**Vitamin D, coronary heart disease, stroke and all-cause mortality: estimating the dose—response curve in observational and genetic analyses**

**Abstract count**: 336

**Text count**: 3268

**Tables**: 3

**Figures:** 2

**ABSTRACT**

**Background:**

Although randomized trials of vitamin D supplementation have generally reported null findings for a range of outcomes, interest in vitamin D supplementation persists. While Mendelian randomization analyses for vitamin D have generally provided null estimates, results may not be generalizable to all subgroups of the population, and in particular to those with low baseline levels of vitamin D. The aim of this investigation was to characterize the dose—response relationship between 25(OH)D concentrations and risk of coronary heart disease (CHD), stroke, and all-cause mortality in conventional and Mendelian randomization frameworks.

**Methods and findings:**

Observational analyses were conducted in long-term, mostly population-based, prospective studies comprising 501,969 individuals with no known history of CHD or stroke at baseline. Genetic analyses were conducted in three large population-based cohorts (UK Biobank, EPIC-CVD, and Copenhagen studies) comprising 386,406 middle-aged individuals of European ancestries, including 33,546 CHD cases, 18,166 stroke cases, and 27,885 all-cause mortality cases. Observational analyses suggested a threshold relationship for incident CHD, stroke, and mortality outcomes, with a negative association for low values of 25(OH)D, and a flatter association for higher values of 25(OH)D. In genetic analyses, there was no association between genetically-predicted 25(OH)D and CHD (odds ratio [OR] per 10 nmol/L higher genetically-predicted 25(OH)D 0.98, 95% confidence interval [CI] 0.95-1.01), stroke (OR 1.01, 95% CI 0.97-1.05), or all-cause mortality (OR 0.99, 95% CI 0.95-1.02) in overall analyses. In the deficient stratum (25(OH)D concentration below 25 nmol/L), there was strong evidence for a protective association with all-cause mortality (OR 0.69, 95% CI 0.59-0.80, p=1×10-6), and weak evidence of a protective association for stroke (OR 0.85, 95% CI 0.70-1.02, p=0.09). Similar associations in the deficient stratum were observed in cause-specific mortality analyses for cardiovascular mortality (OR 0.69, 95% CI 0.52-0.92), cancer mortality (OR 0.81, 95% CI 0.65-1.02), and non-cardiovascular non-cancer mortality (OR 0.68, 95% CI 0.54-0.85).

**Conclusions:**

Genetic evidence suggests that vitamin D supplementation may reduce the risk of all-cause mortality in individuals with low 25(OH)D concentrations. Trials for vitamin D supplementation should focus on vitamin D deficient individuals.

**INTRODUCTION**

Observational studies have found low levels of vitamin D, as measured by concentration of 25‑hydroxyvitamin D (25(OH)D) in blood plasma or serum, are associated with increased risk of cardiovascular disease (CVD) and all-cause mortality [1-6]. This finding is not generally supported by randomized trials of vitamin D supplementation [7-11], although a recent meta-analysis of trials showed limited benefit for all-cause mortality particularly in elderly women [12]. A recent large-scale randomized intervention trial, the Vitamin D and Omega-3 Trial (VITAL), found that daily supplementation with vitamin D3 did not result in lower mortality or morbidity [7]. However, this null result may not necessarily be generalizable to all subgroups of the population.

An alternative approach for assessing the potential causal effect of vitamin D supplementation is Mendelian randomization [13]. Mendelian randomization uses genetic variants specifically related to a particular exposure to compare genetically-defined subgroups of the population 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 randomized trial [14]. Mendelian randomization analyses have generally shown null associations of genetically-predicted 25(OH)D concentrations with coronary heart disease (CHD) [15, 16] and ischaemic stroke [17] in European populations, and similarly in a Chinese population [18]. In contrast, an association corresponding to a protective effect of vitamin D supplementation has been observed between genetically-predicted 25(OH)D and all-cause mortality, although the association with cardiovascular mortality was null [19, 20].

Previous Mendelian randomization analyses have been undertaken in population-based studies assuming a linear dose**—**response relationship between genetically-predicted 25(OH)D and cardiovascular disease. However, some observational analyses have reported a non-linear dose—response relationship [21, 22]. Here, we conduct the largest observational analysis to date to characterize the shape of the association between 25(OH)D concentrations and cardiovascular disease outcomes in an individual-participant data meta-analysis of prospective studies. We then perform non-linear Mendelian randomization analyses in large European-ancestry cohorts to assess whether there is evidence for a causal effect of vitamin D supplementation on cardiovascular disease risk and all-cause mortality for groups in the population with different 25(OH)D concentrations. Analyses are performed for three main outcomes: CHD, stroke, and all-cause mortality. We also consider cause-specific mortality, divided into cardiovascular, cancer, and non-cardiovascular non-cancer mortality outcomes.

**METHODS**

**Study populations**

Observational analyses were conducted in UK Biobank, European Prospective Investigation into Cancer-CVD (EPIC-CVD), and the Vitamin D Studies Collaboration (VitDSC). Genetic analyses were conducted in UK Biobank, EPIC-CVD, and two Copenhagen studies.

UK Biobank is a large-scale prospective cohort study of around half a million people aged 40 to 69 years at baseline, recruited in 2006-2010 from across the United Kingdom and followed-up for a median of 10.9 years [23]. For genetic analyses, we consider data on 333,025 unrelated individuals of European ancestries who passed various quality control filters as described previously [24] and had a valid 25(OH)D measurement. EPIC-CVD is a case-cohort study of 22,142 European-descent individuals derived from a larger cohort (European Prospective Investigation into Cancer, EPIC) of over 500,000 individuals from 23 centres across 10 European countries [25]. Participants were recruited in the 1990s and were followed-up for a median of 9.5 years. VitDSC comprises 32 long-term, mostly population-based, prospective studies with 108,301 individuals in total. Individual-participant data on 25(OH)D concentration, age, sex, conventional cardiovascular risk factors, and major incident cardiovascular morbidity and mortality were available for 67,992 individuals from 25 studies without previous known cardiovascular disease. The Copenhagen City Heart Study (CCHS) and Copenhagen General Population Study (CGPS) are prospective cohort studies in the Danish population [26, 27]. CCHS was initiated in 1977 and participants were followed up periodically until 2018. CGPS was initiated in 2003 and has a median follow-up of 8.8 years. For genetic analyses, we consider 31,262 individuals from both studies with genetic data and a valid 25(OH)D measurement. For all studies, written informed consent was obtained from participants and approval was obtained from relevant ethical committees.

**Vitamin D measurement and classification**

Concentrations of 25(OH)D in blood were measured using the DiaSorin Liaison immunoassay analyser (UK Biobank and Copenhagen studies) or high-performance liquid chromatography (EPIC-CVD). In VitDSC, concentrations were measured by radioimmunoassay, direct chromatographic approaches, or other immunoassays (**Supplementary Table 1**). In genetic analyses, measurements were seasonally adjusted by uniformly shifting all values based on the mean 25(OH)D measurement in that study (and centre for EPIC-CVD) to correspond to a measurement taken in autumn. Stratification of participants into clinical categories was performed using season-shifted measurements based on guidelines set by the National Academy of Medicine in the United States and by the Endocrine Society [28, 29]: adequate:  > 75 nmol/L; sufficient: 50–75 nmol/L; insufficient: 25–50 nmol/L; and deficient: <25 nmol/L.

**Outcomes**

The main outcome measures were CHD, defined as fatal ischaemic heart disease (ICD-10 code: I20-I25) or nonfatal myocardial infarction (I21-I23); stroke, defined as any cerebrovascular disease (I60-I69); and all-cause mortality. We also performed secondary analyses for cause-specific mortality divided based on ICD-10 codes into cardiovascular mortality, cancer mortality, or non-cardiovascular non-cancer mortality. Secondary genetic analyses were performed restricting to incident cardiovascular disease, and subtype analyses for ischaemic and haemorrhagic stroke.

**Genetic variants**

We considered genetic variants from four gene regions that have previously been shown to be associated with 25(OH)D [30] and are implicated in the transport, metabolism, and synthesis of vitamin D [31]: *GC*, *DHCR7*, *CYP2R1*, and *CYP24A1*. The *GC* gene encodes vitamin D binding protein. The *DHCR7* gene product converts 7-dehydrocholesterol to cholesterol, providing a substrate for vitamin D production. CYP2R1 is a regulator of 25(OH)D synthesis through 25-hydroxylation of vitamin D in the liver. CYP24A1 renders the active form of vitamin D (1α25(OH)2D) to be inactive. To maximize the proportion of variance explained by the genetic variants, we considered all variants in the relevant genetic locus (including a 500 kilobasepair flanking window) and selected variants in these gene regions from UK Biobank using a stepwise selection method based on their conditional associations with 25(OH)D, including at each step the variant having the lowest conditional p-value. For UK Biobank and EPIC-CVD, 21 variants were included in the analysis (**Supplementary Table 2**). For the Copenhagen studies, due to limited availability, 3 variants were used based on their conditional associations with 25(OH)D: two from the *CYP2R1* locus (rs12794714 and rs117913124) and one from the *DHCR7* locus (rs7944926).

**Statistical methods**

*Observational analysis*

Observational associations of 25(OH)D and outcomes were assessed by meta-analysis of study-specific hazard ratios (HRs) calculated using Cox proportional hazards regression models stratified by centre, sex and (where appropriate) trial arm. Principal analyses were adjusted for conventional CVD risk factors, namely age, month of 25(OH)D measurement, smoking status (current versus other), total cholesterol, HDL cholesterol, systolic blood pressure, known history of diabetes, and body mass index. To account for the case-cohort study design in EPIC-CVD, Cox models were adapted using Prentice weights and stratified by centre [32]. To avoid overfitting models, studies contributing fewer than 10 incident events to the analysis of a particular outcome were excluded from that analysis.

The primary dose**—**response analyses assessed the continuous shape of association of 25(OH)D and outcomes by meta-analysis of fractional polynomials adjusted for the conventional risk factors [33]. First, the best-fitting fractional polynomial of degree 2 was estimated for each outcome using Cox regression model fitted to the combined dataset stratified by cohort, centre, sex, and trial arm. Next, the coefficients for the best fitting fractional polynomial powers were estimated separately within each study, and then pooled across studies by random effects meta-analysis [34]. The pooled coefficients were used to plot the continuous shape of association relative to the reference value of 50 nmol/L. Further supplementary analyses combined study-specific hazard ratios by deciles of 25(OH)D and plotted the pooled hazard ratios against the pooled mean 25(OH)D by deciles.

*Genetic analysis*

We constructed a genetic risk score (GRS) weighted by the conditional associations of the genetic variants with 25(OH)D concentration in UK Biobank (**Supplementary Table 2**). Mendelian randomization 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 change in genetically-predicted 25(OH)D concentration. Genetic associations with the outcome were estimated using logistic regression, and with 25(OH)D concentration using linear regression. All regression models were adjusted for age, sex, centre (for EPIC-CVD), and 10 principal components.

In addition to analyses conducted in the overall study sample to estimate population-averaged causal effects, we also conducted analyses in strata of the population according to their residual 25(OH)D. Residual 25(OH)D is defined as an individual’s 25(OH)D concentration minus the centred genetic contribution to 25(OH)D from variants included in the GRS. By stratifying on residual 25(OH)D, we compare individuals in the population who would have similar 25(OH)D concentrations (that is, values in the same stratum) if they had the same values of the genetic variants [35]. Stratifying on the exposure directly would distort estimates, as it is on the causal pathway from the genetic variants to the outcome. For each study, we calculated ratio estimates within clinical strata of adequate, sufficient, insufficient, and deficient as defined above and combined them across the three studies using fixed-effect meta-analysis. We also divided each study population into 20 equal strata and calculated ratio estimates within each stratum.

All statistical analyses were performed in R (version 3.4.3) or Stata/SE 15.1 (College Station, TX, USA).

**RESULTS**

**Participant characteristics**

**Table 1** shows the baseline characteristics of participants in the genetic analyses (details of participants in the observational analyses are in **Supplementary Tables 3 and 4**). Mean age of participants was around 55 to 58 years in all studies. Mean season-shifted 25(OH)D concentrations were around 55 nmol/L in the UK Biobank and Copenhagen studies, and around 47 nmol/L in EPIC-CVD. Around 4% of participants in the UK Biobank and Copenhagen studies were in the deficient category, compared with around 7% in EPIC-CVD. In the combined dataset for the observational analysis, 13% of participants were in the deficient category. The GRS explained 4.7% of the variance in 25(OH)D concentrations in UK Biobank, 5.8% in EPIC-CVD, and 1.8% in the Copenhagen studies.

**Observational analysis**

Overall, the shape of the observational association was similar for all outcomes (**Figure 1**): for low values of 25(OH)D, there was an inverse association, and for higher values of 25(OH)D, the association was generally flat. For CHD and stroke, there was no strong association for 25(OH)D concentrations above 50 nmol/L, and but a progressively steeper association was observed below this threshold. For all-cause mortality, the strength of association with lower 25(OH)D concentrations was stronger, and began at a higher threshold 25(OH)D concentration. There was also a slight positive association between higher concentrations of 25(OH)D and all-cause mortality, as well as for cancer and non-cardiovascular non-cancer mortality.

While the shape of the observational association varied between studies, potentially reflecting between-study heterogeneity or chance variation due to small sample sizes for individual studies, the pooled meta-analytic estimates according to the 3 primary data sources (i.e. VitDSC, UK Biobank, and EPIC-CVD) were broadly similar (**Supplementary Figure 3**). Dose-response findings were also similar in supplementary analyses that combined study-specific hazard ratios by deciles of 25(OH)D or according to four clinical categories of 25(OH)D (**Supplementary Figure 4**).

**Genetic analysis**

Mendelian randomization estimates, representing the odds ratio per 10 nmol/L higher genetically-predicted 25(OH)D concentration, are given in **Table 2**. In overall analyses (that is, for population-averaged estimates averaged across all individuals) combined across studies, there was no association between genetically-predicted 25(OH)D and CHD (odds ratio [OR] 0.98, 95% confidence interval [CI] 0.95-1.01), stroke (OR 1.01, 95% CI 0.97-1.05), or all-cause mortality (OR 0.99, 95% CI 0.95-1.02). However, there was some evidence of an overall protective association with all-cause mortality in the Copenhagen study. In combined analyses for clinically-defined strata, there was strong evidence for a protective association with all-cause mortality in the deficient stratum (OR 0.69, 95% CI 0.59-0.80, p=1×10-6), and weak evidence for protective associations in the deficient stratum with CHD (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). In contrast, estimates for other strata were much closer to the null, although there was evidence of a protective association for all-cause mortality in the insufficient stratum (OR 0.94, 95% CI 0.89-0.99, p=0.013) and for CHD in the adequate stratum (OR 0.94, 95% CI 0.89-0.98, p=0.006). Similar results were observed for supplementary analyses in UK Biobank that considered incident stroke outcomes only (**Supplementary Table 5**) and incident CHD outcomes only (**Supplementary Table 6**).

Estimates for all-cause and cause-specific mortality in UK Biobank and Copenhagen studies are presented in **Table 3**. Protective 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 non-cancer mortality (OR 0.68, 95% CI 0.54-0.85, p=0.0009). A suggestive protective association was observed for cancer mortality (OR 0.81, 95% CI 0.65-1.02, p=0.071). Again, overall and stratum-specific estimates in other strata were generally null.

Mendelian randomization estimates for a finer stratification of the UK Biobank population into 20 groups are displayed in **Figure 2**, and for EPIC-CVD and the Copenhagen studies in **Supplementary Figure 1** and **Supplementary Figure 2**. For all-cause mortality, the shape of the dose—response curve appears to have a threshold shape with evidence of a protective association at 25(OH)D values below around 40 nmol/L, and null associations at 25(OH)D values above 40 nmol/L. This suggests that substantial benefit of vitamin D supplementation is restricted to those in the population with 25(OH)D concentrations of 40 nmol/L and below, and any potential benefit for those with higher 25(OH)D concentrations will be lesser in magnitude.

**DISCUSSION**

Overall, our investigation found little consistent evidence to suggest that population-wide vitamin D supplementation will substantially reduce cardiovascular risk or all-cause mortality. However, in line with observational analyses, it did suggest that vitamin D supplementation may reduce risk of all-cause mortality for those with low concentrations of 25(OH)D. Additionally, estimates for the predicted effect of vitamin D supplementation in the deficient stratum (<25 nmol/L) were in the protective direction for cardiovascular, cancer, and non-cardiovascular non-cancer mortality. While estimates for other strata (>25 nmol/L) were generally in the protective direction, they were much closer to the null and generally not significant. While it is difficult to estimate a threshold below which vitamin D supplementation will be beneficial, protective estimates in the genetic analyses were generally observed up to a 25(OH)D concentration of 35-40 nmol/L. In UK Biobank, after correcting for season of blood draw to convert all values to an autumn measurement, this corresponds to 16-25% of the population.

Overall, population-averaged causal estimates from our analyses suggested no benefit of vitamin D supplementation for cardiovascular disease or all-cause mortality, in line with previous large randomized controlled trials and Mendelian randomization analyses. A recent meta-analysis of randomized trials indicated no benefit of vitamin D supplementation for all-cause mortality or cardiovascular mortality, although benefit was observed for cancer mortality [36]. Secondary analyses of randomized trials have suggested reductions in disease outcomes for cancer [7] and diabetes [37] in specific subgroups, as well as improvements in survival after acute stroke in open-label trials [38, 39]. However, large trials have generally not been restricted to vitamin D deficient individuals [36], and previous Mendelian randomization analyses have not considered estimates for strata of the population defined according to 25(OH)D concentrations, and hence have not considered the non-linear shape of the causal relationship between vitamin D and cardiovascular disease or all-cause mortality.

While our analyses suggest that 10 nmol/L higher genetically-predicted concentrations of 25(OH)D is associated with a 31% reduction in all-cause mortality in the deficient stratum (<25 nmol/L), previous Mendelian randomization analyses for other risk factors have generally over-estimated the benefit of real-world interventions [40]. For example, Mendelian randomization estimates for LDL-cholesterol on CHD risk are around 3 times larger than estimates from statin trials with a median follow-up of 5 years [41]. This is because differences in risk factors investigated in Mendelian randomization represent lifelong changes in usual levels of the risk factor. In contrast, randomized trials typically investigate short- to medium-term interventions in a risk factor. For most diseases, it is likely that the benefit of short-term interventions in a risk factor will be less substantial than lifelong changes. Additionally, as our analyses suggested a threshold relationship, it is unlikely that high doses of vitamin D will improve outcomes in a dose—dependent way. This suggests that for a trial to demonstrate the benefit of vitamin D supplementation, it would need to be long in duration, targeted at those with low concentrations of 25(OH)D, and have adequate power to detect a small reduction in mortality or morbidity. The sample size required for a trial may be prohibitively large to be performed.

There are several potential mechanisms by which vitamin D could be protective for cardiovascular mortality. The vitamin D receptor is found in cardiovascular tissues, with animal studies supporting a role vitamin D in the regulation of cardiac metalloproteinases and fibroblasts, and consequently cardiac ventricular size and function [42, 43]. Vitamin D is further implicated in endothelial cell function, and in particular, modulating vascular tone, atherosclerosis and arterial calcification [44-47]. There may also be a role for vitamin D in the regulation of fluid balance via effects on the renin-angiotensin-aldosterone axis [48]. 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 immunomodulation [49].

Our investigation has several strengths, but also limitations. The Mendelian randomization design means that estimates are less susceptible to bias from confounding and reverse causation than those from observational analyses. The availability of individual-level data on large samples enabled detailed analyses in independent datasets. However, there are also limitations. The Mendelian randomization assumptions state that the genetic variants in the investigation act as proxies for vitamin D supplementation. This requires that the only causal pathway from the genetic variants to the outcome is via 25(OH)D concentrations. While the variants that we included in this analysis are all from gene regions that are specifically relevant to vitamin D biology, variants in the *CYP24A1* gene region are known to associate with serum calcium levels. Secondly, even if the Mendelian randomization assumptions are satisfied, it may be that genetic variants influence 25(OH)D concentrations in a different way to dietary supplementation or other clinical interventions to raise 25(OH)D concentrations. Alternatively, it may be that supplementation is only effective at a critical point in the life-course. For example, there is some evidence that multiple sclerosis risk is particularly dependent on 25(OH)D concentrations in early childhood [50]. This means that the protective association between increased genetically-predicted 25(OH)D and multiple sclerosis risk observed in Mendelian randomization [31] may only translate into a clinically beneficial effect of vitamin D supplementation if the intervention is implemented in young children. Thirdly, our analyses were limited to middle-aged participants of European descent. While restricting to a well-defined population is necessary to ensure that genetic associations are not driven by population stratification, this means that our findings may 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 low levels of 25(OH)D. Additionally, UK Biobank is not a fully representative sample of the UK population, further limiting the applicability of findings [51]. Finally, our primary genetic analyses for cardiovascular diseases considered both prevalent and incident events. Stratification into categories according to residual 25(OH)D concentration may 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 the genetic code is fixed from conception.

In conclusion, we found evidence to suggest that vitamin D supplementation may reduce the risk of all-cause and cause-specific mortality in individuals with low concentrations of 25(OH)D. Trials for vitamin D supplementation should focus on vitamin D deficient individuals.

Table 1: Baseline characteristics of participants in the genetic analyses by study

|  |  |  |  |
| --- | --- | --- | --- |
|  | **UK Biobank** | **EPIC-CVD** | **Copenhagen** |
| **Participants** | 333,002 | 22,142 | 31,262 |
| **Age** | 57.1 ± 8.1 | 54.8 ± 9.4 | 57.5 ± 12.9 |
| **Sex** |  |  |  |
| Female | 177,733 (53.4) | 11,426 (51.6) | 17311 (55.4) |
| Male | 155,269 (46.6) | 10,716 (48.4) | 13951 (44.6) |
| **25(OH)D concentration (nmol/L)** | 49.4 ± 20.9 | 46.9 ± 16.4 | 53.8 ± 25.9 |
| **Residual vitamin D classes** |  |  |  |
| <25 nmol/L, Deficient | 12,957 (3.9) | 1522 (6.9) | 1158 (3.7) |
| 25-49 nmol/L, Insufficient | 133,922 (40.2) | 11,845 (53.5) | 9730 (31.1) |
| 50-74 nmol/L, Sufficient | 138,605 (41.6) | 7717 (34.9) | 12,375 (39.6) |
| >75 nmol/L, Adequate | 47,518 (14.3) | 1058 (4.8) | 7999 (25.6) |
| **CHD events** | 22,363 (6.7) | 5942 (26.8) | 5241 (16.8) |
| **Stroke events** | 10,489 (3.1) | 5478 (24.7) | 2199 (7.0) |
| **All-cause mortality events** | 20,340 (6.1) | - | 7545 (24.1) |
| **BMI (kg/m2)** | 27.3 ± 4.8 | 26.7 ± 4.3 | 25.9 ± 4.2 |
| **SBP (mmHg)** | 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** |  |  |  |
| History | 15,822 (4.8) | 1234 (5.6) | 1244 (4.0) |
| No history | 317,203 (95.2) | 20,908 (94.4) | 30,018 (96.0) |

Values represent mean ± standard deviation for continuous traits or N (%) for categorical traits.

Abbreviations: BMI = body mass index, CHD = coronary heart disease, SBP = systolic blood pressure.

Figure 1. Dose**—**response curves from meta-analysis of association of 25(OH)D concentration with outcomes adjusted for conventional cardiovascular risk factors.



Reference value is 50 nmol/L. The shaded area represents the 95% confidence interval for the dose—response curve. Study-specific analyses involved fractional polynomial modelling of continuous associations of 25(OH)D and outcomes using Cox regression stratified by sex and adjusted for conventional risk factors, namely: age, month, current smoking, total cholesterol, HDL cholesterol, systolic blood pressure, history of diabetes, and body mass index, followed by random effects meta-analysis (see **Methods**).

Table 2: Mendelian randomization estimates for overall population and divided into clinical strata according to residual concentration of 25(OH)D.

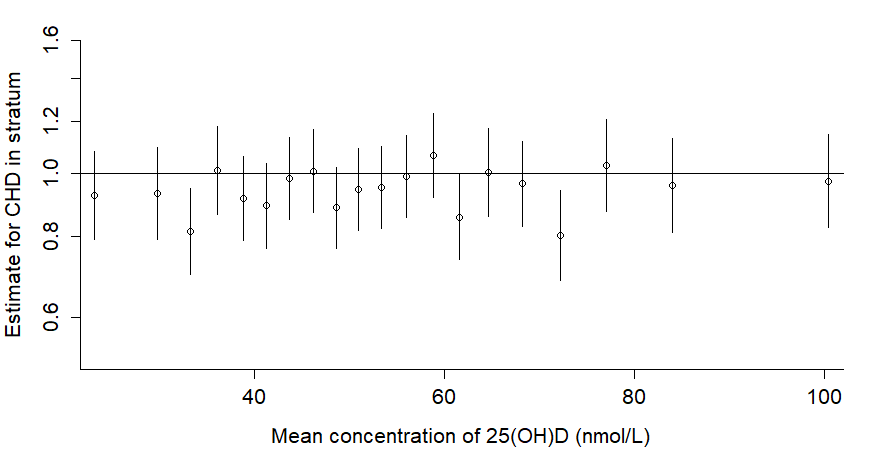
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Study and outcome** | **Overall** | **Deficient (<25 nmol/L)** | **Insufficient (25-49 nmol/L)** | **Adequate (50-74 nmol/L)** | **Sufficient (>=75 nmol/L)** |
| *Combined* |  |  |  |  |  |
| - Coronary heart disease | 0.98 (0.95-1.01) p=0.18 | 0.89 (0.76-1.04) p=0.14 | 0.96 (0.92-1.01) p=0.11 | 0.94 (0.89-0.98) p=0.006 | 1.02 (0.94-1.11) p=0.63 |
| - Stroke | 1.01 (0.97-1.05) p=0.61 | 0.85 (0.70-1.02) p=0.09 | 0.99 (0.94-1.06) p=0.84 | 0.97 (0.91-1.04) p=0.40 | 1.05 (0.93-1.19) p=0.42 |
| - All cause mortality | 0.99 (0.95-1.02) p=0.39 | 0.69 (0.59-0.80) p=1×10-6 | 0.94 (0.89-0.99) p=0.013 | 0.98 (0.93-1.03) p=0.39 | 0.92 (0.83-1.01) p=0.067 |
| *UK Biobank* |  |  |  |  |  |
| - Coronary heart disease | 0.98 (0.95-1.01) p=0.25 | 0.92 (0.77-1.10) p=0.35 | 0.96 (0.91-1.01) p=0.14 | 0.95 (0.90-0.99) p=0.031 | 0.99 (0.90-1.09) p=0.86 |
| - Stroke | 1.01 (0.96-1.05) p=0.81 | 0.79 (0.63-0.99) p=0.040 | 1.00 (0.93-1.08) p=0.99 | 0.96 (0.89-1.03) p=0.29 | 1.06 (0.93-1.22) p=0.38 |
| - All cause mortality | 1.00 (0.96-1.03) p=0.85 | 0.69 (0.60-0.81) p=4×10-6 | 0.95 (0.90-1.00) p=0.067 | 0.98 (0.93-1.04) p=0.55 | 0.96 (0.86-1.05) p=0.35 |
| *EPIC-CVD* |  |  |  |  |  |
| - Coronary heart disease | 0.95 (0.86-1.05) p=0.30 | 0.86 (0.57-1.29) p=0.46 | 0.98 (0.86-1.12) p=0.77 | 0.82 (0.70-0.98) p=0.031 | 1.05 (0.61-1.79) p=0.87 |
| - Stroke | 0.99 (0.90-1.09) p=0.81 | 1.07 (0.72-1.60) p=0.73 | 1.03 (0.90-1.16) p=0.70 | 0.86 (0.71-1.01) p=0.07 | 0.79 (0.47-1.35) p=0.39 |
| *Copenhagen studies* |  |  |  |  |  |
| - Coronary heart disease | 1.00 (0.91-1.10) p=0.95 | 0.67 (0.37-1.21) p=0.18 | 0.95 (0.80-1.12) p=0.54 | 0.96 (0.83-1.11) p=0.58 | 1.19 (0.96-1.49) p=0.11 |
| - Stroke | 1.10 (0.96-1.26) p=0.18 | 0.78 (0.36-1.71) p=0.53 | 0.85 (0.68-1.06) p=0.16 | 1.34 (1.08-1.67) p=0.009 | 1.09 (0.81-1.48) p=0.57 |
| - All cause mortality | 0.89 (0.80-0.99) p=0.030 | 0.58 (0.29-1.18) p=0.14 | 0.78 (0.64-0.94) p=0.009 | 0.93 (0.79-1.09) p=0.37 | 0.76 (0.61-0.95) p=0.018 |

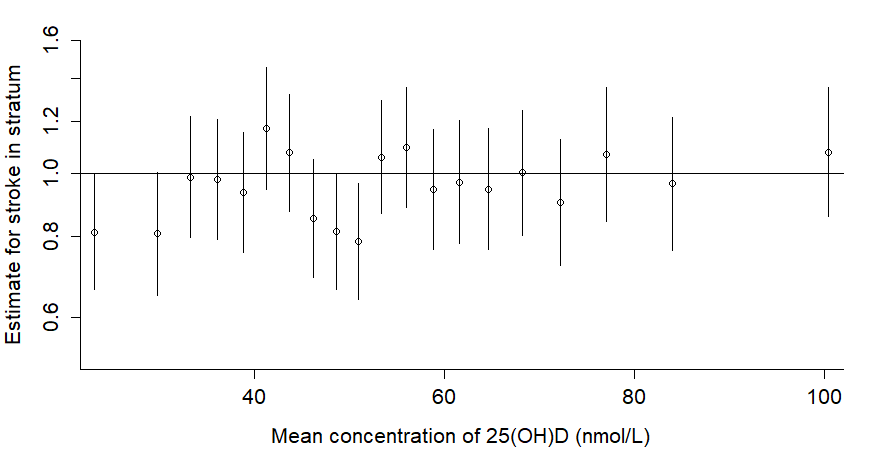
Estimates (95% confidence intervals) represent odds ratio per 10 nmol/L higher genetically-predicted concentration of 25(OH)D.

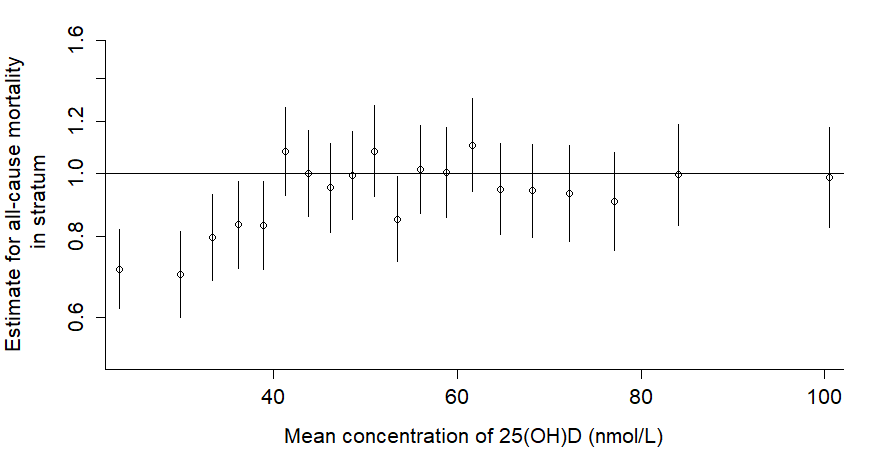
Table 3: Mendelian randomization estimates in UK Biobank and Copenhagen studies for cause-specific mortality: odds ratios (95% confidence intervals) per 10 nmol/L higher genetically-predicted 25(OH)D concentration in strata of the population defined by residual concentration of 25(OH)D.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mortality | All-cause | Cardiovascular | Cancer | Other |
| Sample | **OR (95% CI)** | **OR (95% CI)** | **OR (95% CI)** | **OR (95% CI)** |
| Overall | 0.99 (0.95-1.02) p=0.39 | 1.01 (0.95-1.08) p=0.71 | 0.98 (0.93-1.02) p=0.29 | 1.00 (0.94-1.06) p=0.99 |
| Stratifying on vitamin D levels: |  |  |  |  |
| Deficient  (<25 nmol/L) | 0.69 (0.59-0.80) p=1×10-6 | 0.69 (0.52-0.92) p=0.011 | 0.81 (0.65-1.02) p=0.09 | 0.68 (0.54-0.85) p=0.0009 |
| Insufficient  (25 – 49 nmol/L) | 0.94 (0.89-0.99) p=0.013 | 1.00 (0.90-1.11) p=0.99 | 0.93 (0.87-1.00) p=0.046 | 0.92 (0.84-1.01) p=0.071 |
| Sufficient  (50 – 74 nmol/L) | 0.98 (0.93-1.03) p=0.39 | 0.98 (0.88-1.09) p=0.72 | 0.99 (0.93-1.06) p=0.87 | 0.97 (0.88-1.07) p=0.51 |
| Adequate  (>75 nmol/L) | 0.92 (0.83-1.01) p=0.067 | 0.97 (0.79-1.18) p=0.72 | 0.92 (0.81-1.04) p=0.16 | 0.99 (0.83-1.17) p=0.90 |

Figure 2: Mendelian randomization estimates in UK Biobank: odds ratios (95% confidence intervals) per 10 nmol/L higher genetically-predicted vitamin D concentration in strata of the population defined by residual concentration of 25(OH)D for: (a) coronary heart disease, (b) stroke, (c) all-cause mortality.

a. 

b. 

c. 

REFERENCES

1. Wang, T.J., et al., *Vitamin D deficiency and risk of cardiovascular disease.* Circulation, 2008. **117**(4): p. 503-511.

2. Brøndum-Jacobsen, P., et al., *25-hydroxyvitamin d levels and risk of ischemic heart disease, myocardial infarction, and early death: population-based study and meta-analyses of 18 and 17 studies.* Arterioscler Thromb Vasc Biol, 2012. **32**(11): p. 2794-802.

3. Schottker, B., et al., *Strong associations of 25-hydroxyvitamin D concentrations with all-cause, cardiovascular, cancer, and respiratory disease mortality in a large cohort study.* Am J Clin Nutr, 2013. **97**(4): p. 782-93.

4. Berghout, B.P., et al., *Vitamin D Status and Risk of Stroke.* Stroke, 2019. **50**(9): p. 2293-2298.

5. Anderson, J.L., et al., *Relation of vitamin D deficiency to cardiovascular risk factors, disease status, and incident events in a general healthcare population.* Am J Cardiol, 2010. **106**(7): p. 963-8.

6. Heath, A.K., et al., *Vitamin D status and mortality: A systematic review of observational studies.* International Journal of Environmental Research and Public Health, 2019. **16**(3).

7. Manson, J.E., et al., *Vitamin D Supplements and Prevention of Cancer and Cardiovascular Disease.* New England Journal of Medicine, 2019. **380**(1): p. 33-44.

8. Zittermann, A., 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.* European Heart Journal, 2017. **38**(29): p. 2279-2286.

9. Kopecky, S.L., et al., *Lack of Evidence Linking Calcium With or Without Vitamin D Supplementation to Cardiovascular Disease in Generally Healthy Adults: A Clinical Guideline From the National Osteoporosis Foundation and the American Society for Preventive Cardiology.* Ann Intern Med, 2016. **165**(12): p. 867-868.

10. Pilz, S., et al., *Vitamin D and cardiovascular disease prevention.* Nat Rev Cardiol, 2016. **13**(7): p. 404-17.

11. Scragg, R., 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**(6): p. 608-616.

12. Bjelakovic, G., et al., *Vitamin D supplementation for prevention of mortality in adults.* Cochrane Database Syst Rev, 2014(1): p. Cd007470.

13. Lawlor, D.A., et al., *Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology.* Statistics in Medicine, 2008. **27**(8): p. 1133-1163.

14. Thanassoulis, G. and C.J. O'Donnell, *Mendelian randomization: nature's randomized trial in the post-genome era.* JAMA, 2009. **301**(22): p. 2386-2388.

15. Brøndum-Jacobsen, P., et al., *No evidence that genetically reduced 25-hydroxyvitamin D is associated with increased risk of ischaemic heart disease or myocardial infarction: a Mendelian randomization study.* International Journal of Epidemiology, 2015. **44**(2): p. 651-661.

16. Manousaki, D., et al., *Mendelian randomization studies do not support a role for vitamin D in coronary artery disease.* Circulation: Cardiovascular Genetics, 2016. **9**(4): p. 349-356.

17. Larsson, S.C., et al., *Serum 25-Hydroxyvitamin D Concentrations and Ischemic Stroke and Its Subtypes.* Stroke, 2018. **49**(10): p. 2508-2511.

18. Huang, T., 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 Medicine, 2019. **17**(1): p. 160.

19. Afzal, S., et al., *Genetically low vitamin D concentrations and increased mortality: mendelian randomisation analysis in three large cohorts.* BMJ : British Medical Journal, 2014. **349**: p. g6330.

20. Aspelund, T., et al., *Effect of genetically low 25-hydroxyvitamin D on mortality risk: Mendelian randomization analysis in 3 large european cohorts.* Nutrients, 2019. **11**(1).

21. Zhang, R., 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**(4): p. 810-819.

22. Crowe, F.L., 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**: p. 105480.

23. Sudlow, C., 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 Medicine, 2015. **12**(3): p. e1001779.

24. Astle, W.J., et al., *The allelic landscape of human blood cell trait variation and links to common complex disease.* Cell, 2016. **167**(5): p. 1415-1429. e19.

25. Danesh, J., 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.* European Journal of Epidemiology, 2007. **22**(2): p. 129.

26. Appleyard, M., *The Copenhagen City Heart Study. Østerbroundersøgelsen. A book of tables with data from the first examination (1976-78) and a five year follow-up (1981-83).* Scandinavian Journal of Social Medicine., 1989. **41**: p. 1-160.

27. Nordestgaard, B.G., et al., *The effect of elevated body mass index on ischemic heart disease risk: causal estimates from a Mendelian randomisation approach.* PLOS Medicine, 2012. **9**(5): p. e1001212.

28. Holick, M.F., et al., *Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline.* J Clin Endocrinol Metab, 2011. **96**(7): p. 1911-30.

29. Ross, A.C., et al., *The 2011 Report on Dietary Reference Intakes for Calcium and Vitamin D from the Institute of Medicine: What Clinicians Need to Know.* The Journal of Clinical Endocrinology & Metabolism, 2011. **96**(1): p. 53-58.

30. Wang, T.J., et al., *Common genetic determinants of vitamin D insufficiency: a genome-wide association study.* Lancet, 2010. **376**(9736): p. 180-188.

31. Mokry, L.E., et al., *Vitamin D and risk of multiple sclerosis: a Mendelian randomization study.* PLOS Medicine, 2015. **12**(8): p. e1001866.

32. Prentice, R.L., *A case-cohort design for epidemiologic cohort studies and disease prevention trials.* Biometrika, 1986. **73**(1): p. 1-11.

33. White, I.R., et al., *Meta‐analysis of non‐linear exposure‐outcome relationships using individual participant data: a comparison of two methods.* Statistics in Medicine, 2019. **38**(3): p. 326-338.

34. Thompson, S., et al., *Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies.* International Journal of Epidemiology, 2010. **39**(5): p. 1345-1359.

35. Staley, J.R. and S. Burgess, *Semiparametric methods for estimation of a nonlinear exposure‐outcome relationship using instrumental variables with application to Mendelian randomization.* Genetic Epidemiology, 2017. **41**(4): p. 341-352.

36. Zhang, Y., et al., *Association between vitamin D supplementation and mortality: systematic review and meta-analysis.* Brit Med J, 2019. **366**: p. l4673.

37. Pittas, A.G., et al., *Vitamin D Supplementation and Prevention of Type 2 Diabetes.* N Engl J Med, 2019. **381**(6): p. 520-530.

38. Gupta, A., et al., *Effect of Vitamin D and calcium supplementation on ischaemic stroke outcome: a randomised controlled open-label trial.* Int J Clin Pract, 2016. **70**(9): p. 764-70.

39. Narasimhan, S. and P. Balasubramanian, *Role of Vitamin D in the Outcome of Ischemic Stroke- A Randomized Controlled Trial.* Journal of clinical and diagnostic research : JCDR, 2017. **11**(2): p. CC06-CC10.

40. Burgess, S., et al., *Use of Mendelian randomisation to assess potential benefit of clinical intervention.* British Medical Journal, 2012. **345**: p. e7325.

41. Ference, B.A., *How to use Mendelian randomization to anticipate the results of randomized trials.* European Heart Journal, 2018. **39**(5): p. 360-362.

42. Weishaar, R.E., et al., *Involvement of vitamin D3 with cardiovascular function. III. Effects on physical and morphological properties.* American Journal of Physiology, 1990. **258**(1): p. E134-E142.

43. Mancuso, P., et al., *1, 25-Dihydroxyvitamin-D3 treatment reduces cardiac hypertrophy and left ventricular diameter in spontaneously hypertensive heart failure-prone (cp/+) rats independent of changes in serum leptin.* Journal of Cardiovascular Pharmacology, 2008. **51**(6): p. 559-564.

44. Merke, J., 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.* The Journal of Clinical Investigation, 1989. **83**(6): p. 1903-1915.

45. Somjen, D., et al., *25-Hydroxyvitamin D3-1α-hydroxylase is expressed in human vascular smooth muscle cells and is upregulated by parathyroid hormone and estrogenic compounds.* Circulation, 2005. **111**(13): p. 1666-1671.

46. Danielsson, C., et al., *Potent gene regulatory and antiproliferative activities of 20‐methyl analogues of 1, 25 dihydroxyvitamin D3.* Journal of Cellular Biochemistry, 1996. **63**(2): p. 199-206.

47. Manolagas, S., et al., *The antiproliferative effect of calcitriol on human peripheral blood mononuclear cells.* The Journal of Clinical Endocrinology & Metabolism, 1986. **63**(2): p. 394-400.

48. Li, Y.C., et al., *1, 25-Dihydroxyvitamin D 3 is a negative endocrine regulator of the renin-angiotensin system.* The Journal of Clinical Investigation, 2002. **110**(2): p. 229-238.

49. Fleet, J.C., et al., *Vitamin D and cancer: a review of molecular mechanisms.* Biochemical Journal, 2012. **441**(1): p. 61-76.

50. Holmes, M.V., M. Ala-Korpela, and G.D. Smith, *Mendelian randomization in cardiometabolic disease: challenges in evaluating causality.* Nature Reviews Cardiology, 2017. **14**(10): p. 577-590.

51. Fry, A., et al., *Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population.* American Journal of Epidemiology, 2017. **186**(9): p. 1026-1034.

SUPPLEMENTARY TABLES AND FIGURES

Supplementary Table 1: Methods for measurement of 25(OH)D concentration in the Vitamin D Studies Collaboration

|  |  |
| --- | --- |
| Method | Number of studies |
| Radioimmunoassay | 11 |
| Automated immunoassay | 4 |
| Competitive protein binding | 1 |
| Immunometric assay | 1 |
| High-performance liquid chromatography mass spectrometry | 13 |
| Electro-chemiluminescence immunoassay | 1 |

Supplementary Table 2: List of genetic variants for genetic risk score in UK Biobank and EPIC-CVD

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Chromosome: Position (hg19) | rsID | Effect allele | Other allele | Marginal association with 25(OH)D (nmol/L) |
| 4:72617775 | rs1352846 | G | A | 0.172 |
| 4:72618334 | rs7041 | C | A | -0.045 |
| 4:72634343 | rs4694431 | T | C | -0.034 |
| 4:72770563 | rs139148694 | GTGCTTTTATCAA | G | 0.028 |
| 11:14339328 | rs16913816 | A | G | -0.031 |
| 11:14900931 | rs117913124 | A | G | 0.503 |
| 11:14912573 | rs117576073 | T | G | 0.246 |
| 11:14913575 | rs12794714 | A | G | 0.139 |
| 11:14913645 | rs202122669 | A | G | -0.615 |
| 11:14913900 | rs187639972 | C | G | -0.360 |
| 11:14941652 | rs117115472 | G | C | 0.148 |
| 11:71157867 | rs139168803 | A | G | -0.188 |
| 11:71158672 | rs12573951 | G | A | -0.045 |
| 11:71161063 | rs7928249 | G | A | -0.131 |
| 11:71180762 | rs549000212 | A | C | -0.364 |
| 11:71290740 | rs4081429 | C | A | 0.017 |
| 20:52714706 | rs6123359 | G | A | -0.026 |
| 20:52731402 | rs6127099 | T | A | 0.013 |
| 20:52735238 | rs35870583 | GT | G | 0.027 |
| 20:52737123 | rs2585442 | G | C | -0.025 |
| 20:52788925 | rs2762942 | A | G | -0.053 |

Supplementary Table 3: Baseline characteristics of participants in the observational analyses

|  |  |  |  |
| --- | --- | --- | --- |
| Characteristic | Cohorts | N | Mean (SD) or % |
| 25(OH)D (nmol/L) | 40 | 500692 | 52.0 (21.7) |
| <25 nmol/L, Deficient | 40 | 65313 | 13% |
| 25-49 nmol/L, Insufficient | 40 | 208223 | 42% |
| 50-74 nmol/L, Sufficient | 40 | 165162 | 33% |
| ≥75 nmol/L, Adequate | 40 | 62264 | 12% |
| **Age and physical measures** |  |  |  |
| Age at survey (yrs) | 40 | 500962 | 60.7 (8.7) |
| Height (cm) | 40 | 499369 | 167 (9) |
| Weight (kg) | 40 | 498356 | 74.0 (15.4) |
| BMI (kg/m^2) | 40 | 498019 | 26.6 (4.7) |
| Waist (cm) | 29 | 464017 | 89.0 (13.2) |
| Hip (cm) | 29 | 463946 | 102 (9) |
| Waist:Hip ratio | 29 | 463881 | 0.87 (0.09) |
| SBP (mmHg) | 38 | 488928 | 138 (19) |
| DBP (mmHg) | 37 | 486117 | 80.2 (10.3) |
| **Lipids** |  |  |  |
| Total cholesterol (mmol/l) | 37 | 489120 | 5.92 (1.02) |
| Friedwald LDL cholesterol (mmol/l) | 32 | 431296 | 3.80 (0.89) |
| Measured LDL cholesterol (mmol/l) | 3 | 388309 | 3.37 (0.76) |
| Non-HDL cholesterol (mmol/l) | 35 | 452863 | 4.49 (1.00) |
| HDL-C (mmol/l) | 35 | 453646 | 1.40 (0.38) |
| Log Triglycerides (mmol/l) | 34 | 476028 | 0.34 (0.52) |
| Apolipoprotein A1 (g/l) | 9 | 374898 | 1.51 (0.27) |
| Apolipoprotein B (g/l) | 9 | 407933 | 1.05 (0.22) |
| Log Lp(a) (mg/dl) | 9 | 329672 | 3.18 (1.11) |
| **Glycaemia markers** |  |  |  |
| Log Glucose (mmol/l) | 29 | 426373 | 1.65 (0.19) |
| Log Fasting glucose (mmol/l) | 16 | 39327 | 1.64 (0.19) |
| HbA1c (%) | 9 | 389733 | 5.52 (0.86) |
| **Inflammation markers** |  |  |  |
| Fibrinogen (µmol/l) | 12 | 24246 | 9.15 (2.27) |
| Log CRP (mg/l) | 29 | 447320 | 0.43 (1.06) |
| Log White cell count (x10^9/l) | 13 | 400868 | 1.85 (0.26) |
| Albumin (g/l) | 25 | 411035 | 43.7 (2.8) |
| **Kidney function** |  |  |  |
| Log Creatinine (µmol/l) | 30 | 446152 | 4.40 (0.20) |
| Log eGFR by MDRD (ml/min/1.73m^2) | 30 | 446152 | 4.31 (0.20) |
| **Bone-related markers** |  |  |  |
| Calcium (mmol/l) | 21 | 401249 | 2.38 (0.10) |
| Log PTH (ng/L) | 14 | 21379 | 1.30 (0.44) |
| Log Phosphate (mmol/L) | 13 | 365621 | 0.13 (0.14) |
| Log 1,25(OH)2D (pmol/L) | 7 | 3160 | 4.38 (0.42) |
| Log Alkaline Phosphatase (iu/l) | 10 | 419940 | 4.16 (0.28) |
| **Categorical variables** |  |  |  |
| Sex | 40 | 500962 |  |
| Male | 32 | 225383 | 45% |
| Female | 36 | 275579 | 55% |
| Ethnic group (4 groups) | 36 | 489927 |  |
| White | 36 | 463935 | 95% |
| Asian | 7 | 8504 | 2% |
| Black | 8 | 10749 | 2% |
| Other | 8 | 6739 | 1% |
| Smoking status | 40 | 499495 |  |
| Other | 40 | 431683 | 86% |
| Current | 40 | 67812 | 14% |
| Alcohol status | 34 | 481755 |  |
| Other | 31 | 54161 | 11% |
| Current | 34 | 427594 | 89% |
| History of diabetes | 40 | 488586 |  |
| No | 39 | 465183 | 95% |
| Yes | 40 | 23403 | 5% |
| Season | 40 | 500962 |  |
| Winter | 36 | 99694 | 20% |
| Spring | 35 | 138518 | 28% |
| Summer | 39 | 142646 | 28% |
| Autumn | 38 | 120104 | 24% |
| Level of education reached | 28 | 456164 |  |
| Primary | 25 | 25375 | 6% |
| Secondary | 28 | 176416 | 39% |
| Vocational/University | 24 | 254373 | 56% |

Supplementary Table 4: Details of studies contributing to the observational analyses (to be revised)



Supplementary Table 5: Mendelian randomization estimates for stroke in UK Biobank divided into overall stroke (10,489 events), incident-only stroke (5044 events), ischaemic stroke (including unknown, 4164 events), and haemorrhagic stroke (intracerebral plus subarachnoid haemorrhage, 1194 events): odds ratios (95% confidence intervals) per 10 nmol/L higher genetically-predicted vitamin D concentration

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stroke | Overall | Incident only | Ischaemic | Haemorrhagic |
| Sample | **OR (95% CI)** | **OR (95% CI)** | **OR (95% CI)** | **OR (95% CI)** |
| Overall | 1.01 (0.96-1.05) p=0.81 | 0.98 (0.91-1.04) p=0.49 | 1.01 (0.94-1.09) p=0.73 | 0.91 (0.80-1.04) p=0.19 |
| Females | 0.99 (0.92-1.07) p=0.85 | 0.94 (0.85-1.04) p=0.24 | 1.00 (0.88-1.12) p=0.95 | 0.85 (0.70-1.03) p=0.09 |
| Males | 1.01 (0.96-1.08) p=0.64 | 1.00 (0.92-1.09) p=0.95 | 1.02 (0.93-1.12) p=0.64 | 0.99 (0.82-1.19) p=0.88 |
| Stratifying on vitamin D levels: |  |  |  |  |
| Deficient  (<25 nmol/L) | 0.79 (0.63-0.99) p=0.040 | 0.79 (0.56-1.13) p=0.20 | 0.77 (0.54-1.10) p=0.15 | 0.96 (0.42-2.17) p=0.92 |
| Insufficient  (25 – 49 nmol/L) | 1.00 (0.93-1.08) p=0.99 | 1.01 (0.91-1.12) p=0.90 | 0.98 (0.88-1.10) p=0.77 | 1.09 (0.88-1.35) p=0.45 |
| Sufficient  (50 – 74 nmol/L) | 0.96 (0.89-1.03) p=0.29 | 0.95 (0.86-1.05) p=0.28 | 1.03 (0.92-1.15) p=0.58 | 0.81 (0.66-0.99) p=0.036 |
| Adequate  (>75 nmol/L) | 1.06 (0.93-1.22) p=0.38 | 0.94 (0.78-1.14) p=0.53 | 0.97 (0.78-1.20) p=0.75 | 0.83 (0.57-1.21) p=0.34 |

Supplementary Table 6: Mendelian randomization estimates for coronary heart disease in UK Biobank divided into overall (22,363 events) and incident-only CHD (5447 events): odds ratios (95% confidence intervals) per 10 nmol/L higher genetically-predicted vitamin D concentration

|  |  |  |
| --- | --- | --- |
| CHD | Overall | Incident only |
| Sample | **OR (95% CI)** | **OR (95% CI)** |
| Overall | 0.98 (0.95-1.01) p=0.25 | 1.00 (0.94-1.07) p=0.98 |
| Females | 0.98 (0.92-1.04) p=0.54 | 1.01 (0.89-1.15) p=0.85 |
| Males | 0.98 (0.94-1.02) p=0.34 | 1.00 (0.93-1.07) p=0.92 |
| Stratifying on vitamin D levels: |  |  |
| Deficient  (<25 nmol/L) | 0.92 (0.77-1.10) p=0.35 | 1.04 (0.75-1.45) p=0.80 |
| Insufficient  (25 – 49 nmol/L) | 0.96 (0.91-1.01) p=0.14 | 0.97 (0.88-1.07) p=0.56 |
| Sufficient  (50 – 74 nmol/L) | 0.95 (0.90-0.99) p=0.031 | 1.00 (0.91-1.11) p=0.96 |
| Adequate  (>75 nmol/L) | 0.99 (0.90-1.09) p=0.86 | 0.96 (0.79-1.17) p=0.66 |

Supplementary Figure 1: Mendelian randomization estimates in EPIC-CVD: odds ratios (95% confidence intervals) per 10 nmol/L higher genetically-predicted vitamin D concentration in strata of the population defined by residual concentration of vitamin D for: (a) coronary heart disease, (b) stroke.

a.

b.

Supplementary Figure 2: Mendelian randomization estimates in Copenhagen: odds ratios (95% confidence intervals) per 10 nmol/L higher genetically-predicted vitamin D concentration in strata of the population defined by residual concentration of vitamin D for: (a) coronary heart disease, (b) stroke, (c) all-cause mortality.



Supplementary Figure 3. Study-specific dose**—**response curves of association of 25(OH)D concentration with outcomes by study (top) and pooled estimates by data source (bottom) adjusted for conventional cardiovascular risk factors.



Reference value is 50 nmol/L. The shaded area represents the 95% confidence interval for the dose—response curve. Study-specific analyses involved fractional polynomial modelling of continuous associations of 25(OH)D and outcomes using Cox regression stratified by sex and adjusted for conventional risk factors, namely: age, month, current smoking, total cholesterol, HDL cholesterol, systolic blood pressure, history of diabetes, and body mass index, followed by random effects meta-analysis (**see Methods**).

\* The EPIC-CVD study was specifically designed as a case-cohort study of CVD outcomes therefore does not contribute to analysis of non-CVD outcomes in totality.

Supplementary Figure 4. Dose**—**response association of 25(OH)D concentrations with outcomes by deciles (top) and clinical categories (bottom).

Reference value is decile 5 (close to 50 nmol/L) in top figure and is the Adequate clinical subgroup (>75 nmol/L) in the bottom figure. The associations of 25(OH)D and outcomes were modelled using Cox regression stratified by sex and adjusted for conventional risk factors, namely: age, month, current smoking, total cholesterol, HDL cholesterol, systolic blood pressure, history of diabetes, and body mass index, followed by random effects meta-analysis (**see Methods**).