medium-term interventions in a risk factor. For most diseases, it is probable that the benefit of short-term interventions in a risk factor will be less substantial than lifelong changes, suggesting that it could be challenging for future vitamin D supplementation trials to show reductions in mortality or morbidity, as the trials would need to be targeted at those with low concentrations of 25(OH)D. The trials would also need to have adequate power to detect small reductions in mortality or morbidity, suggesting that both large sample size and extended follow-up duration would be required.

There are several potential mechanisms by which vitamin D could be protective for cardiovascular mortality, including mechanisms linking low vitamin D status with hyperparathyroidism and low serum calcium and phosphate. Animal studies support a role for vitamin D in the regulation of cardiac metalloproteinases and fibroblasts, and consequently cardiac ventricular size and function.[28] Vitamin D is further implicated in endothelial cell function, in particular, modulating vascular tone, atherosclerosis, and arterial calcification.[29,30] There are also potential mechanisms implicating vitamin D for cancer. For example, vitamin D status affects the transcription of genes relating to cell division and apoptosis, including for neoplasms, with other effects including enhanced DNA repair and immuno-modulation.[31]

Our investigation has several strengths. The Mendelian randomisation design means that estimates are less susceptible to bias from confounding and reverse causation than those from conventional observational analyses. The availability of individual-level data on 25(OH)D concentrations, genotypes, and disease outcomes on large samples from several independent

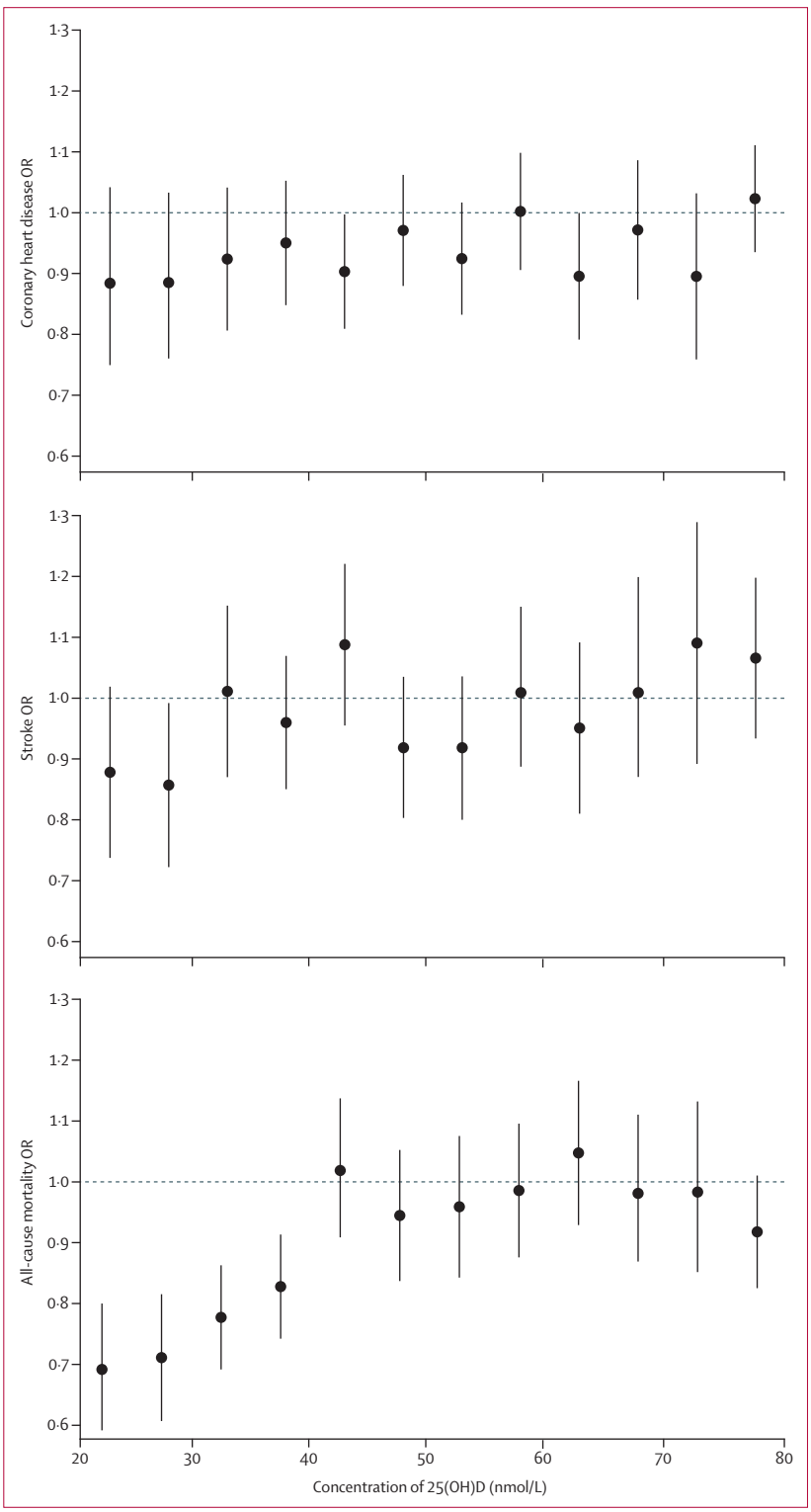

Figure 4: **Mendelian randomisation estimates for finer stratification of residual 25(OH)D concentrations at 5 nmol/L intervals**

Estimates (95% CIs) represent odds ratios per 10 nmol/L increase in genetically-predicted vitamin D concentration in strata of the population defined by residual concentration of 25(OH)D. Points are plotted on the horizontal axis at the central value of each stratum (22·5 nmol/L for the <25 nmol/L stratum, 27·5 nmol/L for the 25–30 nmol/L stratum, and so on). 25(OH)D=25-hydroxyvitamin D. OR=odds ratio.

datasets enabled powerful analysis of people with low 25(OH)D concentrations. Our focused genetic instrument for vitamin D afforded strong statistical power and biological specificity, minimising the potential for bias due to horizontal pleiotropy arising from use of variants that do not have specific effects on vitamin D pathways. The focused score was not associated with major cardiovascular risk factors, providing empirical evidence that genetic associations with all-cause mortality in the deficient stratum were attributable to 25(OH)D concentrations.

However, there are also potential limitations. Firstly, the Mendelian randomisation assumptions state that the only causal pathway from the genetic variants to the outcome is via 25(OH)D concentrations. Although our variants are all from gene regions specifically relevant to vitamin D biology, variants in the *CYP24A1* gene region are known to associate with circulating calcium levels. Secondly, even if the Mendelian randomisation assumptions are satisfied, genetic variants could influence 25(OH)D concentrations in a different way to dietary supplementation or other clinical interventions. Thirdly, to reduce the scope for confounding by ethnicity (population stratification), our analyses were limited to middle-aged participants of European ancestries. This limitation means that our findings might not be applicable to other populations. In particular, further analyses are needed to assess the potential effect of vitamin D supplementation in individuals with dark skin, as this correlates with lower 25(OH)D concentrations. Fourthly, UK Biobank and EPIC-CVD are not fully representative samples of the UK and European populations, further limiting the applicability of findings. Fifthly, fewer genetic variants were available in the Copenhagen studies, limiting comparability between datasets; however, although this will reduce power for analysis, it should not lead to bias. Sixthly, we do not have information from all studies on the accuracy of 25(OH)D measurements from external quality control programmes; however, there was no indication that mean 25(OH)D estimates varied by assay type, and as any such variation is probably non-differential to morbidity and mortality outcomes it would bias the results toward the null and therefore cannot explain our results. Finally, our primary genetic analyses for cardiovascular disease considered both prevalent and incident events. Stratification into categories according to residual 25(OH)D concentration might therefore be affected by reverse causation. However, genetic associations with disease outcomes within each of the strata will not be affected by reverse causation, as genotype is fixed from conception.

In conclusion, we found genetic evidence to suggest a causal relationship between 25(OH)D concentrations and mortality in individuals with low vitamin D status. Our results have implications for the interpretation and design of vitamin D supplementation trials, and potential disease prevention strategies.

For the **steering committee data from the Vitamin D Studies Collaboration** see https://www.phpc.cam.ac.uk/ceu/epic_cvd/

Contributors
ES, SKK, EA, MGA, AMM, RC, LAMP, LS, PW, JD, AMW, EDA, ASB, and SB conceived and designed the study. ES, SKK, SA, TJ, TRB, and SB did the analyses. ES, SKK, DG, AMW, EDA, ASB, JD, and SB drafted the manuscript. ES, SKK, SA, and SB verified the underlying data. SB is responsible for the decision to submit the manuscript. SB has seen and verified all the data. SKK has seen and verified the observational analysis. ES has seen and verified the genetic analyses of UK Biobank and EPIC-CVD. SA has seen and verified the genetic analyses of the Copenhagen studies. All authors acquired and interpreted the data, critically revised the paper, and had final responsibility for the decision to submit for publication.

Emerging Risk Factors Collaboration/EPIC-CVD/Vitamin D Studies Collaborators
Eleni Sofianopoulou*, Stephen K Kaptoge*, Shoaib Afzal, Tao Jiang, Dipender Gill, Thomas E Gundersen, Thomas R Bolton, Elias Allara, Matthew G Arnold, Amy M Mason, Ryan Chung, Lisa A M Pennells, Fanchao Shi, Luanluan Sun, Peter Willeit, Nita G Forouhi, Claudia Langenberg, Stephen J Sharp, Salvatore Panico, Gunnar Engström, Olle Melander, Tammy Y N Tong, Aurora Perez-Cornago, Margareta Norberg, Ingegerd Johansson, Verena Katzke, Bernard Srour, María José Sánchez, Daniel Redondo-Sánchez, Anja Olsen, Christina C Dahm, Kim Overvad, Magritt Brustad, Guri Skeie, Conchi Moreno-Iribas, N Charlotte Onland-Moret, Yvonne T van der Schouw, Konstantinos K Tsilidis, Alicia K Heath, Claudia Agnoli, Vittorio Krogh, Ian H de Boer, Camilla Jannie Kobylecki, Yunus Çolak, Armin Zittermann, Johan Sundström, Paul Welsh, Elisabete Weiderpass, Elom K Aglago, Pietro Ferrari, Robert Clarke, Marie-Christine Boutron, Gianluca Severi, Conor MacDonald, Rui Providencia, Giovanna Masala, Raul Zamora-Ros, Jolanda Boer, W M Monique Verschuren, Peggy Cawthon, Louise L Schierbeck, Cyrus Cooper, Matthias B Schulze, Manuela M Bergmann, Anke Hannemann, Stefan Kiechl, Hermann Brenner, Natasja M van Schoor, Juan R Albertorio, Carlotta Sacerdote, Allan Linneberg, Line L Kårhus, José María Huerta,Liher Imaz, Christel Joergensen, Yoav Ben-Shlomo, Annamari Lundqvist, John Gallacher, Naveed Sattar, Angela M Wood, Nicholas J Wareham, Børge G Nordestgaard, Emanuele Di Angelantonio, John Danesh, Adam S Butterworth, Stephen Burgess.
*Contributed equally.

Declaration of interests
ASB reports grants outside of this work from AstraZeneca, Biogen, BioMarin, Bioverativ, Merck, Novartis, Pfizer, and Sanofi and personal fees from Novartis. JD reports grants, personal fees, and non-financial support from Merck Sharp & Dohme, grants, personal fees, and non-financial support from Novartis, grants from Pfizer, and grants from AstraZeneca outside the submitted work. YÇ reports personal fees from Boehringer Ingelheim, AstraZeneca, and Sanofi Genzyme outside the submitted work. BGN reports consultancies and talks sponsored by AstraZeneca, Sanofi, Regeneron, Akcea, Amgen, Kowa, Denka, Amarin, Novartis, Novo Nordisk, Esperion, and Silence Therapeutics outside submitted work. MA is now an employee of AstraZeneca. DG is employed part-time by Novo Nordisk. PW reports grant income from Roche Diagnostics, AstraZeneca, Boehringer Ingelheim, and Novartis outside the submitted work. NS reports personal fees from Afimmune, Amgen, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Hanmi Pharmaceuticals, Merck Sharp and Dohme, Novartis, Novo Nordisk, Pfizer, and Sanofi, and grant funding paid to his institution from AstraZeneca, Boehringer Ingelheim, and Roche Diagnostics. All other authors declare no competing interests.

Data sharing
Data from UK Biobank is available to any bona fide scientific researcher on application. Applications to access data from EPIC-CVD should be addressed to the steering committee. Data from the Vitamin D Studies Collaboration is available at the discretion of the principal investigators of the individual studies.

Acknowledgments
We thank all EPIC participants and staff for their contribution to the study, the laboratory teams at the Medical Research Council (MRC)

Epidemiology Unit for sample management, Cambridge Genomic Services for genotyping, Matthew Walker and Sarah Spackman (British Heart Foundation [BHF] Cardiovascular Epidemiology Unit, Cambridge, UK) for data management, Nicola Kerrison (MRC Epidemiology Unit, Cambridge, UK) for managing the data for the InterAct Project, the laboratory team at VITAS Analytical Services, especially Anne Marte Haug, for vitamin D assays in EPIC-CVD, and the team at the EPIC-CVD Coordinating Centre for study coordination and administration. We acknowledge the use of data and biological samples from the EPIC-Ragusa cohort and the EPIC-Asturias cohort. A full list of Emerging Risk Factors Collaboration / Vitamin D Studies Collaboration investigators is provided in the appendix. This research has been done using the UK Biobank Resource under Application Number 11833.
No specific funding was obtained for this manuscript. EPIC-CVD was funded by the European Research Council (268834) and the European Commission Framework Programme 7 (HEALTH-F2–2012–279233). The coordinating centre was supported by core funding from the UK Medical Research Council (G0800270; MR/L003120/1), British Heart Foundation (SP/09/002; RG/13/13/30194; RG/18/13/33946), and the UK National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre (BRC-1215–20014). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. AMM is funded by the EC-Innovative Medicines Initiative (BigData@Heart). LS is funded by a BHF Programme Grant (RG/18/13/33946). TRB is funded by the NIHR Blood and Transplant Research Unit in Donor Health and Genomics (NIHR BTRU-2014–10024) and a BHF Programme Grant (RG/18/13/33946). SB is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (award number 204623/Z/16/Z), and by the UK Research and Innovation Medical Research Council (MC_UU_00002/7). RC is funded by a BHF PhD studentship (FS/18/56/34177). EA was funded by the EC-Innovative Medicines Initiative (BigData@Heart) and a BHF Programme Grant (RG/18/13/33946). JD holds a BHF Professorship and a NIHR Senior Investigator Award. TJ is funded by the NIHR Cambridge Biomedical Research Centre (BRC-1215–20014). AMW is supported by a BHF-Turing Cardiovascular Data Science Award and by the EC-Innovative Medicines Initiative (BigData@Heart). LAMP is funded by a BHF Programme Grant (RG/18/13/33946). TYNT is supported by a Nuffield Department of Population Health Intermediate Fellowship. MA is funded by a BHF Programme Grant (RG/18/13/33946). NJW, CL, and NGF acknowledge funding from the MRC Epidemiology Unit MC_UU_12015/1 and MC_UU_12015/5. FS is funded by the Chinese Scholarship Council, and Cambridge Commonwealth European and International Trust. NJW and NGF received support from the NIHR Biomedical Research Centre Cambridge: Nutrition, Diet, and Lifestyle Research Theme (IS-BRC-1215–20014). DG is supported by the BHF Centre of Research Excellence (RE/18/4/34215) at Imperial College London and a National Institute for Health Research Clinical Lectureship (CL-2020–16–001) at St George's, University of London. CM is supported by the Federation Française de Cardiologie,. RZ-R was supported by the Miguel Servet type II programme (CPII20/00009) from the Institute of Health Carlos III (co-funded by the European Social Fund (ESF)—ESF investing in your future). YBS is partially funded by the NIHR School of Population Health Research. The InterAct project was funded by the EU FP6 programme (grant number LSHM_CT_2006_037197). Measurements for 25(OH)D were funded jointly by the InterAct project, the EPIC-CVD project and the MRC Cambridge Initiative (RG71466, SJAH/004). EPIC-Oxford is supported by Cancer Research UK (C8221/A29017) and UK MRC (MR/M012190/1). The Longitudinal Aging Study Amsterdam (LASA) is supported by a grant from the Netherlands Ministry of Health, Welfare and Sport, Directorate of Long-Term Care. The Multi-Ethnic Study of Atherosclerosis is supported by grant R01HL096875 and contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159,75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, and N01-HC-95169 from the National Heart, Lung, and Blood Institute, and by grants UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420 from the National Center for Advancing Translational Sciences (NCATS). The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health (NIH) funding. The following institutes provide support: the National Institute on Aging (NIA), the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NCATS, and NIH Roadmap for Medical Research under the following grant numbers: U01 AG027810, U01 AG042124, U01 AG042139, U01 AG042140, U01 AG042143, U01 AG042145, U01 AG042168, U01 AR066160, and UL1 TR000128. MrOS data is available online. The Study of Osteoporotic Fractures is supported by NIH funding. NIA provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576. The Study of Health in Pomerania-1 is part of the Community Medicine Research network of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs, and the Social Ministry of the Federal State of Mecklenburg–West Pomerania. The coordination of EPIC is financially supported by International Agency for Research on Cancer and also by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, which has additional infrastructure support provided by the NIHR Imperial Biomedical Research Centre. The national cohorts are supported by: Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, and Institut National de la Santé e de la Recherche Médicale (France), German Cancer Aid, German Cancer Research Center, German Institute of Human Nutrition Potsdam-Rehbruecke, Federal Ministry of Education and Research (Germany); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy, Compagnia di SanPaolo, and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports, Netherlands Cancer Registry, LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund, Statistics Netherlands (the Netherlands); Austrian Research Promotion Agency (Austria); Health Research Fund - Instituto de Salud Carlos III, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology (Spain); Swedish Cancer Society, Swedish Research Council, and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C8221/A29017 to EPIC-Oxford), MRC (1000143 to EPIC-Norfolk; MR/M012190/1 to EPIC-Oxford) (UK).

For **MrOS data** see http://mrosdata.sfcc-cpmc.net

References

1 Cashman KD, Dowling KG, Škrabáková Z, et al. Vitamin D deficiency in Europe: pandemic? *Am J Clin Nutr* 2016; **103:** 1033–44.
2 Wang TJ, Pencina MJ, Booth SL, et al. Vitamin D deficiency and risk of cardiovascular disease. *Circulation* 2008; **117:** 503–11.
3 Heath AK, Kim IY, Hodge AM, English DR, Muller DC. Vitamin D status and mortality: a systematic review of observational studies. *Int J Environ Res Public Health* 2019; **16:** 383.
4 Manson JE, Cook NR, Lee I-M, et al. Vitamin D supplements and prevention of cancer and cardiovascular disease. *N Engl J Med* 2019; **380:** 33–44.
5 Zittermann A, Ernst JB, Prokop S, et al. Effect of vitamin D on all-cause mortality in heart failure (EVITA): a 3-year randomized clinical trial with 4000 IU vitamin D daily. *Eur Heart J* 2017; **38:** 2279–86.
6 Scragg R, Stewart AW, Waayer D, et al. Effect of monthly high-dose vitamin D supplementation on cardiovascular disease in the vitamin D assessment study: a randomized clinical trial. *JAMA Cardiol* 2017; **2:** 608–16.
7 Zhang Y, Fang F, Tang J, et al. Association between vitamin D supplementation and mortality: systematic review and meta-analysis. *BMJ* 2019; **366:** l4673.
8 Thanassoulis G, O'Donnell CJ. Mendelian randomization: nature's randomized trial in the post-genome era. *JAMA* 2009; **301:** 2386–88.
9 Brøndum-Jacobsen P, Benn M, Afzal S, Nordestgaard BG. No evidence that genetically reduced 25-hydroxyvitamin D is associated with increased risk of ischaemic heart disease or myocardial infarction: a Mendelian randomization study. *Int J Epidemiol* 2015; **44:** 651–61.
10 Manousaki D, Mokry LE, Ross S, Goltzman D, Richards JB. Mendelian randomization studies do not support a role for vitamin D in coronary artery disease. *Circ Cardiovasc Genet* 2016; **9:** 349–56.

11 Larsson SC, Traylor M, Mishra A, Howson JMM, Michaëlsson K, Markus HS. Serum 25-hydroxyvitamin D concentrations and ischemic stroke and its subtypes. *Stroke* 2018; **49:** 2508–11.

12 Huang T, Afzal S, Yu C, et al. Vitamin D and cause-specific vascular disease and mortality: a Mendelian randomisation study involving 99,012 Chinese and 106,911 European adults. *BMC Med* 2019; **17:** 160.

13 Afzal S, Brøndum-Jacobsen P, Bojesen SE, Nordestgaard BG. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. *BMJ* 2014; **349:** g6330.

14 Ye Y, Yang H, Wang Y, Zhao H. A comprehensive genetic and epidemiological association analysis of vitamin D with common diseases/traits in the UK Biobank. *Genet Epidemiol* 2021; **45:** 24–35.

15 Zhang R, Li B, Gao X, et al. Serum 25-hydroxyvitamin D and the risk of cardiovascular disease: dose-response meta-analysis of prospective studies. *Am J Clin Nutr* 2017; **105:** 810–19.

16 Crowe FL, Thayakaran R, Gittoes N, et al. Non-linear associations of 25-hydroxyvitamin D concentrations with risk of cardiovascular disease and all-cause mortality: results from The Health Improvement Network (THIN) database. *J Steroid Biochem Mol Biol* 2019; **195:** 105480.

17 Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015; **12:** e1001779.

18 Astle WJ, Elding H, Jiang T, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* 2016; **167:** 1415–29.

19 Danesh J, Saracci R, Berglund G, et al. EPIC-Heart: the cardiovascular component of a prospective study of nutritional, lifestyle and biological factors in 520,000 middle-aged participants from 10 European countries. *Eur J Epidemiol* 2007; **22:** 129–41.

20 Langenberg C, Sharp S, Forouhi NG, et al. Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. *Diabetologia* 2011; **54:** 2272–82.

21 Appleyard M. The Copenhagen City Heart Study. Osterbroundersøgelsen. A book of tables with data from the first examination (1976–78) and a five year follow-up (1981–83). *Scand J Soc Med Suppl* 1989; **41:** 1–160.

22 Nordestgaard BG, Palmer TM, Benn M, et al. The effect of elevated body mass index on ischemic heart disease risk: causal estimates from a Mendelian randomisation approach. *PLoS Med* 2012; **9:** e1001212.

23 NICE. Sunlight exposure: risks and benefits. 2016. https://www.nice.org.uk/guidance/ng34/resources/sunlight-exposure-risks-and-benefits-pdf-1837392363205 (accessed Sept 30, 2021).

24 Wang TJ, Zhang F, Richards JB, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet* 2010; **376:** 180–88.

25 Manousaki D, Mitchell R, Dudding T, et al. Genome-wide association study for vitamin D levels reveals 69 independent loci. *Am J Hum Genet* 2020; **106:** 327–37.

26 Martineau AR, Jolliffe DA, Hooper RL, et al. Vitamin D supplementation to prevent acute respiratory tract infections: systematic review and meta-analysis of individual participant data. *BMJ* 2017; **356:** i6583.

27 Ference BA. How to use Mendelian randomization to anticipate the results of randomized trials. *Eur Heart J* 2018; **39:** 360–62.

28 Weishaar RE, Kim SN, Saunders DE, Simpson RU. Involvement of vitamin D3 with cardiovascular function. III. Effects on physical and morphological properties. *Am J Physiol* 1990; **258:** E134–42.

29 Merke J, Milde P, Lewicka S, et al. Identification and regulation of 1,25-dihydroxyvitamin D3 receptor activity and biosynthesis of 1,25-dihydroxyvitamin D3. Studies in cultured bovine aortic endothelial cells and human dermal capillaries. *J Clin Invest* 1989; **83:** 1903–15.

30 Danielsson C, Nayeri S, Wiesinger H, Thieroff-Ekerdt R, Carlberg C. Potent gene regulatory and antiproliferative activities of 20-methyl analogues of 1,25 dihydroxyvitamin D3. *J Cell Biochem* 1996; **63:** 199–206.

31 Fleet JC, DeSmet M, Johnson R, Li Y. Vitamin D and cancer: a review of molecular mechanisms. *Biochem J* 2012; **441:** 61–76.