ESC European Society of Cardiology

European Heart Journal (2021) **00**, 1–10
https://doi.org/10.1093/eurheartj/ehab809

CLINICAL RESEARCH
Epidemiology and prevention

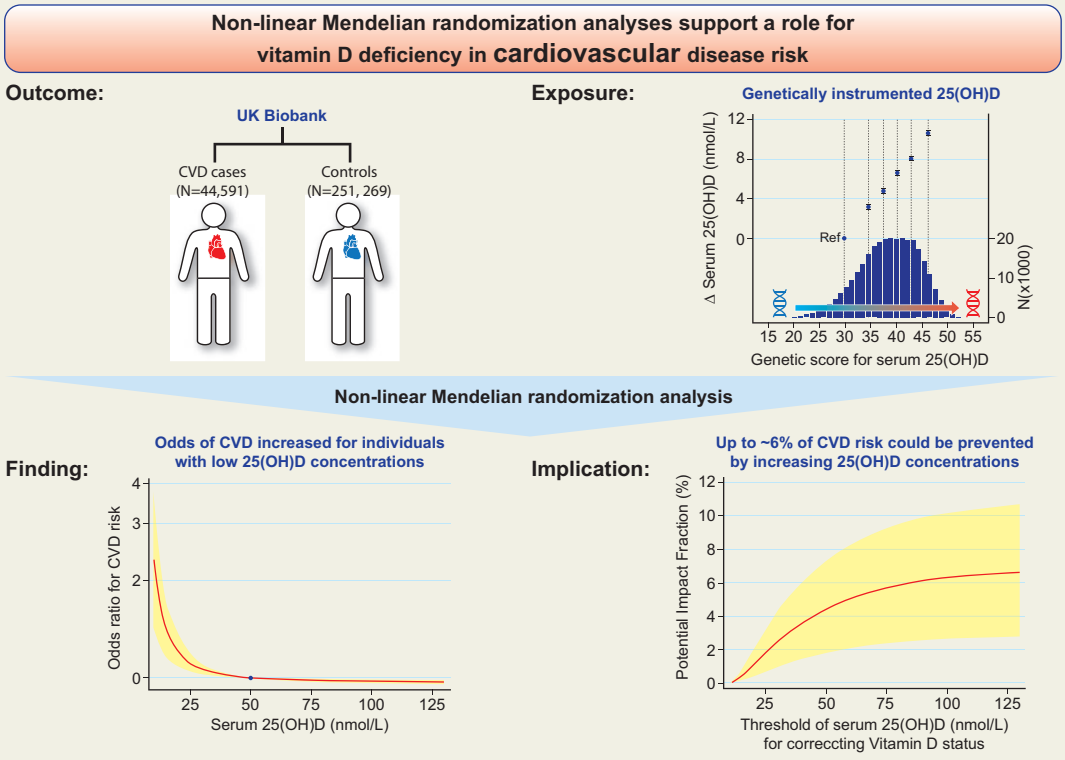
Non-linear Mendelian randomization analyses support a role for vitamin D deficiency in cardiovascular disease risk

Ang Zhou [1,2], Joseph B. Selvanayagam[2,3], and Elina Hyppönen[1,2,4]*

[1]Australian Center for Precision Health, University of South Australia Cancer Research Institute, GPO Box 2471, Adelaide, SA 5001, Australia; [2]South Australian Health and Medical Research Institute, North Terrace, Adelaide, SA 5000, Australia; [3]Department of Cardiovascular Medicine, Flinders Medical Centre, Bedford Park, Adelaide, SA 5042, Australia; and [4]Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, 30 Guilford Street, London, UK

Received 17 June 2021; revised 28 September 2021; editorial decision 28 October 2021; accepted 12 November 2021

Graphical Abstract



Non-linear Mendelian randomization analyses support a L-shaped association of serum 25-hydroxyvitamin D concentration with CVD risk, suggesting that with respect to cardiovascular health the benefits of improving vitamin D status are the strongest for those within the deficiency range.

* Corresponding author. Tel: +61 8 830 22518, Email: elina.hypponen@unisa.edu.au

Aims Low vitamin D status is associated with a higher risk for cardiovascular diseases (CVDs). Although most existing linear Mendelian randomization (MR) studies reported a null effect of vitamin D on CVD risk, a non-linear effect cannot be excluded. Our aim was to apply the non-linear MR design to investigate the association of serum 25-hydroxyvitamin D [25(OH)D] concentration with CVD risk.

Methods and results The non-linear MR analysis was conducted in the UK Biobank with 44 519 CVD cases and 251 269 controls. Blood pressure (BP) and cardiac-imaging-derived phenotypes were included as secondary outcomes. Serum 25(OH)D concentration was instrumented using 35 confirmed genome-wide significant variants.
We also estimated the potential reduction in CVD incidence attributable to correction of low vitamin D status. There was a L-shaped association between genetically predicted serum 25(OH)D and CVD risk ($P_{non\text{-}linear}$ = 0.007), where CVD risk initially decreased steeply with increasing concentrations and levelled off at around 50 nmol/L. A similar association was seen for systolic ($P_{non\text{-}linear}$ = 0.03) and diastolic ($P_{non\text{-}linear}$ = 0.07) BP. No evidence of association was seen for cardiac-imaging phenotypes (P = 0.05 for all). Correction of serum 25(OH)D level below 50 nmol/L was predicted to result in a 4.4% reduction in CVD incidence (95% confidence interval: 1.8–7.3%).

Conclusion Vitamin D deficiency can increase the risk of CVD. Burden of CVD could be reduced by population-wide correction of low vitamin D status.

Keywords Non-linear Mendelian randomization • Vitamin D • Serum 25-hydroxyvitamin Dconcentration • Cardiovascular diseases • Systolic blood pressure • Diastolic blood pressure • Cardiac-imaging phenotypes

Translational perspective

This study employed the non-linear Mendelian randomization approach to characterize the shape of the association of serum 25-hydroxyvitamin D [25(OH)D] concentration with cardiovascular diseases (CVDs). An L-shaped association of genetically predicted serum 25(OH)D concentration with CVD risk and blood pressure was observed, providing evidence for benefits are the strongest for those within the deficiency range. These results suggest that improving vitamin D status among people with low concentrations can support heart health and that a population-wide approach to eradicate vitamin D deficiency could reduce the burden of CVDs.

Introduction

As an essential micronutrient, vitamin D is well-known for its role in calcium homeostasis and bone health. Beyond skeletal health, low vitamin D status has also been associated with higher risk for cardiovascular disease (CVD) and mortality.[1–5] However, the causal nature of such associations remains debated, as no beneficial effect is seen in randomized controlled trials (RCTs) testing the effects of vitamin D supplementation.[6] The presence of a threshold effect has been proposed, where disease risk and benefits of vitamin D supplementation may only surface below certain thresholds of vitamin D status.[7] If this is true, possible benefits of vitamin D supplementation cannot be excluded, given the RCTs have been carried out in populations that are not clinically vitamin D deficient.[7,8] Indeed, recruiting participants with vitamin D deficiency in supplementation trials has been acknowledged to raise serious practical and ethical issues,[9] as severe deficiency is relatively rare and it is not acceptable to subject participants to undue harm.

Mendelian randomization (MR) is increasingly used to provide causal evidence for exposures, such as smoking[10] and alcohol intake,[11] where it is either unethical or infeasible to conduct RCTs. Sitting at the interface between observational studies and interventional trials, MR is considered as a natural analogue of the classical RCTs.[12] It uses genetic variants associated with the exposure of interest to approximate the exposure, and conditional on the key method assumptions being met, MR has the benefit of reducing bias due to confounding and reverse causation, commonly seen in observational studies.[12] However, the standard MR approach assumes linearity, and hence, it cannot describe the shape of the association. If the proposed threshold effect is true, where the disease risks are only related to a deficiency state,[7] the standard MR approach is likely to overlook such non-linear effect. This may explain the null findings from MR studies assessing the effect of vitamin D status on CVD risk.[13,14] A recently developed extension to the MR analyses allows for non-linear associations between the exposure and the outcome.[15] This promising new approach has already been used to investigate the association between body mass index (BMI) and mortality,[16] where it confirmed the well-established J-shaped association where both being underweight and excess weight increase the risks.[17–19] To better characterize the association between vitamin D status and CVD risk, we used the non-linear MR approach to investigate the causal association between serum 25-hydroxyvitamin D [25(OH)D, an established marker for nutritional vitamin D status] concentrations and CVD risk. We conducted the analyses in the UK Biobank, where information on genetic variants, 25(OH)D concentrations, and disease status are available for up to 295 788 participants. As secondary analyses, we

also examined the association of 25(OH)D with blood pressure (BP) and cardiovascular imaging phenotypes.

Methods

Study participants

The UK Biobank is a large prospective cohort study with over 500 000 participants aged 37–73 years recruited from 22 assessment centres across the UK between 13 March 2006 and 1 October 2009 with a goal to improve the prevention, diagnosis, and treatment of diseases of middle and old age.[20,21] Participants filled in questionnaires to provide broad information on health and lifestyles at baseline survey and provided blood samples for biomarker and genetic assays. Information on health outcomes has been enhanced by the linkage to electronic health records and mortality registrations. Sub-groups of participants also took part in a multimodal imaging study, which involved a cardiac magnetic resonance (CMR) imaging scan and carotid ultrasound imaging. Genetic profiling identified patterns of relatedness in the UK Biobank.[22] We restricted the analyses to unrelated individuals, who were identified as white British based on self-report and genetic profiling,[22] and excluded participants with mismatched information between self-reported and genetic sex. Final genetic analyses were conducted among individuals with complete information on serum 25(OH)D [calcidiol, 25(OH)D concentration, and relevant covariates (*n* up to 295 788, see Supplementary material online, *Figure S1*)]. The present study was conducted under UK Biobank application number 20175. The UK Biobank study was approved by the National Information Governance Board for Health and Social Care and North West Multicentre Research Ethical Committee (11/NW/0382). All participants provided informed consent to participate.

Cardiovascular traits

Our primary outcome measure is CVD, with BP and cardiovascular imaging phenotypes being secondary outcome measures. CVD events were identified using data linkage to Hospital Episodes Statistics and mortality data.[23] Cases were classified using International Classification of Disease 9 and 10 codes, as well as Classification of Interventions and Procedures (see Supplementary material online, *Table S1*). Controls were those who were free from a CVD diagnosis. For phenotypic analyses, to minimize potential reverse causality, we restricted CVD events to the incident cases, who had their first CVD-related diagnosis after the baseline visit, and removed prevalent CVD (i.e. those who had their first CVD-related diagnosis before the baseline visit) from the analyses. BP was measured during the baseline assessment using digital monitors (HEM-7015IT; Omron Healthcare Inc). We calculated the mean systolic (SBP) and diastolic BP (DBP) values from two BP measurements. For participants reported to be taking BP-lowering medication (21%), we accounted for medication use by adding 15 and 10 mmHg to SBP and DBP, respectively.[23,24] We included the following cardiac-imaging phenotypes in our analysis: augmentation index (AI, %) and central augmentation pressure (mmHg) obtained from the pulse wave analysis (PWA), mean carotid intima-medial thickness (μm) taken from the carotid imaging, and cardiac output (L/min) and left ventricular end-diastolic volume (mL), end-systolic volume (mL), and ejection fraction (%) attained from CMR imaging scan. The imaging studies were carried out at the UK Biobank imaging facility at Cheadle, Stockport, UK, with CMR being performed on a clinical wide bore 1.5 Tesla scanner (MAGNETOM Aera, Syngo Platform VD13A, Siemens Healthcare, Erlangen, Germany), carotid imaging being performed using a CardioHealth Station (Panasonic Healthcare Corporation of North America, Newark, NJ, USA) with a 9 MHz linear array transducer, and PWA being performed using a VICORDER® device (http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/vicorder_in_cmri.pdf). Protocols for the CMR scan and carotid imaging can be found elsewhere.[25,26] For each of these image-derived phenotypes, participants with values beyond 3 SD were removed from the analysis.

Serum 25-hydroxyvitamin D concentration

Blood samples of participants were collected at the time of recruitment. Serum 25(OH)D concentration (nmol/L) was measured using the LIAISON XL 25(OH)D assay (DiaSorin, Stillwater, USA).

Genetic instrument for serum 25-hydroxyvitamin D concentration

We constructed a weighted genetic score (VitaminD-GS) consisting of 35 single nucleotide polymorphisms (SNPs) to instrument serum 25(OH)D concentration. All 35 SNPs are genome-wide significant variants, discovered in a recent genome-wide association analysis (GWAS) for serum 25(OH)D concentration in the UK Biobank.[27] In the original GWAS, a total of 143 independent loci were identified.[27] We restricted variants to autosomal common SNPs (minor allele frequency >5%), which can also be replicated (with a consistent direction and a *P*-value < 0.05) in the earlier GWAS by the SUNLIGHT Consortium,[28] leaving us with 35 variants (see Supplementary material online, *Figure S2*). Benefits of replication in the SUNLIGHT Consortium are two-fold. It ensures the robustness of the GWAS signals, and also allows us to take weights for the vitaminD-GS from an independent sample, avoiding bias arising from using internal weights.[29] Information on the 35 variants can be found in Supplementary material online, *Table S2*. VitaminD-GS was constructed by first computing the weighted average of the number of 25(OH)D-increasing alleles for an individual, and then multiplying it by the number of available variants. As aforementioned, the weight for each SNP was the effect estimate of the association of the SNP with serum 25(OH)D in the SUNLIGHT Consortium.[28] As a sensitivity analysis, we constructed an alternative instrument using a broader set of SNPs consisting of 122 autosomal variants (excluding 1 variant on the sex chromosome and 20 insertion/deletion variants from the original 143 GWAS variants, see Supplementary material online, *Text S1*, Supplementary material online, *Figure S2*, and Supplementary material online, *Table S2*).

Covariates

For the phenotypic analyses, we adjusted for a wide range of potential confounders, which were obtained during the baseline assessment. These covered basic and sociodemographics (age, sex, assessment centre, area deprivation, and education), anthropometric measures (BMI and waist circumference), lifestyle factors (smoking, alcohol intake, physical activity), and general health indicators (self-reported health status and long-term illness). Smoking was grouped into 10 categories (non-smokers, ex-smokers, current smokers with no information on the type of tobacco that they smoke, cigar/pipe smokers, cigarette smokers <1–5 cigs/day, 6–10 cigs/day, 11–15 cigs/day, 16–20 cigs/day, 21–25 cigs/day, and >25 cigs/day). Alcohol intake was categorized as 'never', 'special occasions only', '1–3 times per month', '1 or 2 times/week', '3–4 times/week', or '≥5 times/week', and intensity of physical activity as 'light', 'moderate', or 'vigorous'. Self-reported health was classified as 'poor', 'fair', 'good', or 'excellent', and long-term illness as binary 'no' or 'yes'. We used Townsend deprivation index to indicate area deprivation,[30] and grouped education as 'None or vocational education', 'CSE (secondary education)', or 'A-levels or higher (further education)'. We also included nuisance variables that could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement (http://biobank.

ndph.ox.ac.uk/showcase/showcase/docs/biomarker_issues.pdf). For the genetic analyses, we included the following covariates: age, sex, assessment centre, birth location, SNP array, top 40 genetic principal components, and nuisance factors, which could affect serum 25(OH)D measurements. Adjustment of birth location and 40 genetic components are recommended to account for latent population structure in the UK Biobank.[31]

Statistical analysis

Phenotypic associations of serum 25(OH)D concentration with cardiovascular traits were examined by fitting logistic (for CVD risk) or linear (for continuous traits) regression models, adjusting for covariates covering demographic, anthropometric, lifestyle, general health, and socioeconomic aspects of participants, as well as nuisance factors affecting serum 25(OH)D measurements (see covariates section). Fractional polynomial models were used to determine the appropriate function form for serum 25(OH)D concentration.[32] Likelihood ratio test was used to compare model fit between the best-fitting fractional polynomial model and the linear model, taking $P < 0.05$ to indicate a non-linear association.

We performed linear and non-linear MR analyses to examine genetic evidence for the associations of serum 25(OH)D concentration with cardiovascular traits. For the linear MR analysis, we computed the MR estimate using the ratio of coefficients method.[33] Linear regression was used to estimate the association of vitaminD-GS with serum 25(OH)D and continuous cardiovascular traits, and logistic regression to estimate the association of vitaminD-GS with CVD risk. We adjusted these analyses for age, sex, assessment centre, birth location, SNP array, top 40 genetic principal components, and nuisance factors related to serum 25(OH)D measurements. We also performed the stratified analyses, in which MR estimates were computed within categories of residual 25(OH)D concentration (<25.0 nmol/L, 25–49.9 nmol/L, 50.0–74.9 nmol/L, and ≥75.0 nmol/L). Residual 25(OH)D concentration is defined as participants' serum 25(OH)D concentration minus the centred genetic contribution by vitaminD-GS. Stratification was performed using residual 25(OH)D concentrations [rather than raw 25(OH)D] to avoid collider bias.[34]

For non-linear MR, we used the fractional polynomial method to capture non-linear genetic associations.[15] Briefly, we stratified our sample into 100 strata using the residuals of serum 25(OH)D after regressing on vitaminD-GS. Within each stratum, we computed the localized average causal effect (LACE), which is the ratio of coefficient of vitaminD-GS–outcome association to that of vitaminD-GS–25(OH)D association. LACE is equivalent to the traditional linear MR estimate within a stratum. We performed meta-regression of LACE against stratum-specific mean 25(OH)D by fitting a range of fractional polynomial exposure–outcome models of Degree 1 and 2, and selected the best-fitting model based on the likelihood ratio test. The fractional polynomial test is reported for non-linearity, which compares the best-fitting fractional polynomial model of Degree 1 against the linear model.[15] Non-linear MR analyses assume that the association of genetic instrument with the exposure is constant over the entire distribution of exposure. To test this assumption, we computed vitaminD-25(OH)D estimate in each of the 100 strata and then examined the heterogeneity between strata using the trend test and Cochran's Q test.[15] Furthermore, as sensitivity analyses, we performed leave-one-out and leave-block-out analyses to examine if our primary non-linear MR analysis were driven by any particular variant or groups of variants (see Supplementary material online, *Texts S2* and *S3*). For the leave-block-out analyses, variants were assigned to functional blocks based on the traits that they were associated with in the PhenoScanner search[35] (see Supplementary material online, *Text S3*). We identified four functional blocks, including blocks for 'blood', 'metabolic', and 'renal' traits, with variants whose associated traits did not fall into one of these three blocks grouped together as the 'unclassified' block (see Supplementary material online, *Table S3*). Moreover, we also implemented non-linear MR analyses where we computed the LACE within each stratum of residuals of serum 25(OH)D using two-sample approaches, including inverse variance weighted, MR-Egger, weighted median, weighted mode, and MR-Presso (see Supplementary material online, *Text S4*). As the two-sample methods have different assumptions on horizontal pleiotropy, a good agreement across these methods implies that the result is robust to different patterns of horizontal pleiotropy.[36]

We calculated the potential impact fraction (PIF)[37] to estimate the average reductions in the incidence of CVD attributable to correction of low vitamin D status (see Supplementary material online, *Text S5*).

Analyses for phenotypic associations, linear MR, and PIF were performed using STATA, version 14.1 (StataCorp LP, College Station, TX, USA). Non-linear MR analyses were conducted in R (version 4.0.2) using the NLMR package (version 2.0).[15]

Results

Up to 267 980 participants were involved in phenotypic analysis (see Supplementary material online, *Figure S1*). The average 25(OH)D concentration was 50.0 nmol/L (range: 10–340 nmol/L), and 11.4% ($n = 32\ 868$) of the participants had concentrations < 25 nmol/L, 41.3% ($n = 119\ 243$) 25–49.9 nmol/L, 35.3% ($n = 101\ 848$) 50–74.9 nmol/L, 10.5% ($n = 30\ 314$) 75–99.9 nmol/L, 1.4% ($n = 4110$) 100–124.9 nmol/L, and 0.2% ($n = 570$) ≥125 nmol/L. Only 107 individuals had concentrations above 150 nmol/L. There were notable variations in serum 25(OH)D concentration and the incidence of CVD with respect to distributions of demographics, lifestyle, general health, and socioeconomic factors (*Table 1*). For the phenotypic association analyses, we have adjusted for all these covariates.

Instrument validation

VitaminD-GS was robustly associated with serum 25(OH)D concentration in the UK Biobank, explaining 2.8% of variation (F statistic = 8672, $P < 1.0E\text{-}300$, Supplementary material online, *Figure S3*). We examined the association of vitaminD-GS with serum 25(OH)D across 100 strata of residuals of serum 25(OH)D. We detected evidence for heterogeneity ($P_{\text{Cochran's } Q} = 4.15E\text{-}9$; $P_{\text{trend}} = 0.38$), where vitaminD-GS–25(OH)D association in the 1st and 100th stratum appeared to be outliers (see Supplementary material online, *Figure S4*). We found little evidence that vitaminD-GS was associated with potential confounders in the UK Biobank, including BMI, smoking, alcohol intake, physical activity, education, and Townsend deprivation index (see Supplementary material online, *Table S4*, uncorrected $P \geq 0.044$ for all). We also examined vitaminD-GS–confounder associations across 100 strata of residuals of serum 25(OH)D, and found no evidence of association after accounting for multiple testing (see Supplementary material online, *Figure S5*).

Phenotypic associations

Every 10 nmol/L increase in serum 25(OH)D was associated with a 1.6% lower odds of CVD [odds ratio (OR): 0.98, 95% confidence interval (CI): 0.98–0.99, *Figure 1A*, $P_{\text{linear}} = 0.0001$, $P_{\text{non-linear}} = 0.48$]. Non-linear associations were evident for SBP and DBP (*Figure 1B and C*, $P_{\text{non-linear}} \leq 1.8E\text{-}04$). Null associations were observed for cardiac-

Table 1 Serum 25-hydroxyvitamin D and cardiovascular disease by baseline characteristics in the UK Biobank

		25(OH)D (nmol/L)	CVD	Incident CVD
	n (%)	Mean (SD)	Case (%)	Case (%)
Age				
<65 years	232 521 (80.47)	49.49 (20.99)	26 353 (11.33)	13 791 (6.27)
≥65 years	56 432 (19.53)	52.23 (20.51)	15 425 (27.33)	7014 (14.61)
P[a]		2.15E-188	<1.0E-300	<1.0E-300
Sex				
Male	135 912 (47.04)	50.06 (21.00)	27 060 (19.91)	13 110 (10.75)
Female	153 041 (52.96)	50.00 (20.85)	14 718 (9.62)	7695 (5.27)
P[a]		0.021	<1.0E-300	<1.0E-300
BMI				
<18.5 kg/m^2	1427 (0.49)	51.04 (24.10)	166 (11.63)	102 (7.48)
18.5–25 kg/m^2	96 302 (33.33)	53.06 (21.78)	9197 (9.55)	5030 (5.46)
25–30 kg/m^2	123 593 (42.77)	50.68 (20.54)	18 585 (15.04)	9330 (8.16)
≥30 kg/m^2	67 631 (23.41)	44.50 (19.17)	13 830 (20.45)	6343 (10.55)
P[a]		<1.0E-300	<1.0E-300	1.26E-194
Smoking				
Non-smokers	159 209 (55.10)	50.19 (20.61)	17 591 (11.05)	9398 (6.22)
Ex-smokers	101 758 (35.22)	50.98 (21.01)	18 622 (18.30)	8419 (9.20)
Smokers[b]	7371 (2.55)	49.33 (21.59)	1097 (14.88)	545 (7.99)
Cigars/Pipes	1626 (0.56)	45.93 (21.14)	392 (24.11)	212 (14.66)
<1 to 15 cigs/day	11 404 (3.95)	45.65 (21.64)	2110 (18.50)	1163 (11.12)
>15 cigs/day	7585 (2.62)	41.93 (21.76)	1966 (25.92)	1068 (15.97)
P[a]		<1.0E-300	<1.0E-300	6.64E-286
Alcohol intake				
Non-drinkers	17 985 (6.22)	46.29 (20.90)	3748 (20.84)	1632 (10.28)
Special occasions or 1–3 times/month	61 149 (21.16)	47.38 (20.31)	9389 (15.35)	4498 (8.00)
1 or 2 times/week	76 506 (26.48)	50.54 (20.72)	10 301 (13.46)	5185 (7.26)
3 or 4 times/week	70 947 (24.55)	51.50 (20.87)	9075 (12.79)	4651 (6.99)
Daily or almost daily	62 366 (21.58)	51.40 (21.46)	9265 (14.86)	4839 (8.35)
P[a]		<1.0E-300	2.04E-269	2.47E-76
Physical activity				
Light	87 259 (30.20)	46.34 (20.19)	13 923 (15.96)	6657 (8.32)
Moderate	144 193 (49.90)	50.64 (20.80)	19 497 (13.52)	9720 (7.23)
Vigorous	57 501 (19.90)	54.08 (21.42)	8358 (14.54)	4428 (8.27)
P[a]		<1.0E-300	3.11E-73	2.64E-22
Education				
None	47 373 (16.39)	50.90 (21.36)	11 622 (24.53)	5153 (12.60)
NVQ/CSE/A-levels	103 346 (35.77)	50.76 (21.16)	14 027 (13.57)	7169 (7.43)
Degree/professional	138 234 (47.84)	49.18 (20.56)	16 129 (11.67)	8483 (6.50)
P[a]		1.92E-98	5.39E-294	3.50E-94
Townsend deprivation index quartiles				
Q1 lowest	72 235 (25.00)	52.00 (20.69)	9253 (12.81)	4860 (7.16)
Q2	72 238 (25.00)	51.66 (20.67)	9818 (13.59)	5007 (7.43)
Q3	72 235 (25.00)	50.14 (20.80)	10 151 (14.05)	5173 (7.69)
Q4 highest	72 245 (25.00)	46.31 (21.04)	12 556 (17.38)	5765 (8.81)
P[a]		<1.0E-300	3.44E-285	3.76E-82
Self-rated health				
Excellent	50 613 (17.52)	53.03 (21.14)	3302 (6.52)	2300 (4.64)
Good	171 701 (59.42)	50.71 (20.70)	20 613 (12.01)	11 591 (7.13)
Fair	56 331 (19.49)	46.58 (20.61)	13 748 (24.41)	5647 (11.71)
Poor	10 308 (3.57)	42.81 (21.18)	4115 (39.92)	1267 (16.98)
P[a]		<1.0E-300	<1.0E-300	<1.0E-300

Continued

Table 1 *Continued*

	n (%)	25(OH)D (nmol/L) Mean (SD)	CVD Case (%)	Incident CVD Case (%)
Long-term illness				
No	197 440 (68.33)	50.78 (20.74)	18 268 (9.25)	11 835 (6.20)
Yes	91 513 (31.67)	48.39 (21.22)	23 510 (25.69)	8970 (11.65)
P[a]		1.53E-285	<1.0E-300	1.66E-305

25(OH)D, 25-hydroxyvitamin D; A-levels, Advanced levels; BMI, body mass index; cig, cigarette; CSE, Certificate of Secondary Education; CVD, cardiovascular disease; NVQ, National Vocational Qualification; Q, quartiles; SD, standard deviation.
[a]For serum 25(OH)D, *P*-values have been adjusted for age, sex, assessment centre, and nuisance factors, which could affect serum 25(OH)D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement; for CVD and incident CVD, *P*-values have been adjusted for age, sex, and assessment centre.
[b]Current smokers without information on types of tobacco that they smoke.

imaging phenotypes (see Supplementary material online, *Table S5*, P≥0.05).

Mendelian randomization

MR analyses suggested a non-linear association between 25(OH)D and CVD risk (*Figure 2A*). The association was 'L-shaped' with the highest CVD risk with the lowest concentrations, and levelling off at about 50 nmol/L (*Figure 2A*, $P_{non\text{-}linear}$ = 0.007). Based on the non-linear MR, individuals with serum 25(OH)D at 25 nmol/L had 11% (95% CI: 1.05–1.18) higher odds of CVD compared with those with 50 nmol/L. There appeared to be a very slight further lowering in the odds of CVD with higher concentrations, and for example participants with 75 nmol/L had 2% lower odds (95% CI: 0.97–0.99) compared with 50 nmol/L. Similar curved associations were also observed for 25(OH)D with SBP (*Figure 2B*, $P_{non\text{-}linear}$ = 0.03) and DBP (*Figure 2C*, $P_{non\text{-}linear}$ = 0.07), with individuals with 25 nmol/L estimated to have 0.70 mmHg (95% CI: 0.15–1.26) and 0.25 mmHg (95% CI: –0.02 to 0.51) higher BP compared with 50 nmol/L. There was no statistical evidence for non-linear associations in the sub-sample available for analyses on cardiac-imaging phenotypes (*n* up to 16 489), although the pattern with AI was similar to that seen for CVD and BPs (see Supplementary material online, *Figure S6* and Supplementary material online, *Table S6*, $P_{non\text{-}linear}$ ≥0.053 for all). Linear MR did not provide any support for an association between 25(OH)D and any of the outcomes (see Supplementary material online, *Tables S5* and *S6*).

Sensitivity analyses

As substantial heterogeneity of vitaminD-GS–25(OH)D association across 100 strata was detected (see Supplementary material online, *Figure S4*), we conducted a sensitivity analysis excluding local causal estimates from the outlying strata (i.e. the 1st and 100th stratum, after exclusion $P_{Cochran's\ Q}$ = 0.97; P_{trend} = 0.44), and this did not affect our findings (see Supplementary material online, *Figure S7*). Furthermore, analyses using the alternative instrument with 122 variants provided similar results (see Supplementary material online, *Figure S8*). There was no evidence that our non-linear MR analysis for CVD risk was driven by a particular variant or a block of variants (see Supplementary material online, *Tables S7* and *S8*, $P_{non\text{-}linear}$≤ 0.031 for all). The two-sample non-linear MR analyses showed a consistent L-shaped association between serum 25(OH)D and CVD risk across all methods (see Supplementary material online, *Figure S9*, $P_{non\text{-}linear}$≤0.026 for all).

Potential impact fraction

Figure 3 shows the relationship between PIF for CVD and thresholds of serum 25(OH)D for correcting vitamin D status. As expected, PIF increased as the correction threshold was raised. In line with the 25(OH)D–CVD association pattern observed in the non-linear MR analysis, the most rapid growth of PIF occurred when the correction threshold was raised from 10 nmol/L to 50 nmol/L. PIF reached 4.4% (95% CI: 1.8–7.3%) at 50 nmol/L, and further increased to 5.7% (95% CI: 2.4–9.3%) at 75 nmol/L. Little increments were seen after 100 nmol/L, with PIF at 100 nmol/L being 6.3% (95% CI: 2.7–10.2%). Insufficient data were available to allow the calculation of PIF for concentrations much higher than 125 nmol/L.

Discussion

The effect of severe vitamin D deficiency on disease risk has been an elusive target to study. Clinical trials have typically either failed or been prevented from recruiting people with deficiency, with studies left to investigate effects of intakes, which are likely to be a surplus to the actual nutritional requirement. In this study, we use a large-scale genetic design and show evidence for a causal role of vitamin D deficiency in cardiovascular health, with individuals at the lowest concentrations having increased risk for CVD and higher BP. Our findings strongly support the previously proposed threshold effect,[7] suggesting that correction of vitamin D deficiency in the affected individuals is likely to support cardiovascular health among other downstream consequences (*Graphical Abstract*).

Health effects of vitamin D on CVD risk have previously been investigated in RCTs[6] and MR framework,[13,14,29,38] mostly reporting null findings, which seemingly contradicts our findings. However, if the causal effect of vitamin D on CVD risk is truly L-shaped as reported in our study, it would have been overlooked by the existing supplementation trials and MR studies. Although a RCT is the gold-standard approach to establish causality, its validity in capturing

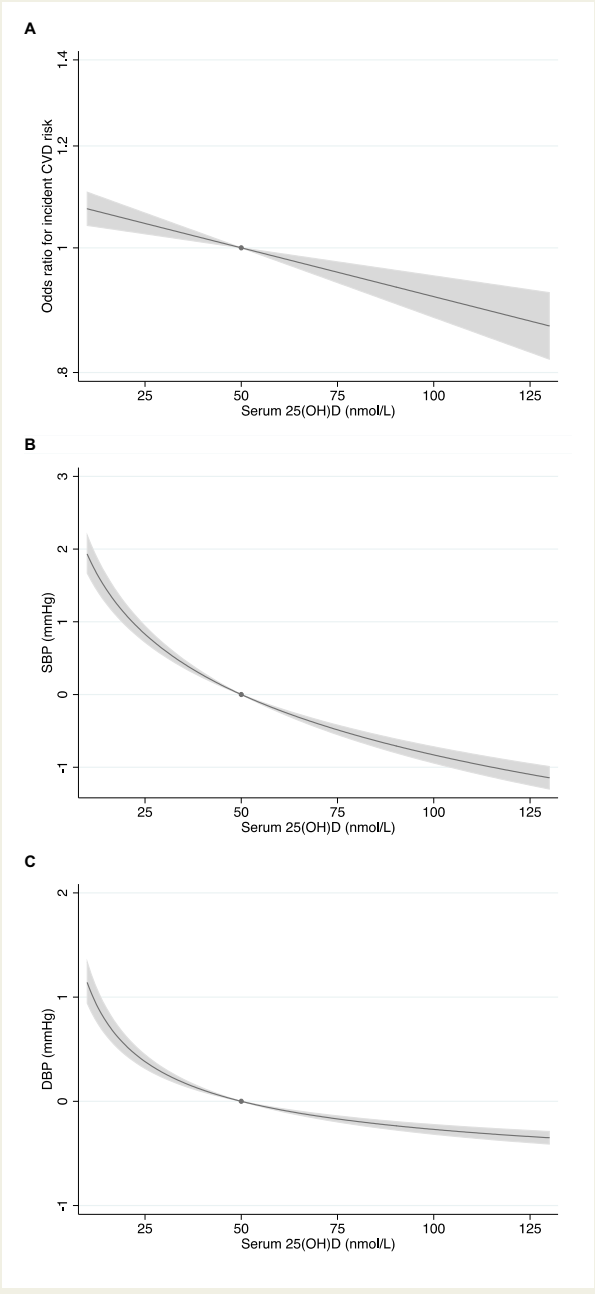


Figure 1 Phenotypic association of serum 25-hydroxyvitamin D with cardiovascular disease risk (*A*), systolic blood pressure (*B*), and diastolic blood pressure (*C*). The dot represents the reference point of serum 25-hydroxyvitamin D of 50 nmol/L. The shaded areas represent the 95% confidence intervals. Adjustment includes age, sex, body mass index, waist circumference, smoking, alcohol intake, physical activity, self-reported health status, long-term illness, assessment centre, area deprivation and education, and nuisance factors, which could affect serum 25-hydroxyvitamin D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

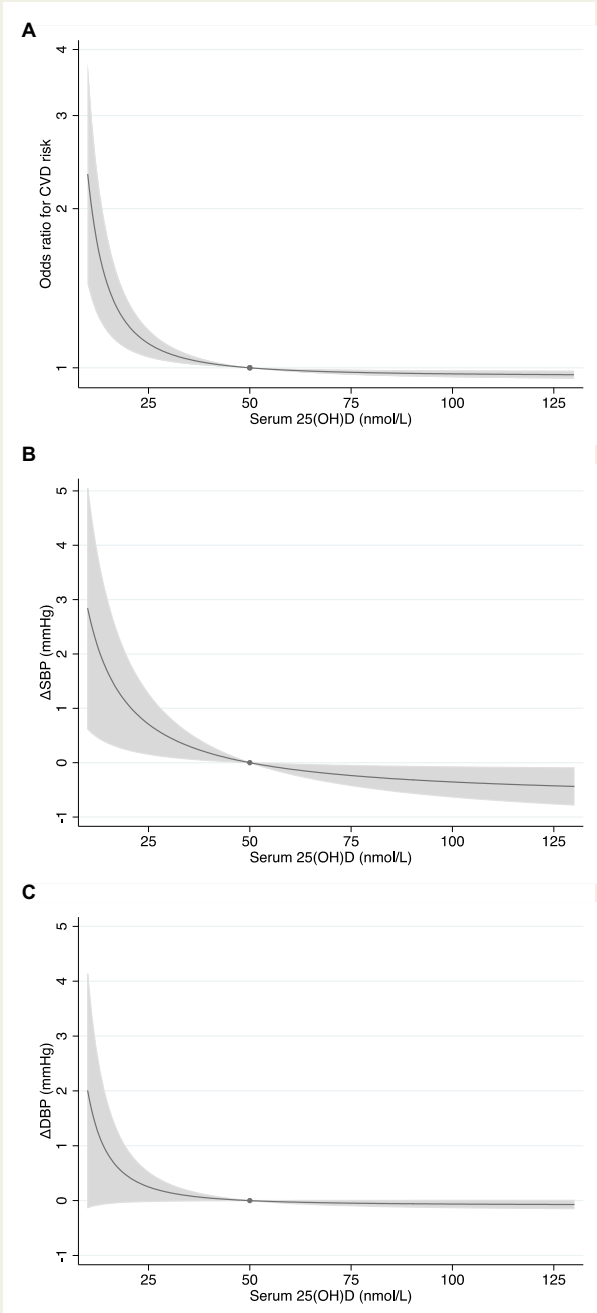


Figure 2 Genetic association of serum 25-hydroxyvitamin D with cardiovascular disease risk (*A*), systolic blood pressure (*B*), and diastolic blood pressure (*C*). The dot represents the reference point of serum 25-hydroxyvitamin D of 50 nmol/L. The shaded areas represent the 95% confidence intervals. Adjustment includes age, sex, assessment centre, birth location, single nucleotide polymorphism array, Top 40 genetic principal components, and nuisance factors which could affect serum 25-hydroxyvitamin D measurements, including month in which blood sample was taken, fasting time before blood sample was taken, and sample aliquots for measurement.

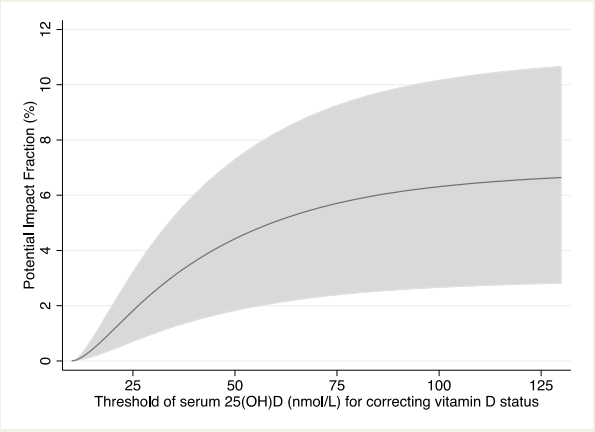


Figure 3 Potential impact fraction for cardiovascular disease in the UK Biobank at different thresholds of serum 25-hydroxyvitamin D for correcting vitamin D status. The shaded areas represent the 95% confidence intervals.

potential benefits of vitamin D supplementation has been questioned,[39] as most trials include individuals who are already vitamin D replete.[7,8] This is further compounded by contamination of the placebo group and unblinding, given that vitamin D testing and supplementation are easily accessible.[39] Furthermore, previous MR studies assessing the effect of vitamin D on CVD risk have been employed the standard MR framework, which assumes linearity and is not designed to describe or capture a non-linear association. Consistent with these earlier studies, our linear MR analysis also failed to detect any evidence for an association. Therefore, earlier studies do not discount our findings, and are compatible with the true effects of vitamin D on CVD risk, being largely restricted to individuals with the lowest concentrations.

Association of vitamin D deficiency with CVD risk is biologically plausible. The biologically active form of vitamin D is 1,25-dihydroxyvitamin D [calcitriol, 1,25$(OH)_2$D], which is formed from 25(OH)D through hydroxylation by 1-α-hydroxylase. 1,25$(OH)_2$D exerts its biological actions through binding to the vitamin D receptor (VDR).[40] Both VDR and 1-α-hydroxylase are actively expressed and functional in the cardiovascular tissues,[41–43] with evidence from animal and cell line experiments suggesting that depletion of vitamin D could adversely affect functions and behaviours of cardiomyocytes, and endothelial and vascular smooth muscle cells.[40] Furthermore, mounting evidence also suggests that calcitriol is a negative regulator of the renin–angiotensin system (RAS),[44–46] a key regulatory system in maintaining normal cardiovascular functions.[47] In VDR or 1-α-hydroxylase knock-out mice,[44–46] RAS activity is up-regulated, and these mice also subsequently develop hypertension and cardiac hypertrophy, which can be corrected by the treatment of RAS blockers[44,45] and exogenous 1,25$(OH)_2$D.[46] Association of vitamin D with RAS has also been reported in human observational studies,[48,49] although it was not observed in a recent RCT with a small sample size and a short treatment duration.[50] Furthermore, given its diverse role in multiple organs and systems,[51] calcitriol could also affect cardiovascular health indirectly via its effects on other systems, for example, immune system. Indeed, these indirect effects might explain the lack of the associations of 25(OH)D with structural and functional changes in the heart and vessels indexed by cardiac-imaging-derived phenotypes in our study. That said, our study may have been underpowered to detect related effects, as the imaging data were only available for a sub-sample, and these phenotypes may also be insensitive to modest changes. It is worthwhile to re-visit these analyses when more imaging data become available, with potential to provide more mechanistic insights to the 25(OH)D–CVD association.

While severe vitamin D deficiency is relatively rare, low concentrations are common. For example, the prevalence of serum 25(OH)D level < 50 nmol/L has been estimated to be 23% for Australia, 24% for USA, 37% for Canada, 6–76% for Europe, and 6–70% for South East Asia,[52–54] with these estimates varying by age, geographical location, and ethnicity. In the UK Biobank, 55% of participants had serum 25(OH)D concentrations < 50 nmol/L, with 13% < 25 nmol/L.[55] As participants in the UK Biobank in general are healthier than the general public,[56] the true prevalence of low vitamin D status in the UK population is likely to be higher. Given low vitamin D status is common around the world, our findings have significant public health implications. Indeed, our study implies that population-wide correction of low vitamin D status (for example, by food fortification) could be a cost-effective measure to reduce the burden of CVDs. Based on the population impact fraction calculation, we estimate that 4.4% of CVD incidence could have been prevented in this population by increasing serum 25(OH)D to be at least 50 nmol/L for all individuals. Benefits for shifting 25(OH)D concentrations even higher had less impact on CVD risk; however, we saw no evidence for harm and the predicted reduction reached 6.3% for those with 100 nmol/L. However, our study had few individuals with concentrations >100 nmol/L (1.5%), and hence extrapolation of our findings to higher concentrations needs to be done with caution.

Strengths and limitations

To our best knowledge, our study is the first genetic analysis using the non-linear MR framework to explore the shape of the association of 25(OH)D with CVD risk and to demonstrate the L-shaped association. Our findings are unlikely to be due to chance, given robust statistical evidence and the existing knowledge on the biological basis for the association. MR analyses can be biased by horizontal pleiotropy where variants influence the outcome through pathways other than that via the exposure.[12] We implemented several strategies to alleviate related issues and to gauge the robustness of our finding to horizontal pleiotropy. First, we restricted our instrument to 35 variants with robust replicated evidence for an association with 25(OH)D concentrations. Second, we found no evidence for an association between vitaminD-GS with potential confounders in the whole cohort or across stratums of residuals of serum 25(OH)D. Furthermore, neither leave-one-out nor leave-block-out analyses suggest our finding was driven by a particular variant or a group of variants. Moreover, the L-shaped association of 25(OH)D with CVD risk was highly consistent across complementary pleiotropy-robust MR methods. Despite these strengths, our study also has some limitations. We restricted our analysis to participants of White-British descent. While this minimizes the bias due to population stratification, it may also limit the transferability of our findings to other ethnic groups. As with all MR studies, genetic instruments approximate the average

effects over the life-course. The shape and strength of the true biological association between serum 25(OH)D and CVD risk could vary by life-stage and potentially be more complex than that indexed in our study. With only 5% response rate, UK Biobank is not representative of the general public in the UK[56] despite its large sample size. It is uncertain to what extent this selection could affect the non-linear MR analysis. However, given that risk factor–disease associations show close agreement between UK Biobank and nationally representative studies,[57] and that an earlier publication from the UK Biobank using the same non-linear MR approach has replicated the expected J-shape association between BMI and mortality,[16] this lack of representativeness may have little influence on our findings.

Conclusions

In conclusion, using a large prospective cohort, we provide genetic evidence for an L-shaped association of 25(OH)D with CVD risk, with increased CVD risk largely restricted to individuals with low vitamin D status. While improving vitamin D status among people with the lowest concentrations is likely to have the strongest effects, population-wide approach to eradicate vitamin D deficiency could reduce the burden of CVDs.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Funding

This study was financially supported by the National Health and Medical Research Council, Australia (Grant number 1123603). The funder had no role in the design, implementation, analysis, interpretation of the data, approval of the manuscript, and decision to submit the manuscript for publication.

Conflict of interest: none declared.

Data availability

Data are available from the UK Biobank (https://www.ukbiobank.ac.uk/) for researchers who meet the criteria and gain approvals to access the research database from the UK Biobank access management committee at the University of Oxford.

References

1. Wang TJ, Pencina MJ, Booth SL *et al.* Vitamin D deficiency and risk of cardiovascular disease. *Circulation* 2008;**117**:503–511.
2. Brøndum-Jacobsen P, Benn M, Jensen GB, Nordestgaard BG. 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**:2794–2802.
3. Chowdhury R, Kunutsor S, Vitezova A *et al.* Vitamin D and risk of cause specific death: systematic review and meta-analysis of observational cohort and randomised intervention studies. *BMJ* 2014;**348**:g1903.
4. Schöttker B, Jorde R, Peasey A *et al.*; Consortium on Health and Ageing: Network of Cohorts in Europe and the United States. Vitamin D and mortality: meta-analysis of individual participant data from a large consortium of cohort studies from Europe and the United States. *BMJ* 2014;**348**:g3656.
5. Fan X, Wang J, Song M *et al.* Vitamin D status and risk of all-cause and cause-specific mortality in a large cohort: results from the UK Biobank. *J Clin Endocrinol Metab* 2020;**105**:dgaa432.
6. Barbarawi M, Kheiri B, Zayed Y *et al.* Vitamin D supplementation and cardiovascular disease risks in more than 83000 individuals in 21 randomized clinical trials: a meta-analysis. *JAMA Cardiol* 2019;**4**:765–776.
7. Scragg R. Emerging evidence of thresholds for beneficial effects from vitamin D supplementation. *Nutrients* 2018;**10**:561.
8. Quyyumi AA, Al Mheid I. The demise of vitamin D for cardiovascular prevention. *JAMA Cardiol* 2019;**4**:776–777.
9. Frame LA, Fischer JP, Geller G, Cheskin LJ. Use of placebo in supplementation studies—vitamin D research illustrates an ethical quandary. *Nutrients* 2018;**10**:347.
10. King C, Mulugeta A, Nabi F, Walton R, Zhou A, Hyppönen E. Mendelian randomization case-control PheWAS in UK Biobank shows evidence of causality for smoking intensity in 28 distinct clinical conditions. *EClinicalMedicine* 2020;**26**:100488.
11. Millwood IY, Walters RG, Mei XW *et al.* Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet* 2019;**393**:1831–1842.
12. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;**362**:k601.
13. Huang T, Afzal S, Yu C *et al.*; China Kadoorie Biobank Collaborative Group. 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.
14. 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**:349–356.
15. Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. *Genet Epidemiol* 2017;**41**:341–352.
16. Sun YQ, Burgess S, Staley JR *et al.* Body mass index and all cause mortality in HUNT and UK Biobank studies: linear and non-linear Mendelian randomisation analyses. *BMJ* 2019;**364**:l1042.
17. Di Angelantonio E, Bhupathiraju S, Wormser D *et al.*; Global BMI Mortality Collaboration. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet* 2016;**388**:776–786.
18. Aune D, Sen A, Prasad M *et al.* BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *BMJ* 2016;**353**:i2156.
19. Bhaskaran K, Dos-Santos-Silva I, Leon DA, Douglas IJ, Smeeth L. Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3·6 million adults in the UK. *Lancet Diabetes Endocrinol* 2018;**6**:944–953.
20. 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.
21. Littlejohns TJ, Sudlow C, Allen NE, Collins R. UK Biobank: opportunities for cardiovascular research. *Eur Heart J* 2019;**40**:1158–1166.
22. Bycroft C, Freeman C, Petkova D *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;**562**:203–209.
23. Warren HR, Evangelou E, Cabrera CP *et al.*; UK Biobank CardioMetabolic Consortium BP working group. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet* 2017;**49**:403–415.
24. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med* 2005;**24**:2911–2935.
25. Petersen SE, Matthews PM, Francis JM *et al.* UK Biobank's cardiovascular magnetic resonance protocol. *J Cardiovasc Magn Reson* 2016;**18**:8.
26. Coffey S, Lewandowski AJ, Garratt S *et al.* Protocol and quality assurance for carotid imaging in 100,000 participants of UK Biobank: development and assessment. *Eur J Prev Cardiol* 2017;**24**:1799–1806.
27. Revez JA, Lin T, Qiao Z *et al.* Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. *Nat Commun* 2020;**11**:1647.
28. Jiang X, O'Reilly PF, Aschard H *et al.* Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat Commun* 2018;**9**:260.
29. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* 2013;**42**:1134–1144.
30. Townsend P, Beattie A, Phillimore P. *Health and Deprivation: Inequality and the North*. London: Routledge; 1987.
31. Haworth S, Mitchell R, Corbin L *et al.* Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nat Commun* 2019;**10**:333.
32. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol* 1999;**28**:964–974.

33. Palmer TM, Sterne JA, Harbord RM *et al.* Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. *Am J Epidemiol* 2011;**173**:1392–1403.
34. Dudbridge F, Allen RJ, Sheehan NA *et al.* Adjustment for index event bias in genome-wide association studies of subsequent events. *Nat Commun* 2019;**10**: 1561.
35. Kamat MA, Blackshaw JA, Young R *et al.* PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 2019;**35**: 4851–4853.
36. Slob EAW, Burgess S. A comparison of robust Mendelian randomization methods using summary data. *Genet Epidemiol* 2020;**44**:313–329.
37. Barendregt JJ, Veerman JL. Categorical versus continuous risk factors and the calculation of potential impact fractions. *J Epidemiol Community Health* 2010;**64**: 209–212.
38. Vimaleswaran KS, Cavadino A, Berry DJ *et al.* Association of vitamin D status with arterial blood pressure and hypertension risk: a Mendelian randomisation study. *Lancet Diabetes Endocrinol* 2014;**2**:719–729.
39. Scragg R. Limitations of vitamin D supplementation trials: why observational studies will continue to help determine the role of vitamin D in health. *J Steroid Biochem Mol Biol* 2018;**177**:6–9.
40. Al Mheid I, Quyyumi AA. Vitamin D and cardiovascular disease: controversy unresolved. *J Am Coll Cardiol* 2017;**70**:89–100.
41. Chen S, Law CS, Grigsby CL *et al.* Cardiomyocyte-specific deletion of the vitamin D receptor gene results in cardiac hypertrophy. *Circulation* 2011;**124**:1838–1847.
42. Zehnder D, Bland R, Chana RS *et al.* Synthesis of 1,25-dihydroxyvitamin D(3) by human endothelial cells is regulated by inflammatory cytokines: a novel autocrine determinant of vascular cell adhesion. *J Am Soc Nephrol* 2002;**13**:621–629.
43. Somjen D, Weisman Y, Kohen F *et al.* 25-hydroxyvitamin D3-1alpha-hydroxylase is expressed in human vascular smooth muscle cells and is upregulated by parathyroid hormone and estrogenic compounds. *Circulation* 2005;**111**:1666–1671.
44. Li YC, Kong J, Wei M, Chen ZF, Liu SQ, Cao LP. 1,25-Dihydroxyvitamin D(3) is a negative endocrine regulator of the renin-angiotensin system. *J Clin Invest* 2002; **110**:229–238.
45. Xiang W, Kong J, Chen S *et al.* Cardiac hypertrophy in vitamin D receptor knockout mice: role of the systemic and cardiac renin-angiotensin systems. *Am J Physiol Endocrinol Metab* 2005;**288**:E125–E132.
46. Zhou C, Lu F, Cao K, Xu D, Goltzman D, Miao D. Calcium-independent and 1,25(OH)2D3-dependent regulation of the renin-angiotensin system in 1alpha-hydroxylase knockout mice. *Kidney Int* 2008;**74**:170–179.
47. Wu CH, Mohammadmoradi S, Chen JZ, Sawada H, Daugherty A, Lu HS. Renin-angiotensin system and cardiovascular functions. *Arterioscler Thromb Vasc Biol* 2018;**38**:e108–e116.
48. Tomaschitz A, Pilz S, Ritz E *et al.* Independent association between 1,25-dihydroxyvitamin D, 25-hydroxyvitamin D and the renin-angiotensin system: the Ludwigshafen Risk and Cardiovascular Health (LURIC) study. *Clin Chim Acta* 2010;**411**:1354–1360.
49. Forman JP, Williams JS, Fisher ND. Plasma 25-hydroxyvitamin D and regulation of the renin-angiotensin system in humans. *Hypertension* 2010;**55**: 1283–1288.
50. McMullan CJ, Borgi L, Curhan GC, Fisher N, Forman JP. The effect of vitamin D on renin-angiotensin system activation and blood pressure: a randomized control trial. *J Hypertens* 2017;**35**:822–829.
51. Umar M, Sastry KS, Chouchane AI. Role of vitamin D beyond the skeletal function: a review of the molecular and clinical studies. *Int J Mol Sci* 2018;**19**: 1618.
52. Nimitphong H, Holick MF. Vitamin D status and sun exposure in southeast Asia. *Dermatoendocrinol* 2013;**5**:34–37.
53. Australian Government Department of Health. *Vitamin D Status*. 2021. https://www.health.gov.au/resources/pregnancy-care-guidelines/part-g-targeted-maternal-health-tests/vitamin-d-status (26 October 2021).
54. Amrein K, Scherkl M, Hoffmann M *et al.* Vitamin D deficiency 2.0: an update on the current status worldwide. *Eur J Clin Nutr* 2020;**74**:1498–1513.
55. Sutherland JP, Zhou A, Leach MJ, Hyppönen E. Differences and determinants of vitamin D deficiency among UK biobank participants: a cross-ethnic and socioeconomic study. *Clin Nutr* 2021;**40**:3436–3447.
56. Fry A, Littlejohns TJ, Sudlow C *et al.* Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol* 2017;**186**:1026–1034.
57. Batty GD, Gale CR, Kivimäki M, Deary IJ, Bell S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ* 2020;**368**:m131.