Bone 126 (2019) 59–77

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Full Length Article

The genetics of vitamin D

Xia Jiang[a,b,*], Douglas P. Kiel[c,d,e], Peter Kraft[a]

[a] Program in Genetic Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Brookline, Boston 02115, USA
[b] Unit of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institute, Nobels vagen 13, Stockholm 17177, Sweden
[c] Institute for Aging Research, Hebrew SeniorLife, 1200 Centre Street, Boston, MA 02131, United States
[d] Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA 02115, United States
[e] Broad Institute of Harvard and Massachusetts Institute of Technology, Boston, MA 02142, United States

ARTICLE INFO

Keywords:
Vitamin D
Genetics
Genome-wide association studies
Whole genome sequencing
Mendelian randomizations

ABSTRACT

Vitamin D plays an essential role in human health as it influences immune function, cell proliferation, differentiation and apoptosis. Vitamin D deficiency has been associated with numerous health outcomes, including bone disease, cancer, autoimmune disease, cardiovascular conditions and more. However, the causal role of vitamin D beyond its importance for bone health remains unclear and is under much debate. Twin and familial studies from past decades have demonstrated a nontrivial heritability of circulating vitamin D concentrations. Several large-scale genome-wide association studies (GWAS) have discovered associations of *GC*, *NADSYN1/DHCR7*, *CYP2R1*, *CYP24A1*, *SEC23A*, *AMDHD1* with serum levels of vitamin D. A recent whole genome sequencing (WGS) study, combined with deep imputation of genome-wide genotyping, has identified a low-frequency synonymous coding variant at *CYP2R1*. Information on these genetic variants can be used as tools for downstream analysis such as Mendelian randomization. Here, we review the genetic determinants of circulating vitamin D levels by focusing on new findings from GWAS and WGS, as well as results from Mendelian randomization analyses conducted so far for vitamin D with various traits and diseases. The amount of variation in vitamin D explained by genetics is still small, and the putative causal relationship between vitamin D and other diseases remains to be demonstrated.

1. Introduction

Vitamin D is an important fat soluble vitamin and steroid pro-hormone that plays a critical role in bone mineralization. It regulates the blood and extracellular concentrations of calcium and phosphorus, and ensures their adequacy for the deposition of calcium hydroxyapatite to the bone matrix. Vitamin D deficiency has long been observed to be associated with rickets in children and osteomalacia in adults [1]. In addition to musculoskeletal disorders, a variety of common diseases that are of important public health concerns, including but not limited to, cancer, autoimmune inflammatory disease, infectious disease, cardiovascular condition and diabetes, have been reported to be linked with vitamin D [2–6]. The serum concentration of vitamin D is largely influenced by environment and nutrition (geographical latitude, sun exposure of the skin, dietary pattern, and intake of supplements). It is however, also determined by genetic background, as twin studies have estimated a substantial quantity of heritability [7–9].

Understanding the genetic determinants of vitamin D could increase our knowledge to the trait itself, aid in vitamin D deficiency screening, as well as promoting personalized recommendations for the utility of supplementation. Genetic variants associated with vitamin D could also serve as instrumental variables for downstream Mendelian randomization analyses, and help understand the causal relationships between vitamin D and other traits, including bone health. Further, the shared etiology between vitamin D and other traits can be elucidated through the quantification of genome-wide genetic correlation or identification of pleotropic loci. The many recent advances and efforts in the field of genetics have already highlighted several vitamin D associated genes.

In this review, we first briefly introduce the physiology of vitamin D metabolism, and then briefly summarize earlier twin and familial studies of vitamin D. We next outline the novel findings from genome-wide association studies and whole genome sequencing studies conducted to date. Finally, we illustrate the results from Mendelian randomization studies which examine the putative causal relationship between vitamin D and other diseases.

* Corresponding author at: Program in Genetic Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Brookline, Boston 02115, USA.
E-mail address: xiajiang@hsph.harvard.edu (X. Jiang).

https://doi.org/10.1016/j.bone.2018.10.006
Received 13 August 2018; Received in revised form 6 October 2018; Accepted 7 October 2018
Available online 12 October 2018

Table 1
Twin and familial studies conducted in circulating vitamin D levels and estimates of heritability.

Author	Year	Population	Design	Sample size	Measurement	Heritability	Covariates	Statistical methods
Livshits et al.	1999	Caucasian	Familial study	355 (95 families)	25(OH)D	0.22	52.6% females. 18–91 yr.	Complex segregation analysis
Hunter et al.	2001	Caucasian	Twin study	2136 (384 MZ/684 DZ)	25(OH)D $1,25(OH)_2D$	0.43 (0.28–0.57) 0.65 (0.53–0.74)	98.3% females. 18–71 yr. Estimates were age, and BMI adjusted.	Compare intra-class correlations and ACE model
Wjst et al.	2006	Caucasian	Familial study	947 (124 families)	$25(OH)D_3$ $1,25(OH)_2D_3$	0.8 0.3	Asthma patients. Methods for statistical analysis unclear.	unclear
Orton et al.	2008	Caucasian	Twin study	198 (40 MZ/59 DZ)	25(OH)D	0.77	65.6% females. 34–78 yr. multiple sclerosis patients. Blood drawn at the end of winter. Estimates were age, and sex adjusted.	Compare intra-class correlations
Engelman et al.	2008	Hispanic, Africa America	Familial study	1530 from three studies (513, 513, 504); 130 families from three studies (42, 30, 58)	25(OH)D 25(OH)D 25(OH)D $1,25(OH)_2D$ $1,25(OH)_2D$ $1,25(OH)_2D$	0.23 ± 0.11 0.41 ± 0.10 0.28 ± 0.10 0.20 ± 0.09 0.16 ± 0.08 0.48 ± 0.10	57.6% female. Mean age 40 yr. Insulin Resistance Atherosclerosis Family Study. Estimates were adjusted for sex, age, solar radiation, BMI, and physical activity.	Residual heritability (SOLAR)
Snellman et al.	2009	Caucasian	Twin study	204	25(OH)D	0.39 (0.32–0.75)	57.8% females. Mean age 58 yr. Seasonal difference was observed: summer 0.48 (0.00–0.85), winter 0.00 (0.00–0.77).	Compare intra-class correlations and ACE model
Shea et al.	2009	Caucasian	Community-based study	1762 (264 sib-pairs)	25(OH)D	0.29 ± 0.11	52.1% female. Mean age 59 yr.	Residual heritability (SOLAR)
Arguelles et al.	2009	Asian	Twin study	226 (51 MZ/58 DZ)	25(OH)D	0.69 (0.26–0.79)	52.6% female. Adolescents mean age 16 yr. Estimates were adjusted for age, sex, physical activity, and season. Gender difference was observed: men 0.86 (0.62–0.93), women 0.59 (0.34–0.76).	Compare intra-class correlations and ACE model
Karohl et al.	2010	Caucasian	Twin study	510 (155 MZ/100 DZ)	25(OH)D 25(OH)D	0.70 (0.31–0.80) 0 (0–0.35)	100% male; mean age 55 yr. Seasonal difference was observed: winter 0.7, summer 0.	Compare intra-class correlations and ACE model
Mills et al.	2014	Caucasian	Twin study	378 (70 MZ/118 DZ)	$25(OH)D_3$	0.86 (0.61–0.94)	Adolescents mean age 16 yr.	Compare intra-class correlations and ACE model

MZ: monozygotic; DZ: dizygotic.

2. Physiology of vitamin D

There are two forms of vitamin D, vitamin D_2 (ergocalciferol), mainly acquired from the diet; and vitamin D_3 (cholecaciferol), mainly synthesized in the skin from pro-vitamin D_3 (7-dehydrocholesterol) when exposed to ultraviolet B radiation from sunlight. The cutaneously generated vitamin D_3 and the dietary ingested vitamin D_2 are then transported to the liver where they are metabolized by the hepatic enzyme 25-hydroxylase (CYP2R1) to its major circulating form, 25-hydroxyvitamin D [25(OH)D]. 25(OH)D is biologically impotent on calcium metabolism at physiological concentrations and requires a further hydroxylation by 25(OH)D-1α-hydroxylase enzyme (CYP27B1) in the kidney to produce its biologically active form 1,25-hydroxyvitamin D [1,25$(OH)_2$D]. The metabolic pathway leading to the synthesis of active vitamin D has been known for 30 years and reviewed in details by Jones et al. [10].

Both vitamin D metabolites [25(OH)D and 1,25$(OH)_2$D] are commonly measured in the serum, and served as the main targets for most genetic studies of vitamin D metabolism. Albeit 1,25$(OH)_2$D is the most active form of vitamin D and undertakes the major biological function, it has a much shorter half-life and a 1000-fold lower concentration than 25(OH)D. Hence, total-body vitamin D stores are best reflected by the circulating level of the more stable and common 25(OH)D [11].

3. Heritability of vitamin D – twin and familial studies

The extent to which observed phenotypic variance of a trait can be explained by genetic effects (heritability) can be estimated through twin studies. Using identical (monozygotic, MZ) and non-identical (dizygotic, DZ) twin pairs, twin studies partition the variance of a trait into genetic, environmental and residual variation, known as the "ACE" model, where A denotes the additive genetic variance resulted from the sum of allelic effects; C denotes shared environmental variance such as prenatal environment, home environment, socioeconomic status and residential area; and E denotes residual variance that are specific to individuals (not shared). Twin studies have several important assumptions. First, they assume "equal environment", meaning that the trait-relevant environments are similar in both MZ and DZ twin pairs. Second, it assumes that MZ twins share 100% of their genes, while DZ twins share on average 50% of their genes. Third, it only calculates additive genetic effects but not dominant or epistatic effects, nor does it take into account gene-environment interactions. Based on these assumptions, if the variance of a trait were solely influenced by genetics, a phenotypic correlation of 1.0 for MZ twins and 0.5 for DZ twins would be expected. On the other hand, if the shared environmental effects were the only source of variance, a phenotypic correlation of 1.0 for both MZ and DZ pairs would be observed. Verweij et al. [12] give a detailed review of the classical twin design.

The heritability of circulating vitamin D levels has been quantified in several twin studies, with estimates varying from < 20% to as high as above 85% [7–9,13–18]. The first twin study was conducted by Hunter et al., using 1068 twin pairs (384 MZ and 684 DZ, 98.3% females, all were European ancestry) from the TwinsUK, and identified a heritability of 43% for 25(OH)D and 65% for 1,25$(OH)_2$D. This study was also able to adjust for key confounders such as age and body mass index (BMI) [7]. Subsequent two smaller studies (sample sizes: 124 families and 99 twin pairs, respectively), conducted using individuals with co-morbidities, reported higher yet consistent estimates: Wjst et al. found a heritability of 80% for 25(OH)D among asthma patients; and Orton et al. found a heritability of 77% for 25(OH)D among multiple sclerosis patients [16,18]. Likewise, comparable figures have been reported by Karohl et al. [70% for 25(OH)D], using 255 male twin pairs (155 MZ and 100 DZ) with blood drawn in winter [9]; as well as by Mills et al. [86% for 25(OH)D], using 188 adolescent twin pairs (70 MZ and 118 DZ) at their mean age of 16 years [15].

While these numbers reflect the higher end of heritability estimates for plasma vitamin D among the Caucasian population, different estimates also have been reported by others. For example, Snellman et al. analyzed 204 twins living at north latitude 60 degree, and identified a heritability of 39% for 25(OH)D [17]. Shea et al. examined 1762 participants (of which, 265 sib-pairs) in a cross-sectional study, and reported a heritability of 29% for 25(OH)D [8]. Livshits et al. conducted a familial study consisting of 95 nuclear pedigrees and reported that 22% of the variation was accounted for by a putative major gene effect [19].

Evidence from non-European populations is relatively sparse. To date, only one twin study has been conducted among Hispanics and Africa Americans, aggregating data from California, Colorado, and Texas over 1530 individuals from 130 families, and reported heritabilities of 23%, 41% and 28% for 25(OH)D across the three geographical areas [14]. Similarly, the only available twin study of vitamin D among the Asians has examined 109 Chinese twin pairs and reported a heritability of 69% for 25(OH)D [13]. Details of the twin and familial studies in vitamin D were described in Table 1.

The considerable variation in the estimates of heritability can be explained by several reasons. First, the status of circulating vitamin D is known to be influenced by environmental factors such as age, sex and seasons [20], therefore, blood sampling in different seasons and co-morbidities are likely to affect the estimates. In addition, the magnitude and precision of the estimates are also largely influenced by sample size and statistical models. Therefore, well-powered twin studies with adequate control for confounding factors are needed to provide a firm upper bound of the genetic contribution in circulating vitamin D. Despite all the inter-study discrepancies, results from these studies collectively demonstrate a nontrivial role of genetic components in circulating vitamin D levels.

4. Genetic studies of vitamin D

4.1. Candidate gene based study

Earlier studies have explored the genetics of vitamin D through candidate gene approach, where single nucleotide polymorphisms (SNPs) or genes, usually selected based on prior knowledge to their biological functions, were assessed for the association with target disease or trait. Several genes that are closely related to vitamin D synthesis, activation and degradation have been investigated (e.g., *CYP2R1* and *CYP27B1*, two genes that are responsible for vitamin D hydroxylation; group-specific component (*GC*), a gene encodes vitamin D binding protein [21]; *VDR*, a gene encodes for vitamin D receptor [22,23]; and *CYP24A1*, a cytochrome P450 genes [24]). Studies using such an approach have been reviewed in details by Dastani et al. elsewhere [25], and will therefore not be mentioned comprehensively by the current review.

Results from candidate gene studies have been inconsistent, possibly influenced by low statistical power, ethnic differences in study populations, the characterization of phenotype, and the measurement of genotype. Further, the lack of proper multiple testing corrections increases false positive rate and makes the interpretation of results difficult. More importantly, candidate gene approach is substantially restricted by established knowledge. It is likely that genes other than those directly linked to the metabolism of vitamin D play a role but could not be selected for assessment due to a lack of prior knowledge. Similar concerns hold true for the selection of target SNPs. Despite that SNPs from promoter and exons have been selected, they may only tag the effect of a causal variant nearby instead of being the true casual variant themselves.

4.2. Genome-wide association study

The first GWAS of vitamin D was conducted in 2007 among 1012 related individuals of European ancestry, recruited to the Framingham Offspring study, consisting of children and their spouses from

Framingham Original Cohort. This small-scale GWAS measured plasma 25(OH)D on 517 individuals, and examined 70,987 SNPs after quality control. Due to limited power and limited coverage, it is no surprise that none of the SNPs has passed genome-wide significant p-threshold at 5×10^{-8}; only one suggestive signal in an intron of *LOC105377885* (chromosome 6, rs10485165) was identified [26].

Two subsequent independent meta-GWASs, both conducted on participants of European ancestry, have reported key findings [27,28]. The first meta-GWAS consisted of 4501 persons drawn from 6 cohorts with prospectively collected 25(OH)D levels for the discovery stage, and an additional 2221 individuals drawn from 3 cohorts for the replication stage. The second meta-GWAS was conducted by the SUNLIGHT consortium (Study of Underlying Genetic Determinants of Vitamin D and Highly Related Traits), aggregating data from 15 cohorts on a total of 33,996 individuals, and divided into a discovery stage of 16,125 participants, an in-silico replication stage of 9367 participants, and a de-novo replication stage of 8504 participants. Pooling together the discovery and replication samples, both meta-GWASs have confirmed strong genome-wide associations with 25(OH)D at three loci, *GC* (index SNP: rs2282679), *DHCR7/NADSYN1* (rs12785878) and *CYP2R1* (rs10741657); in addition, the SUNLIGHT meta-GWAS has identified one more locus at *CYP24A1* (rs17216707). These four susceptibility loci, all being in or near genes encoding well established steps in vitamin D synthesis, transportation and degradation, have demonstrated a crucial role for common genetic variants in the regulation of circulating 25(OH)D concentrations. They have also been served as main tools for downstream analysis such as construction of polygenic risk score and Mendelian randomization since 2010.

Earlier this year (2018), a five-fold augmented meta-GWAS was published by the SUNLIGHT consortium. This GWAS expanded the previous SUNLIGHT meta-GWAS discovery sample size from 16,125 to 79,366, as well as included two separate replication samples with 40,562 and 2195 individuals [29]. In addition to the four known loci, this larger study has yielded two novel loci harboring genome-wide significant variants at *SEC23A* (Sec23 homolog A, coat protein complex II component, rs8018720) and *AMDHD1* (amidohydrolase domain containing 1, rs10745742). Mutations on these two genes have been reported to cause craniolenticulosutural dysplasia, a disease characterized by craniofacial and skeletal malformation due to defective collagen secretion; and atypical lipomatous tumor, a cancer of connective tissues. It is also for the first time that genes outside of the vitamin D metabolism pathway have been revealed, improving our understandings to the genetic architecture of serum 25(OH)D homeostasis.

While a majority of the current GWASs in circulating vitamin D levels have been conducted in Caucasian adults, a few studies have been carried out with specific focus on Hispanics, Asian Indians, children and women [30–33]. These studies, however, were all based on small sample sizes ranging from several hundreds to thousands, and lacked proper replications, precluding significant findings. For example, a GWAS conducted in 1190 Hispanic participants has identified suggestive associations at *A2BP1* and *GPR114* for 25(OH)D, as well as at *DAB1* and *MLL3* for $1,25(OH)_2D$, none reached genome-wide significance in the replication sample [30]. Similarly, a GWAS comprised of 1387 Indian subjects followed by 2151 replication samples has identified 24 putative SNPs in the discovery stage, yet only 1 novel locus at chromosome 20 (*FOXA2/SSTR4*) has been replicated [32]. For GWAS restricted to Caucasian women and children, no genomic regions other than those previous reported loci has been identified, probably due to limited power [31,33]. Large-scale collaborations are warranted to elucidate the genetic basis of circulating vitamin D for underrepresented populations such as minorities, women and children, whom might be benefited to a greater extent in terms of risk predictions and personalized dietary recommendations. A summary of the hitherto genome-wide scans performed in circulating vitamin D levels was shown in Table 2.

Despite the enormous progress in GWAS of vitamin D, common genome-wide SNPs (with minor allele frequency > 5%) have collectively only accounted for a small proportion of phenotypic variance. Jiang et al. has calculated the heritability using SNPs all over the genome, and reported a SNP-heritability for 25(OH)D of 7.5% [29], an estimate far lower than the traditional twin and familial studies. Other than the possibility that estimates from twin studies are inflated, this gap may reflect the proportion of heritability explained by rare SNPs or structural variants that have not been tagged by GWAS data, as well as gene-gene interactions that remain to be identified. We anticipate that additional studies involving large global collaborative efforts, such as the release of UK Biobank GWAS data on more than half a million individuals in the near future, would help solve part of this missing heritability issue.

4.3. Whole genome sequencing

To the best of our knowledge, to date, only one study has examined the effect of low-frequency (MAF < 5%) and rare variants (MAF < 1%) in 25(OH)D, using whole-genome sequencing data from 2619 individuals through the UK10K program, and deep-imputation data from 39,655 individuals with genome-wide genotypes [34]. Meta-analysis across the 19 participating cohorts has identified a low-frequency synonymous coding variant at *CYP2R1* (MAF = 2.5%, rs117913124). This SNP confers a large effect on 25(OH)D levels, which is four times stronger and independent of the effect of a previous reported common SNP at *CYP2R1* (rs10741657). This SNP has also been found to increase the risk of vitamin D insufficiency by 2-folds, a magnitude that achieves clinically relevant degree. These results, together with previous GWASs findings, are consistent with an oligogenetic architecture for vitamin D, that is, vitamin D is influenced by a few genetic variants, some of which have comparatively large effects, thus indicating an oligo-genic feature [35].

However, we caution that the GWAS and sequencing studies are relatively small, and more variants may yet be found in future studies. Moreover, in the one WGS study to date, the subjects in the UK10K sample were sequenced to a relatively low read depth of 6.7×, and imputation was performed using older reference panels. We expect that the power to capture very rare variants could be enhanced when large-scale, high read depth sequencing studies could be conducted in the future.

5. Mendelian randomization

The role of vitamin D beyond its importance for bone health is under much debate. Accumulating evidence for an association between vitamin D deficiency and various health related outcomes has been revealed from epidemiological investigations. The observational nature of these studies, however, hinders causal interpretation, as the validity of results could be plagued by measurement error, confounding and reverse causality (i.e. if 25(OH)D is measured close to when preclinical changes of diseases have already started). Although the effect of circulating 25(OH)D on disease risk can be demonstrated by traditional randomized controlled trials (RCT), large-scale RCTs of vitamin D supplementation are not currently widely available due to their high cost and long duration. For example, so far there are only two comprehensive ongoing trails involving both men and women, the “VITAL” RCT launched in 2010 that enrolled 25,871 participants [36] and the “D-Health” RCT launched in 2014 that recruited 21,315 individuals [37]. Both RCTs will be followed by years before they can provide any evidence on the role of vitamin D in health outcomes (VITAL is about to be completed later 2018, and D-Health is still ongoing). Yet, both RCTs are likely to be underpowered given the relatively low incidence of certain endpoint phenotypes such as cancers.

The current genetic discoveries in 25(OH)D may help disentangle the causal relationships between vitamin D and other traits through

Table 2
Genome-wide association studies conducted in circulating vitamin D levels and loci identified.

Author	Year	Population	Phenotype	Discovery sample size	Replication sample size	Discovery results	Replication results
Benjamin et al.	2007	Caucasian	25(OH)D	1012	NA	Suggestive loci ($p = 1.4 \times 10^{-6}$): rs10485165 at chromosome 6.	
Ahn et al.	2010	Caucasian	25(OH)D	4501	A separate replication sample of 2221 individuals	GWAS-significant loci ($p < 5 \times 10^{-8}$): rs2282679 in GC. Suggestive signals ($p < 5 \times 10^{-5}$): rs3829251 in NADSYN1 (in high LD with rs1790349 in DHCR7); rs6599638 in the region harboring C10orf88 in the vicinity of ACADSB; and rs2060793 in CYP2R1.	Strong genome-wide significant associations with 25(OH)D were confirmed through meta-GWAS pooling together the discovery and replication data, for loci **GC**, **NADSYN1/DHCR7**, and **CYP2R1**.
Wang et al.	2010	Caucasian	25(OH)D	16,125	Two separate replication samples: an in-silico sample of 9367 individuals, and a de-novo sample of 8504 individuals	GWAS-significant loci ($p < 5 \times 10^{-8}$): rs2282679 GC, rs12785878 NADSYN1/DHCR7, rs10741657 CYP2R1; Selected candidate genes: VDR, CYP27B1, CYP24A1, CYP27A1.	Genome-wide significant associations at three loci (**GC**, **NADSYN1/DHCR7**, and **CYP2R1**) were confirmed in replication samples; In addition, **CYP24A1** showed genome-wide significant association in the replication samples.
Engelman et al.	2010	Hispanic	25(OH)D, $1,25(OH)_2D$	229	1190 (including those 229 individuals in the discovery sample)	GWAS-significant loci ($p < 5 \times 10^{-8}$): DAB1 (rs6680429) for $1,25(OH)_2D$; Suggestive loci ($p < 5 \times 10^{-4}$): rs2806508, rs10141935, rs4778359, rs1507023 (A2BP1), and rs9937918 (GPR114) for 25(OH)D; rs1348864, rs4559029, rs12667374, rs7781309 (MLL3), rs10505337, rs2486443, and rs2154175 for $1,25(OH)_2D$.	None of these loci reached genome-wide significance in the replication samples.
Andersson et al.	2014	Caucasian; children at age 6 and age 14	25(OH)D	Age 6: 673; age 14: 1140	NA	GWAS-significant loci ($p < 5 \times 10^{-8}$): GC for both the age 6 years (rs17467825) and the age 14 years (rs1155563); CYP2R1 for both the age 6 years (rs1007392) and the age 14 years (rs11023332).	NA
Sapkota et al.	2016	Asian Indian	25(OH)D	1387	A separate replication sample of 2151 individuals	Suggestive loci ($p < 10^{-4}$): 24 SNPs. FOXA2/SSTR4 (rs6048371, rs2207173), IVL (rs11586313), GLIS1 (rs12744257), ABCA4 (rs3827713), HMCN1 (rs16824427, rs76632730), HHAT (rs67585440, rs17260466), STAB1 (rs56960576), NT5DC2 (rs80140864, rs35920544, rs74690188), CD180 (rs2230524, rs2230520), ZKSCAN8 (rs112243146), OR12D3/OR12D2 (rs114026340, rs238874), HLA-DQA2 (rs2051600), N4BP2L1 (rs3752448), SLC25A21 (rs1367032), SLC24A3 (rs3748481), FHIT (rs453951), C1orf87 (rs683311), HIST1H2AE (rs11753131).	A robust novel association signal at the intergenic region **FOXA2/SSTR4** was identified, through the meta-analysis pooling the discovery and replication data.
Manousaki et al.	2017	Caucasian	25(OH)D	Whole-genome sequencing data on 2619 individuals and deep-imputation data from 39,655 individuals	Identified single-nucleotide variants were replicated in 8711 individuals of the effect on vitamin D insufficiency	GWAS-significant loci ($p < 5 \times 10^{-8}$): a signal at the chromosomal locus 11p.15.2 (CYP2R1), rs117913124, rs116970203, rs117361591, rs117621176, rs142830933, rs117672174.	rs117913124 at **CYP2R1** was strongly associated with an increased risk of vitamin D insufficiency ($p = 1.2 \times 10^{-12}$).
Jiang et al.	2018	Caucasian	25(OH)D	79,366	Two separate replication samples with 40,562 individuals, and 2195 individuals, respectively	GWAS-significant loci ($p < 5 \times 10^{-8}$): GC (rs2282679), NADSYN1/DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs17216707), AMDHD1 (rs10745742), and SEC23A (rs8018720).	In addition to the previous GWAS-identified loci at **GC**, **NADSYN1/DHCR7**, **CYP2R1**, and **CYP24A1**, meta-GWAS using both the discovery and replication data confirmed two genome-wide significant associations at loci **AMDHD1**, and **SEC23A**.

(*continued on next page*)

Table 2 (*continued*)

Author	Year	Population	Phenotype	Discovery sample size	Replication sample size	Discovery results	Replication results
O'Brien et al.	2018	80% Caucasian; women only	25(OH)D	1829	A separate replication sample of 1534 individuals	GWAS-significant loci ($p < 5 \times 10^{-8}$): GC (rs4588, rs2282679, rs1155563, rs705120, rs2201124, rs4694105, rs1526692); Suggestive signals ($p < 5 \times 10^{-5}$): CYP2R1 (rs12794714).	Strong genome-wide significant associations with 25(OH)D were confirmed through meta-GWAS pooling together the discovery and replication datasets, for loci **GC**, and **CYP2R1**.

Red bold: genome-wide significant loci with successful replication.

Mendelian Randomization (MR), an approach that uses genetic variants as instrumental variables (IV) for assessing the causal effect of a risk factor on an outcome from observational data [38]. Under certain assumptions, an un-confounded estimate of the causal effect of an exposure on an outcome can be made using the observed IV-exposure and IV-outcome associations [39]. Although the assumptions underlying the validity of MR estimates are often unverifiable, a series of recent papers have proposed sensitivity analyses to test the robustness of MR results when the assumptions fail [40].

Despite the well-established role of vitamin D deficiency in bone health, current MR analyses have not provided any evidence for a genetically predicted level of 25(OH)D to be associated with neither bone mineral density (BMD) nor bone metabolism markers [41–43]. One of the largest studies, used GWAS summary statistics based on 32,965 individuals from the Genetic Factors for Osteoporosis Consortium, and 142,487 individuals from the UK Biobank, and used the previous identified 4 loci as IVs, found that genetically predicted 1 standard deviation increment of 25(OH)D was not associated with higher femoral neck BMD (change per SD = 0.02, p = 0.37) or lumber spine BMD (0.02, p = 0.49), and only suggestively with estimated BMD (−0.03, p = 0.02), which did not pass multiple correction [42]. It is argued by the authors that the null findings are probably due to the inclusion of generally health-conscious and healthy adults, among whom, the prevalence of vitamin D deficiency is low; and accurate cutoffs for vitamin D deficiency should be defined. Similarly, a recent published MR leveraging GWAS summary-level data on a discovery set of 37,857 fracture cases and 227,116 controls, with replication in up to 147,200 fracture cases and 150,085 controls, did not find genetically predicted vitamin D levels to be linearly associated with increased fracture risk (OR per SD decrease = 0.84 (0.70–1.02), p = 0.07) [44]. For a detailed and in-depth overview of MR studies in the bone field, please read the review written by Susanna Larsson (published in the same issue).

Similarly, in cancer, only one MR study, involving GWAS summary statistics based on 10,065 ovarian cancer cases and 21,654 controls, has reported a potential causal effect of vitamin D. The odds ratio for epithelial ovarian cancer risk was estimated to be 1.27 (1.06–1.51) per 20 nmol/L decrease in 25(OH)D concentration, and for high-grade serous subtype the risk was 1.54 (1.19–2.01) [45]. These results are in line with a Danish MR study which demonstrated a causal effect of low circulating 25(OH)D with cancer-specific mortality at a borderline significance [1.03 (1.00–1.06)] [46]. However, in that Danish study, only two loci (DHCR7 and CYP2R1) instead of the four known loci were selected. In addition, none of the MR studies conducted in other cancer types (with comparable sample sizes as the aforementioned ovarian cancer) is able to detect any significant effect of 25(OH)D, including breast, prostate, colorectal, lung, pancreatic, oral, glioma, neuroblastoma and non-melanoma skin cancer [47–53]. These results suggest that although a very small causal effect of 25(OH)D on malignant diseases cannot be ruled out, the potential benefits of 25(OH)D on lowering cancer risk might be limited.

Likewise, the causal link of vitamin D with metabolic traits or diseases remains also unclear. To date, only one MR study has reported that each 10% increase in genetically instrumented 25(OH)D was associated with a change of 0.029 mmHg in diastolic blood pressure, and an 8.1% decreased OR of hypertension [0.92 (0.87–0.97)] [54]. Another MR study has found a 50% decrease in genetically predicted 25(OH)D to be associated with a 6.0% (p = 0.001) lower high-density lipoprotein levels [55]. No obvious effect of genetically predicted vitamin D level has been found for type 2 diabetes [56,57], BMI [58], cardiovascular diseases [59–61] or other metabolites [62–64].

A more solid causal role for vitamin D has been found in multiple sclerosis (MS), an autoimmune neurodegenerative disorder. Three independent MR studies have consistently identified an association between genetically predicated 25(OH)D and MS. The first study conducted by Mokry et al. used the four SUNLIGHT GWAS-identified SNPs

Table 3
Mendelian randomization analysis conducted to investigate a putative causal role of circulating vitamin D in various disease endpoints.

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
Bone health										
Li et al.	2016	Asian	One-sample MR	1824	4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)	Serum 25(OH)D concentrations	Bone mineral density, and bone metabolism markers	None of the two stage least squares models provided evidence for associations between serum 25(OH)D and either BMD or bone metabolism markers (all $p > 0.05$).	Two-stage least square multiple instrument approach	NS
Larsson et al.	2018	Caucasian	Two-sample MR, using GWAS summary level data	GEFOS: 32,965: UK Biobank: 142,487	GWAS-reported effect sizes of IV-25(OH)D including 5 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657, or rs117913124), CYP24A1 (rs6013897), CYP19A1 (rs727479) and GC (rs2282679)		Femoral neck and lumbar spine bone mineral density; estimated bone mineral density. The effect sizes of IV-outcome were extracted from GWASs conducted by the Genetic Factors for Osteoporosis (GEFOS) Consortium, and the UK Biobank	Genetically predicted 1 standard deviation increment of serum 25(OH)D was not associated with higher femoral neck BMD (SD change in BMD 0.02; −0.03 to 0.07; $p = 0.37$), lumbar spine BMD (SD change in BMD 0.02; −0.04 to 0.08; $p = 0.49$), or estimated BMD (g/cm^2 change in BMD −0.03; −0.05 to −0.01; $p = 0.02$).	Two-sample MR; An inverse variance weighted analysis, weighted median, and MR-Egger methods	NS
Trajanoska et al.	2018	Caucasian	Two-sample MR, using GWAS summary level data	A discovery set of 37,857 fracture cases and 227,116 controls; with replication in up to 147,200 fracture cases and 150,085 controls.	4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Fracture	Vitamin D levels assessed by use of 25-hydroxyvitamin D variants were not found to be linearly associated with increased fracture risk (OR per SD decrease = 0.84, 95% CI: 0.70 to 1.02, $p = 0.07$).	Two-sample MR; An inverse variance weighted analysis, and MR-Egger methods	NS
Dudding et al.	2015	Caucasian	One-sample MR	5545	An unweighted genetic risk score constructed by using 3 SNPs in GC (rs2282679), CYP2R1 (rs10741657) and NADSYN1 (rs7944926)	Serum 25(OH)D concentrations	Caries experience; dental general anesthetic; caries onset	The odds ratio (OR) for caries experience per 10 nmol/L increase in 25(OH)D was 0.93 (0.83–1.05; $p = 0.26$) and the OR for dental general anesthetic per 10 nmol/L increase in 25(OH)D was 0.96 (0.75–1.22; $p = 0.72$).	Logistic structural mean models	NS
Malignancy										
Theodoratou et al.	2012	Caucasian	One-sample MR	4238 (2001 cases and 2237 controls)	4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679); and an allele score	Plasma 25(OH)D concentrations	Colorectal cancer	Using an allele score that combined all 4 SNPs as the IV, the estimated causal effect was OR 1.16 (0.60–2.23), while it was 0.94 (0.46–1.91) and 0.93 (0.53–1.63) when using an upstream (rs12785878, rs10741657) and a downstream allele score (rs2282679, rs6013897).	Two-stage MR, control function instrumental variable estimator	NS
Trummer et al.	2015	Caucasian	One-sample MR	702	1 SNP, rs2282679 in GC		Prostate cancer prognosis: biochemical recurrence, development of metastases, and overall survival	GC rs2282679 were not associated with biochemical recurrence (hazard ratios 0.91, 0.73–1.12; $p = 0.36$), development of metastases (1.20, 0.88–1.63; $p = 0.25$) or overall survival (1.10, 0.84–1.43; $p = 0.50$).	A simple Cox regression of outcome ~ SNP	NS
Ong et al.	2016	Caucasian	Two-sample MR, using GWAS summary level data	31,719 (10,065 cases and 21,654 controls)	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Overall ovarian cancer, and high grade serous ovarian cancer. The effect sizes of IV-outcome were extracted from a GWAS conducted by the Ovarian	The OR for epithelial ovarian cancer risk estimated by combining the individual SNP associations using inverse variance weighting was 1.27 (1.06–1.51) per 20 nmol/L decrease in 25(OH)D concentration. The estimated OR	Wald-type ratio estimator. Inverse variance weighted analysis.	[b]

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
							Cancer Association Consortium	for high-grade serous epithelial ovarian cancer was 1.54 (1.19–2.01).		
Wang et al.	2017	African American	Two-sample MR, using GWAS summary level data	3686 (1657 cases and 2029 controls)	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Breast cancer. The effect sizes of IV-outcome were calculated by using data from the Root consortium	A decrease in predicted 25(OH)D levels by one SD on the natural log scale was not associated with breast cancer (OR = 1.04, 0.97–1.11, p = 0.23).	Methods unclear, most probably through inverse variance weighted analysis	NS
Dimitrakopoulou et al.	2017	Caucasian	Two-sample MR, using GWAS summary level data	70,563 cancer cases (22,898 prostate, 15,748 breast, 12,537 lung, 11,488 colorectal, 4369 ovarian, 1896 pancreatic, and 1627 neuroblastoma) and 84,418 controls	Literature-reported effect sizes of IV-25(OH)D, including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Seven cancers. Effect sizes of the IV-outcome were extracted from the meta-GWAS conducted for each cancer by the relevant consortia (GAME-ON, GECCO, CORECT, ELLIPSE, PRACTICAL, FOCI-OCAC, TRICL-ILCCO, PANSCAN1)	There was little evidence that the multi-polymorphism score of 25(OH)D was associated with risk of any of the seven cancers or their subtypes. Specifically, the ORs per 25 nmol/L increase in genetically determined 25(OH)D concentrations were 0.92 (0.76–1.10) for colorectal cancer; 1.05 (0.89–1.24) for breast cancer; 0.89 (0.77–1.02) for prostate cancer; and 1.03 (0.87–1.23) for lung cancer.	Two-sample MR; An inverse variance weighted average of associations for specific SNPs and a likelihood based method	NS
Winslow et al.	2017	Caucasian	One-sample MR	103,084	DHCR7 (rs7944926, rs11234027) and CYP2R1 (rs10741657, rs12794714)	20 nmol/L higher plasma 25(OH)D	Non-melanoma skin cancer	A 20 nmol/L higher genetically determined plasma 25(OH)D was associated with adjusted OR of 1.11 (0.91–1.35) for developing non-melanoma skin cancer.	Wald-type ratio estimator for coefficients, and delta methods for standard errors	NS
Dudding et al.	2018	Caucasian	Two-sample MR, using GWAS summary level data	5133 cases and 5984 controls	5 SNPs in GC (rs4588), CYP2R1 (rs10741657), CYP24A1 (rs6013897), DHCR7 (rs4423214), and CYP2R1 (rs116970203). The effect sizes of IV-exposure were extracted from an existing GWAS		Oral cancer. The effect sizes of IV-outcome were calculated using data from GAME-ON consortium, as well as the UK Biobank	There was little evidence of a causal association between circulating 25(OH)D and oral cancer (OR = 0.86(0.68–1.09), p = 0.22); oropharyngeal cancer (OR = 1.28(0.72–2.26), p = 0.40); or when sites were combined (1.01(0.74–1.40), p = 0.93). Replication in UK Biobank and pooled estimates produced similar results.	Two-sample MR, inverse variance weighted analysis	NS
Takahashi et al.	2018	Caucasian	Two-sample MR, using GWAS summary level data	30,657 (12,488 cases and 18,169 controls)	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Glioma. The effect sizes of IV-outcome were extracted from a recent meta-GWAS of European individuals comprising 8 datasets.	A non-significant association between 25(OH)D levels and glioma risk was shown using both the IVW (OR = 1.21 (0.90–1.62), p = 0.20) and MLE (OR = 1.20 (0.98–1.48), p = 0.08) methods. In an exploratory analysis of tumor subtype, an inverse relationship between 25(OH)D levels and glioblastoma (GBM) risk was identified using the MLE method (OR = 0.62 (0.43–0.89), p = 0.01), but not the IVW method (OR = 0.62 (0.37–1.04), p = 0.07). No statistically significant association was shown	Inverse variance weighted analysis, maximum likelihood estimation	NS

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
								between 25(OH)D levels and non-GBM glioma.		
Sun et al.	2018	Caucasian	One-sample MR	54,580 (676 incident lung cancer)	rs2282679, rs12785878 and rs10741657	10% increase genetically determined 25(OH)D level	Lung cancer and its subtypes	There was no association between the allele score and lung cancer incidence overall, with HR 0.99 (95% CI: 0.93–1.06) per allele score. A 25 nmol/L increase in genetically determined 25(OH)D level was not associated with the incidence of lung cancer overall (MR estimate HR 0.96, 95% CI : 0.54–1.69) or any histologic type.	A genetic risk score approach; Inverse-variance weighted and median-based methods.	NS
Mortality										
Trummer et al.	2013	Caucasian	One-sample MR	3316	3 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), and GC (rs2282679); and a weighted genetic score	Plasma 25(OH)D concentrations	All-cause mortality; cause specific mortality	In a multivariate Cox regression adjusted for classical risk factors, GC, CYP2R1, and DHCR7 genotypes were not associated with all-cause mortality, cardiovascular mortality, or non-cardiovascular mortality. The mortality hazard ratio per 2.5 nmol/L genetically determined 25(OH)D concentration was 1.01 (0.96–1.07; p = 0.59).	Genetic risk score approach; outcome ~ vitamin D genetic risk score, through Cox regression	NS
Afzal et al.	2014	Caucasian	One-sample MR	95,766	DHCR7 (rs7944926, rs11234027) and CYP2R1 (rs10741657, rs12794714); unweighted allele counts (0–4)	20 nmol/L lower plasma concentration of 25(OH)D	All-cause mortality; cause specific mortality	Genetically low 25(OH)D concentrations were associated with increased all-cause mortality, cancer mortality, and other mortality but not with increased cardiovascular mortality. The hazard ratio per one DHCR7/CYP2R1 allele score increase was 1.02 (1.00–1.03; p = 0.03) for all-cause mortality, 0.98 (0.96–1.01) for cardiovascular mortality, 1.03 (1.00–1.06; p = 0.04) for cancer mortality, and 1.03 (1.00–1.06; p = 0.07) for other mortality.	Wald-type ratio estimator for coefficients, and delta methods for standard errors	[b]
Autoimmune										
Hysinger et al.	2016	A mixture of Caucasian and non-Caucasian	One-sample MR	1388	A vitamin D propensity score constructed by using 2 SNPs at gene CYP2R1 (rs10741657) and GC (rs2282679)		Asthma; severe asthma	Vitamin D propensity score was not associated with asthma (p = 0.85) or severe asthma (p = 0.86).	A genetic risk score approach, outcome ~ vitamin D propensity score, through Logistic regression	NS
Manousaki et al.	2017	Caucasian	Two-sample MR, using GWAS summary level data	Asthma: 146,761 (25,471 cases and 121,290 controls); Atopic dermatitis (10,788 cases and 30,047 controls): 40,835; Serum IgE levels: 12,853 (5888 cases and 6965 controls)	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Asthma (overall, childhood-onset); elevated immunoglobulin E levels; atopic dermatitis. The effect sizes of IV-outcome were extracted from GWASs conducted by GABRIEL consortium, and EAGLE eczema consortium	The OR per SD decrease in log-transformed 25(OH)D was 1.03 (0.90–1.19, p = 0.63) for asthma; 0.95 (0.69–1.31, p = 0.76) for childhood-onset asthma; and 1.12 (0.92–1.37, p = 0.27) for atopic dermatitis; and the effect size on log-transformed IgE levels was −0.40 (−1.65–0.85, p = 0.54).	Inverse variance weighted analysis	NS
Mao et al.	2017	Caucasian								NS

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
			Two-sample MR, using GWAS summary level data	26,475 (10,365 cases and 16,110 controls)	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Asthma. The effect sizes of IV-outcome were extracted from a meta-GWAS conducted by GABRIEL consortium	None of the SNPs were significantly associated with asthma (OR 1.00 (0.99–1.01), p = 0.45 for inverse variance weighting method; 1.00 (0.99–1.01), p = 0.65 for weighted median method; 1.00 (0.93–1.08), p = 0.98 for MR-Egger; 1.00 (0.94–1.06), p = 0.99 for robust MR-Egger).	Inverse variance weighted analysis, weighted median methods, MR-Egger regression, and robust MR-Egger regression	
Mokry et al.	2015	Caucasian	Two-sample MR, using GWAS summary level data	38,589 (14,498 cases and 24,091 controls)	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Multiple sclerosis. The effect sizes of IV-outcome were extracted from a GWAS conducted by the International Multiple Sclerosis Genetic Consortium	Each genetically determined 1-SD decrease in log transformed 25(OH)D level conferred a 2.0-fold increase in the odds of MS (1.7–2.5; $p = 7.7 \times 10^{-12}$).	Inverse variance weighted analysis	[b]
Rhead et al.	2016	Caucasian	One-sample MR	A US population: 10,071 (1056 cases and 9015 controls); a Swedish population: 12,097 (6335 cases and 5762 controls)	3 SNPs in GC (rs2282679), CYP2R1 (rs2060793 or rs10741657), and NADSYN1/DHCR7 (rs3829251); and a genetic risk score		Multiple sclerosis risk; age at onset; disease severity	Increasing levels of 25(OH)D are associated with a decreased risk of MS in both populations. In the Northern California, the OR was 0.79 (0.64–0.99, p = 0.04). In the Swedish population, the OR was 0.86 (0.76–0.98, p = 0.03). A meta-analysis of the two populations gave a combined OR of 0.85 (0.76–0.94, p = 0.003). No association was observed for age at onset or disease severity.	A genetic risk score approach; outcome ~ vitamin D genetic risk score, through Logistic regression	[b]
Gianfrancesco et al.	2017	Caucasian	One-sample MR	A US population: 11,269 (394 cases and 10,875 controls); a Swedish population: 5551 (175 cases and 5376 controls)	A genetic risk score constructed by using 3 SNPs in GC (rs2282679), CYP2R1 (rs2060793), and NADSYN1 (rs3829251)		Pediatric onset of multiple sclerosis	A vitamin D GRS associated with increasing levels of 25(OH)D in serum decreased the odds of pediatric-onset MS (OR = 0.72 (0.55–0.94); p = 0.02) after controlling for sex, genetic ancestry, HLADRB1*15:01, and over 100 non-human leukocyte antigen MS risk variants.	Genetic risk score approach; outcome ~ vitamin D genetic risk score, through logistic regression	[b]
Viatte et al.	2013	Caucasian	One-sample MR	NOAR: 1433; ERAS: 443	20 SNPs at CYP24A1 (rs6013897, rs17217119); CYP2R1 (rs10500804, rs1993116, rs10741657, rs7116978, rs12794714); DHCR7 (rs12785878, rs12800438, rs4944076, rs4944997, rs7944926, rs4945008); NADSYN1 (rs3829251); GC (rs2298850, rs3755967, rs2282679, rs1155563, rs17467825, rs7041)		Rheumatoid arthritis outcomes: Larsen score, and presence of erosions	Only DHCR7 (rs4944997 and its perfect proxy rs4944076) showed a consistent effect size over the two populations. However, the estimates did not achieve the Bonferroni corrected level of significance after meta-analysis using a fixed effects model (p-threshold < 0.05/22).	Simple regression (generalized estimating equation model, generalized linear latent and mixed model), each outcome ~ each SNP	NS
Yarwood et al.	2013	Caucasian	One-sample MR	1396	6 SNPs in CYP2R1 (rs10741657), NADSYN1 (rs3829251), DHCR7		Rheumatoid arthritis anti-TNF treatment response: DAS28, EULAR response criteria	One SNP showed modest association with absolute change in DAS28 (rs10741657, p = 0.04).	Simple regression (multivariate regression, logistic	NS

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
					(rs1790349, rs12785878), GC (rs1155563, rs7041)			However, the result did not survive multiple correction.	regression), each outcome ~ each SNP	
Bae et al.	2018	Caucasian	Two-sample MR, using GWAS summary level data	RA: 4564; SLE: 41,282	3 SNPs in SSTR4 (rs2207173), GC (rs2282679), NADSYN1 (rs3829251)		Rheumatoid arthritis, systemic lupus erythematosus	No evidence to support a causal association between vitamin D level and risk of SLE (beta = 0.032, SE = 0.119, p = 0.789) or RA (beta = 0.026, SE = 0.061, p = 0.664).	Inverse variance weighted analysis, weighted median, MR-Egger	NS
Lund-Nielsen et al.	2018	Caucasian	One-sample MR	120,013 individuals, among which, 653 CS and 1265 UC; 408,455 individuals from UK biobank, 1707 CD and 3147 UC.	rs7944926 and rs11234027 in DHCR7 and rs10741657 and rs12794714 in CYP2R1	10 nmol/L higher plasma 25(OH)D	Crohn's disease (CD) and ulcerative colitis (UC)	The multivariable adjusted hazard ratios for 10 nmol/L higher 25-hydroxyvitamin D were 1.04 (95% CI: 0.93–1.16) for CD and 1.13 (1.06–1.21) for UC. A combined 25-hydroxyvitamin D allele score was associated with a 1.4 nmol/L increase in plasma 25-hydroxyvitamin D and hazard ratios of 0.98 (0.94–1.03) for CD and 1.01 (0.97–1.05) for UC.	A combined unweighted allele score as the sum of the number of plasma 25(OH)D increasing alleles (0–8)	NS
Metabolic										
Afzal et al.	2014	Caucasian	One-sample MR	96,423	DHCR7 (rs7944926, rs11234027) and CYP2R1 (rs10741657, rs12794714); unweighted allele counts (0–4)	20 nmol/L lower plasma concentration of 25(OH)D	Type 2 diabetes	Genetic variants associated with low plasma 25(OH)D concentrations are associated with type 2 diabetes, although the results were not significant. The OR for a genetically determined 20 nmol/L lower plasma 25(OH)D concentration via endogenous production (DHCR7) was 1.51 (0.98–2.33) and via liver conversion (CYP2R1) was 1.02 (0.75–1.37).	Wald-type ratio estimator for coefficients, and delta methods for standard errors	NS
Ye et al.	2015	Caucasian	Two-sample MR, using GWAS summary level data	28,144 cases and 76,344 non-cases	4 SNPs as primary genetic instruments: rs12785878 near DHCR7, rs10741657 near CYP2R1, rs4588 of DBP, and rs17217119 near CYP24A1. The effect sizes of IV-exposure were derived using data from the Ely, and the EPIC-Norfolk studies		Type 2 diabetes; concentrations of fasting glucose; 2-h glucose; fasting insulin; and HbA1c. Associations of SNPs with risk of type 2 diabetes were based on a case-cohort study (EPIC-InterAct) and four case-control studies (the DIAGRAM consortium, ADDITION-Ely, Norfolk Diabetes, and Cambridgeshire)	There was no significant associations of 25(OH)D with risk of type 2 diabetes. The summary OR per 1 SD lower 25(OH)D concentration was 1.01 (0.75–1.36; p = 0.94). Similarly, the results were not significant for individual alleles or any allelic scores. In the secondary analysis for glycemic traits, we identified no significant causal associations for the 4 SNPs combined or for 2 SNPs related to 25(OH)D metabolism.	Bayesian likelihood-based method	NS
Lu et al.	2018	Chinese	One-sample MR	55,46 cases and 76,655 controls	4 SNPs as primary genetic instruments: rs12785878 near DHCR7, rs10741657 near CYP2R1, rs2282679 of DBP/GC, and rs6013897 near CYP24A1. Genetic scores were estimated for the 2		Type 2 diabetes	In CKB, none of the individual variants were significantly associated with risk of diabetes (n = 5566), with per allele adjusted ORs (95% CIs) of 0.97 (0.93–1.01) for DHCR7-rs12785878, 1.00 (0.96–1.04) for CYP2R1-rs10741657, 1.00	Instrumental variable analysis, genetic risk score approach	NS

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
					synthesis SNPs and all four 25(OH)D SNPs.			(0.95–1.05) for CYP24A1-rs6013897, and 1.01 (0.96–1.05) for GC/DBP -rs2282679. The per allele adjusted OR (95% CI) of the genetic score for the 2 synthesis SNPs was 0.97 (0.93–1.01) and for the genetic score for all 4 SNPs was 0.99 (0.96–1.02).		
Vimaleswaran et al.	2013	Caucasian	Two-sample MR, using GWAS summary level data	123,864	Effect sizes of 4 GWAS-identified SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679) were calculated using D-CarDia studies; Two allelic scores (separately for genes encoding synthesis or metabolism) were generated		Body mass index. The effect sizes of IV-outcome were calculated by meta-analyzing studies involved in the GIANT consortium.	The two vitamin D allele scores were strongly associated with 25(OH)D but not with BMI (synthesis score, $p = 0.88$; metabolism score, $p = 0.08$) in the meta-analysis.	IV ratio method, Taylor expansion to calculate variance	NS
Vimaleswaran et al.	2014	Caucasian	Two-sample MR, using GWAS summary level data	108,173	Effect sizes of 4 GWAS-identified SNPs in DHCR7 (rs12785878), CYP2R1 (rs12794714), CYP24A1 (rs6013897) and GC (rs2282679), were calculated by using D-CarDia studies; Two allelic scores (separately for genes encoding synthesis or metabolism) were generated		Systolic blood pressure, diastolic blood pressure, and hypertension. The effect sizes of IV-outcome were calculated by meta-analyzing D-CarDia studies and several consortia (CHARGE, Global BPGen, and ICBP).	In instrumental variable analysis, each 10% increase in genetically instrumented 25(OH)D concentration was associated with a change of −0.29 mmHg in diastolic blood pressure (−0.52 to −0.07; $p = 0.01$), a change of −0.37 mmHg in systolic blood pressure (−0.73 to 0.003; $p = 0.052$), and an 8.1% decreased odds of hypertension (0.92, 0.87–0.97; $p = 0.002$).	IV ratio method, Taylor expansion to calculate variance	[b]
Kunutsor et al.	2014	Caucasian	Two-sample MR, using GWAS summary level data	69,395	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Systolic blood pressure; diastolic blood pressure. The effect sizes of IV-outcome were extracted from a GWAS conducted by the International Consortium of Blood Pressure	No significant results were observed for the causal effects of genetically-determined vitamin D deficiency on systolic ($p = 0.42$), or diastolic blood pressure ($p = 0.25$).	Likelihood based method	NS
Brondum-Jacobsen et al.	2015	Caucasian	One-sample MR	92,416	DHCR7 (rs7944926, rs11234027) and CYP2R1 (rs10741657, rs12794714); unweighted allele counts (0–4)	10 nmol/L lower, 25 nmol/L lower, or 50% reduction of plasma concentration of 25(OH)D	Ischemic heart disease; myocardial infarction	No evidence to suggest that genetically reduced plasma 25(OH)D is associated with increased risk of ischemic heart disease or myocardial infarction. The OR for ischemic heart disease for a 25 nmol/L genetic decrease in 25(OH)D concentration was 0.98 (0.76–1.26); for myocardial infarction 1.15 (0.83–1.59).	Wald-type ratio estimator for coefficients, and Fieller's theorem to derive the 95% confidence intervals	NS
Cuellar-Partida et al.	2017	Caucasian, and Asian (separately)	Two-sample MR, using GWAS summary level data	37,382 Caucasian, and 8376 Asian participants	Literature-reported effect sizes of IV-25(OH)D, including 6 SNPs in DHCR7 (rs12785878, rs7944926), CYP2R1		Myopic refractive error. Effect sizes of the IV-outcome were extracted from a meta-GWAS of myopic refractive error	A causal link is unlikely. The estimate for the effect of 25(OH)D on refractive error was −0.02 (−0.09, 0.04) diopters (D) per 10 nmol/L increase in 25(OH)D	Wald-type ratio estimator for coefficients, and delta methods for standard errors	NS

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
					(rs10741657, rs12794714), CYP24A1 (rs6013897) and GC (rs2282679)			concentration in Caucasians; and 0.01 (−0.17, 0.19) D per 10 nmol/L increase in Asians.		
Manousaki et al.	2016	Caucasian	Two-sample MR, using GWAS summary level data	86,599 (22,233 cases and 64,762 controls)	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Coronary artery disease. The effect sizes of IV-outcome were extracted from a GWAS conducted by CARDIoGRAM	None of the SNPs associated with 25(OH)D levels were associated with CAD (all p values > 0.6). The Mendelian randomization OR for CAD was 0.99 (0.84–1.17; $p = 0.93$) per SD decrease in log-transformed 25(OH)D levels.	Inverse variance weighted analysis	NS
Magnus et al.	2018	Caucasian	Both One-sample MR, and two sample MR	7389 women (751gestational hypertension, 135preeclampsia) for one-sample MR; 1513 preeclampsia cases and 971 controls and 1875 preeclampsia cases and 5088 controls for two-sample MR.	SNPs in genes associated with vitamin D synthesis (rs10741657, rs12785878) and metabolism (rs6013897, rs2282679)		Gestational hypertension and pre-eclampsia	The one sample MR using the total genetic risk score as an instrument did not provide strong evidence of a linear effect of 25-hydroxyvitamin D on the risk of gestational hypertension or preeclampsia: OR = 0.90 (95% CIs: 0.78 to 1.03) and 1.19 (0.92 to 1.52) per 10% decrease, respectively. The two sample MR gave an OR for preeclampsia of 0.98 (0.89 to 1.07) per 10% decrease in 25-hydroxyvitamin D level, an OR of 0.96 (0.80 to 1.15) per unit increase in the log(odds) of 25-hydroxyvitamin D level < 75 nmol/L, and an OR of 0.93 (0.73 to 1.19) per unit increase in the log(odds) of 25-hydroxyvitamin D levels < 50 nmol/L.	Genetic risk score approach; Wald ratio, MR-Egger	NS
Husemoen et al.	2014	Caucasian	One-sample MR	Inter99: 6405; MONICA10: 2656	Inter99 study, 7 SNPs (rs2282679 in GC, rs2060793 in CYP2R1, rs3829251 in NADSYN1, rs6599638 in C10orf88, rs6013897 in CYP24A1, rs12785878 near DHCR7 and rs10741657 near CYP2R1); MONICA10 study, rs2282679	Plasma 25(OH)D concentrations	Total adiponectin for Inter99 study; high molecular weight adiponectin for MONICA10 study	Although significant result was observed for rs2282679 in Inter99 study using IV regression (the causal effect was estimated to 53.9%, $p = 0.019$), it was not replicated by MONICA10 study (29.4%, $p = 0.21$).	Two-stage least square multiple instrument approach	NS
Ooi et al.	2014	Caucasian	One-sample MR	85,868	DHCR7 (rs7944926, rs11234027) and CYP2R1 (rs10741657, rs12794714)	50% decrease in plasma 25(OH)D concentrations	Non-fasting remnant cholesterol, low-density lipoprotein-cholesterol, and high-density lipoprotein-cholesterol	There were no associations between 25(OH)D-lowering alleles for either non-fasting remnant or LDL-cholesterol. A 50% decrease in plasma 25(OH)D levels was genetically associated with −6.0% (−10% to −2.3%, $p = 0.001$) lower HDL-cholesterol levels.	Two-stage least square multiple instrument approach	[b]
Liefaard et al.	2015	Caucasian	Bi-directional MR	9649	A genetic risk score constructed by using 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1	Plasma 25(OH)D concentrations	C-reactive protein	Bi-directional Mendelian randomization analyses showed no association between the vitamin D genetic risk score and ln CRP (Beta per SD = −0.018; $p = 0.082$) or	Bi-directional Mendelian randomization analysis using linear regression	NS

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
					(rs6013897) and GC (rs2282679)			the CRP genetic risk score and 25(OH)D (Beta per SD = 0.001; p = 0.99).		
Wang et al.	2018	Asian	One-sample MR	9182	A weighted genetic risk score constructed by using 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)	Plasma 25(OH)D concentrations	Nonalcoholic fatty liver disease	In the IV analysis, the causal OR of genetically determined 25(OH)D for risk of nonalcoholic fatty liver disease was 1.03 (0.99–1.07).	IV ratio method	NS
Cognition										
Jorde et al.	2015	Caucasian	One-sample MR	5980	9 SNPs in genes VDR Apal (rs7975232), Bsml (rs1544410), Fokl (rs2228570), taql (rs731236, rs7968585), NADSYN1 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs603897), DBP (rs2298850); and a SNP score		Word recall test; Digit symbol coding test; Finger tapping test	No SNP showed consistent results over the three cognitive tests. Nominally significant results were observed but did not withstand multiple correction.	Statistical methods unclear	NS
Kueider et al.	2015	Caucasian	One-sample MR	848	2 SNPs in GC (rs2282679, rs7041); and a GC composite score summing the minor alleles in each SNP	Serum 25(OH)D concentrations	Cognition, mental status measured with the Mini-Mental State Examination (22 outcomes)	Although significant results were observed for composite GC score and several outcomes including trail making test, Boston naming test, Rey-O long delay test, clock drawing test, and Pegboard non-dominant hand, none survived multiple correction.	Two-stage least square multiple instrument approach	NS
Maddock et al.	2017	Caucasian	One-sample MR	172,349	2 SNPs in DHCR7 (rs12785878), and CYP2R1 (rs12794714); and their combined synthesis score	Plasma 25(OH)D concentrations	Cognition: global and memory cognitive function	Associations with global or memory-related cognitive function were nonsignificant and in opposing directions for DHCR7 and CYP2R1, with no overall association observed for the synthesis score.	Associations between SNPs/synthesis score and cognitive function adjusted for a number of variables, through linear regressions	NS
Larsson et al.	2017	Caucasian	Two-sample MR, using GWAS summary level data	17,352 (5333 cases and 12,019 controls)	GWAS-reported (D-CarDia) effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Parkinson's disease. The effect sizes of IV-outcome were extracted from a GWAS conducted by the International Parkinson's Disease Genomic Consortium	The OR of Parkinson's disease per genetically predicted 10% lower 25(OH)D concentration, based on the 4 SNPs, was 0.98 (0.93–1.04; p = 0.56).	Inverse variance weighted analysis, MR-Egger, and weighted median approach	NS

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Population	Data type	Sample size[a]	Instrumental variables	Exposure	Outcome	Results	Methods	Sig
Mokry et al.	2016	Caucasian	Two-sample MR, using GWAS summary level data	54,162 (17,008 cases and 37,154 controls)	GWAS-reported effect sizes of IV-25(OH)D including 4 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657), CYP24A1 (rs6013897) and GC (rs2282679)		Alzheimer's disease. The effect sizes of IV-outcome were extracted from a GWAS conducted by the International Genomics of Alzheimer's Project	A 1-SD decrease in natural log-transformed 25(OH)D increased AD risk by 25% (OR = 1.25 (1.03–1.51), p = 0.021).	Two-sample MR, inverse variance weighted analysis	[b]
Larsson et al.	2018	Caucasian	Two-sample MR, using GWAS summary level data	54,162 (17,008 cases and 37,154 controls)	GWAS-reported effect sizes of IV-25(OH)D including 7 SNPs in DHCR7 (rs12785878), CYP2R1 (rs10741657, rs117913124), CYP24A1 (rs17216707), GC (rs3755967), AMDHD1 (rs1968487) and SEC23A (rs8018720)		Alzheimer's disease. The effect sizes of IV-outcome were extracted from a GWAS conducted by the International Genomics of Alzheimer's Project	The OR of AD per genetically-predicted 1-SD increase in 25OHD was 0.86 (95% CI: 0.78–0.94; p = 0.002).	Two-sample MR, inverse variance weighted analysis, MR-Egger	[b]

[a] For two-sample MR, sample size of the outcome was reported.
[b] Significant results reported.

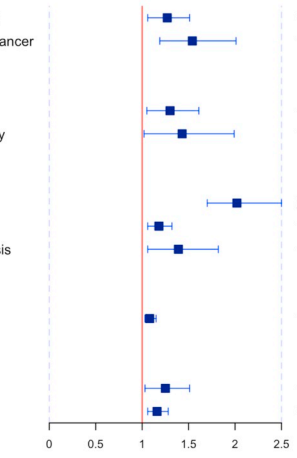
(from year 2010), and GWAS summary data from the International Multiple Sclerosis Genetic Consortium [65]. MR analysis revealed that a standard deviation decrease in natural log-transformed 25(OH)D levels doubled the risk of MS [OR = 2.02 (1.65–2.46); $p = 7.72 \times 10^{-12}$]. Another group conducted MR analyses in two different populations, the Kaiser Permanente Medical Care Plan in Northern California (KPNC) and two Swedish case-control studies [66]. The combined populations included 7391 MS cases and 14,777 controls. The authors calculated a weighted genetic score for each individual based on the number of 25(OH)D increasing alleles at three different loci (*GC*, *DHCR7* and *CYP2R1*). Meta-analysis across both populations showed that an increased genetic score was protective against MS [OR = 0.85 (0.76–0.94); p = 0.003]. Similarly, an increase in the same genetic score was also associated with a decreased risk of pediatric-onset MS [OR = 0.72 (0.55–0.94); p = 0.02] in a study of 569 cases and 16,251 controls [67]. Moreover, Individuals carrying one copy of the rare variant (CYP2R1 rs117913124_A) as identified by Manousaki et al. also had increased risk of multiple sclerosis (OR = 1.4, 95% CI = 1.19–1.64, $p = 2.63 \times 10^{-5}$) in a separate sample of 5927 MS case and 5599 control subjects [34]. Consistent with these results, and using data from the International Genomics of Alzheimer's Project, Mokry et al. further found that a standard deviation decrease in natural log-transformed 25(OH)D levels increased the risk of Alzheimer's disease [OR = 1.25(1.03–1.51), p = 0.02] [68]. Similar findings on the Alzheimer's disease were further confirmed by Larsson et al. using the 6 newly identified vitamin D index SNPs [OR = 1.16(1.06–1.28), p = 0.002] [69]. However, no significant finding has been revealed for other autoimmune disease such as asthma [70–72], rheumatoid arthritis [73,74], or other neurodegenerative disorder such as Parkinson's disease [75] or cognition status [76–78].

Nearly all outcomes assessed by MR studies of vitamin D have been negative. This is inconsistent with many observational studies, which may be affected by residual confounding. However, findings from MR studies need to be interpreted in the context of other evidence related to the particular issue under investigation. While these results may help to prevent costly RCTs that would be paid for by publicly funded agencies, it also could help to prioritize the most promising target diseases for vitamin D intervention. It is however, also worth noting that among almost all MR studies conducted to date (findings summarized in Table 3, significant findings summarized in Fig. 1), only four genetic variants (rs2282679, rs12785878, rs6013897 and rs10741657 identified by the 2010 SUNLIGHT meta-GWAS) or even fewer were used to build the instrument. The recent identification of two novel genetic loci associated with circulating 25(OH)D level, as well as the rapidly accumulating GWAS data for various disease endpoints of interest, provide an exceptional opportunity to re-evaluate the casual effect of 25(OH)D by incorporating additional genetic variants which, would provide an improved instrument strength and accuracy of estimation.

6. Challenges and future directions

Recent GWAS have shed light on the role of genetic variants in vitamin D regulation, suggesting a relatively modest SNP-heritability rate of 25(OH)D when considering only common variants. The marked gap between estimated SNP-heritability and twin-heritability are likely to be narrowed with the identification of rare or structural variants, as well as gene-gene interactions that remain to be identified. Well-powered population-based exome sequencing and whole genome sequencing studies are needed to uncover the role of low-frequency and rare variants in vitamin D. Larger GWAS, for example the emerging UK Biobank data, are still likely to identify new promising susceptibility loci which could improve our understanding to the genetic regulation of serum vitamin D homeostasis. An important public health message conveyed by the past genetic studies is that despite variation in 25(OH)D by common genetic loci, exposure to sunlight and intake of supplements remain as the key determinants of vitamin D status.

Author	Year	Population	Methods	Scale of Exposure (vitamin D levels)	Health Outcome	Estimates*	P-values
A. Binary outcomes							
Malignant disease							
Ong et al.	2016	Caucasian	Two-sample MR	per 20 nmol/L decrease	Overall ovarian cancer	1.27 (1.06-1.51)	0.0088
				per 20 nmol/L decrease	High grade serous ovarian cancer	1.54 (1.19-2.01)	0.0011
Mortality							
Afzal et al.	2014	Caucasian	One-sample MR	per 20 nmol/L decrease	All cause mortality	1.30 (1.05-1.61)	0.03
				per 20 nmol/L decrease	Cancer specific mortality	1.43 (1.02-1.99)	0.04
Autoimmune disease							
Mokry et al.	2015	Caucasian	Two-sample MR	per 25 nmol/L decrease	Multiple sclerosis	2.02 (1.70-2.50)	7.7×10−12
Rhead et al.	2016	Caucasian	One-sample MR	per SD decrease in multi-SNP score	Multiple sclerosis	1.18 (1.06-1.32)	0.003
Gianfrancesco et al.	2017	Caucasian	One-sample MR	per SD decrease in multi-SNP score	Pediatric multiple sclerosis	1.39 (1.06-1.82)	0.02
Metabolic traits							
Vimaleswaran et al.	2014	Caucasian	Two-sample MR	per 10% decrease	Hypertension	1.08 (1.03-1.15)	0.002
Cognition							
Mokry et al.	2016	Caucasian	Two-sample MR	per SD decrease	Alzheimer's disease	1.25 (1.03-1.51)	0.021
Larsson et al	2018	Caucasian	Two-sample MR	per SD decrease	Alzheimer's disease	1.16 (1.06-1.28)	0.002
B. Continuous outcomes							
Metabolic traits							
Vimaleswaran et al.	2014	Caucasian	Two-sample MR	per 10% decrease	Systolic blood pressure	0.29 (0.070-0.52)	0.01
				per 10% decrease	Diastolic blood pressure	0.37 (0.003-0.73)	0.05
Ooi et al.	2014	Caucasian	One-sample MR	per 50% decrease	High-density lipoprotein-cholesterol	0.06 (0.023-0.10)	0.001



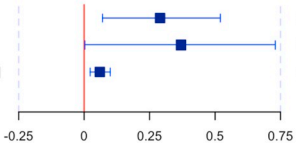


Fig. 1. Putative causal relationships between circulating vitamin D levels and health outcomes. Only significant findings from published Mendelian randomizations (MR) were plotted. There is little evidence in support for a causal link between vitamin D and bone mineral density, caries experience; malignant diseases such as breast, colorectal, prostate, lung, pancreatic, oral cancers, neuroblastoma, or glioma; inflammatory diseases such as asthma, or rheumatoid arthritis; metabolic traits such as type 2 diabetes, heart disease, fatty liver disease, or levels of small metabolites; cognitive features or Parkinson's disease. For details of all MR studies please see Table 3. *Estimates are Odds ratios for overall ovarian cancer, high grade serous ovarian cancer, multiple sclerosis, pediatric multiple sclerosis, hypertension and Alzheimer's disease; hazard ratios for all-cause mortality and cancer specific mortality; beta-coefficients (% increase) for systolic blood pressure, diastolic blood pressure and high-density lipoprotein-cholesterol.

The results of genetic studies are useful in several aspects. First of all, individual response to vitamin D administration has been demonstrated to be highly variable. Women in general show a lower level of serum vitamin D than men. It remains challenging to efficiently identify individuals who are predisposed to vitamin D deficiency and would benefit most from supplements. The novel susceptibility loci could help improve prediction power of genetic risk score and hence contribute to screening, personalized risk prediction and targeted intervention. Moreover, with the release of several high-quality functional annotations into the public domain and the development of novel analytical methods tailored to genomic data (for example, LD score regression [79,80], LD-hub etc. [81]), we are now not only able to quantify the genetic correlation between traits, but also analyze the component of SNP-heritability to unravel the functional architecture of these complex traits. Jiang et al. have observed a significant enrichment of 25(OH)D heritability in immune and hematopoietic tissues, as well as the clustering of 25(OH)D with autoimmune diseases, indicating a shared genetic basis between these traits and informs a potential direction for future studies. Finally, MR of vitamin D is tremendously important to the field of vitamin D research because RCTs generally will not be done for many diseases and particularly those which are uncommon or have a long lag-time in onset. GWAS findings have provided an unprecedented opportunity to understand a putative causal relationship between vitamin D and diseases through Mendelian randomizations, the results of which, contribute significantly to prevention and drug development. We expect some of the ambiguous causal links between vitamin D and diseases can become apparent when large-scale GWAS continue to increase the number of instruments and precision of estimation.

While acknowledging these advances, we will conclude our review by summarizing challenges. Although in principle gene-gene and gene-environment interactions help identify potential biological mechanisms underlying complex traits, and a large number of methods have been developed to detect these interactions, none have yet to be discovered in vitamin D genetics [82]. Two studies have searched for gene-by-diet interaction in vitamin D but did not discover any significant effects. For example, Jiang et al. examined G-by-E effect for the 6 GWAS-identified significant SNPs from the main association analysis, and except for the lead SNP in *CYP2R1* showing nominal significance for interaction with dietary vitamin D intake (rs10741657, $p = 0.028$), no interactions were observed for other SNPs [29]. Manousaki et al. hypothesized that carriers of *CYP2R1* low-frequency allele might respond poorly to vitamin D replacement therapy and tested it in G-by-E analysis, yet no convincing evidence for interaction was found ($p = 0.60$) [34]. It has long been recognized that studies of gene-environment interactions are generally underpowered, whose statistical power is highly dependent on the variance of exposure. There are several obvious barriers to the identification of interaction effects [82], for example, measurement errors in dietary data is common, and interaction is limited by time difference between the assessment of dietary intake and measurement of 25(OH)D levels. While one cannot completely rule out the possibility of modest interaction, future work is needed to capture such effects. Notwithstanding the massive investment in large-scale GWASs and emergence of whole genome sequencing data for vitamin D, a majority of these studies are performed exclusively among Caucasians, limiting the generalizability of findings. It will be of great importance to explore the genetic determinants of vitamin D among non-Europeans in the future.

Acknowledgement

XJ is funded by an International Postdoc Grant from the Swedish Research Council (VetenskapsRådet).

References

[1] S.H. Doppelt, Vitamin D, rickets, and osteomalacia, Orthop. Clin. North Am. 15 (1984) 671–686.
[2] M. Scaranti, G. de C. Júnior, A.O. Hoff, Vitamin D and cancer: does it really matter? Curr. Opin. Oncol. 28 (2016) 205–209, https://doi.org/10.1097/CCO.0000000000000282.
[3] B. Altieri, G. Muscogiuri, L. Barrea, C. Mathieu, C.V. Vallone, L. Mascitelli, G. Bizzaro, V.M. Altieri, G. Tirabassi, G. Balercia, S. Savastano, N. Bizzaro, C.L. Ronchi, A. Colao, A. Pontecorvi, S. Della Casa, Does vitamin D play a role in autoimmune endocrine disorders? A proof of concept, Rev. Endocr. Metab. Disord. 18 (2017) 335–346, https://doi.org/10.1007/s11154-016-9405-9.
[4] P.H.F. Gois, D. Ferreira, S. Olenski, A.C. Seguro, Vitamin D and infectious diseases: simple bystander or contributing factor? Nutrients 9 (2017), https://doi.org/10.3390/nu9070651.
[5] I. Al Mheid, A.A. Quyyumi, Vitamin D and cardiovascular disease: controversy unresolved, J. Am. Coll. Cardiol. 70 (2017) 89–100, https://doi.org/10.1016/j.jacc.2017.05.031.
[6] C.M. Issa, Vitamin D and type 2 diabetes mellitus, Adv. Exp. Med. Biol. 996 (2017) 193–205, https://doi.org/10.1007/978-3-319-56017-5_16.
[7] D. Hunter, M. De Lange, H. Snieder, A.J. MacGregor, R. Swaminathan, R.V. Thakker, T.D. Spector, Genetic contribution to bone metabolism, calcium excretion, and vitamin D and parathyroid hormone regulation, J. Bone Miner. Res. 16 (2001) 371–378, https://doi.org/10.1359/jbmr.2001.16.2.371.
[8] M.K. Shea, E.J. Benjamin, J. Dupuis, J.M. Massaro, P.F. Jacques, R.B. D'Agostino, J.M. Ordovas, C.J. O'Donnell, B. Dawson-Hughes, R.S. Vasan, S.L. Booth, Genetic and non-genetic correlates of vitamins K and D, Eur. J. Clin. Nutr. 63 (2009) 458–464, https://doi.org/10.1038/sj.ejcn.1602959.
[9] C. Karohl, S. Su, M. Kumari, V. Tangpricha, E. Veledar, V. Vaccarino, P. Raggi, Heritability and seasonal variability of vitamin D concentrations in male twins, Am. J. Clin. Nutr. 92 (2010) 1393–1398, https://doi.org/10.3945/ajcn.2010.30176.
[10] G. Jones, S.A. Strugnell, H.F. DeLuca, Current understanding of the molecular actions of vitamin D, Physiol. Rev. 78 (1998) 1193–1231, https://doi.org/10.1152/physrev.1998.78.4.1193.
[11] P. Lips, Relative value of 25(OH)D and 1,25(OH)2D measurements, J. Bone Miner. Res. 22 (2007) 1668–1671, https://doi.org/10.1359/jbmr.070716.
[12] K.J.H. Verweij, M.A. Mosing, B.P. Zietsch, S.E. Medland, Estimating heritability from twin studies, Methods Mol. Biol. 850 (2012) 151–170, https://doi.org/10.1007/978-1-61779-555-8_9.
[13] L.M. Arguelles, C.B. Langman, A.J. Ariza, F.N. Ali, K. Dilley, H. Price, X. Liu, S. Zhang, X. Hong, B. Wang, H. Xing, Z. Li, X. Liu, W. Zhang, X. Xu, X. Wang, Heritability and environmental factors affecting vitamin D status in rural Chinese adolescent twins, J. Clin. Endocrinol. Metab. 94 (2009) 3273–3281, https://doi.org/10.1210/jc.2008-1532.
[14] C.D. Engelman, T.E. Fingerlin, C.D. Langefeld, P.J. Hicks, S.S. Rich, L.E. Wagenknecht, D.W. Bowden, J.M. Norris, Genetic and environmental determinants of 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D levels in Hispanic and African Americans, J. Clin. Endocrinol. Metab. 93 (2008) 3381–3388, https://doi.org/10.1210/jc.2007-2702.
[15] N.T. Mills, M.J. Wright, A.K. Henders, D.W. Eyles, B.T. Baune, J.J. McGrath, E.M. Byrne, N.K. Hansell, E. Birosova, J.G. Scott, N.G. Martin, G.W. Montgomery, N.R. Wray, A.A.E. Vinkhuyzen, Heritability of transforming growth factor-β1 and tumor necrosis factor-receptor type 1 expression and vitamin D levels in healthy adolescent twins, Twin Res. Hum. Genet. 18 (2015) 28–35, https://doi.org/10.1017/thg.2014.70.
[16] S.-M. Orton, A.P. Morris, B.M. Herrera, S.V. Ramagopalan, M.R. Lincoln, M.J. Chao, R. Vieth, A.D. Sadovnick, G.C. Ebers, Evidence for genetic regulation of vitamin D status in twins with multiple sclerosis, Am. J. Clin. Nutr. 88 (2008) 441–447.
[17] G. Snellman, H. Melhus, R. Gedeborg, S. Olofsson, A. Wolk, N.L. Pedersen, K. Michaëlsson, Seasonal genetic influence on serum 25-hydroxyvitamin D levels: a twin study, PLoS ONE 4 (2009) e7747, , https://doi.org/10.1371/journal.pone.0007747.
[18] M. Wjst, J. Altmüller, T. Faus-Kessler, C. Braig, M. Bahnweg, E. André, Asthma families show transmission disequilibrium of gene variants in the vitamin D metabolism and signalling pathway, Respir. Res. 7 (2006) 60, , https://doi.org/10.1186/1465-9921-7-60.
[19] G. Livshits, D. Karasik, M.J. Seibel, Statistical genetic analysis of plasma levels of vitamin D: familial study, Ann. Hum. Genet. 63 (1999) 429–439.
[20] H.-J. Yu, M.-J. Kwon, H.-Y. Woo, H. Park, Analysis of 25-hydroxyvitamin D status according to age, gender, and seasonal variation, J. Clin. Lab. Anal. 30 (2016) 905–911, https://doi.org/10.1002/jcla.21955.
[21] S.P. Daiger, M.S. Schanfield, L.L. Cavalli-Sforza, Group-specific component (Gc) proteins bind vitamin D and 25-hydroxyvitamin D, Proc. Natl. Acad. Sci. U. S. A. 72 (1975) 2076–2080.
[22] P.F. Brumbaugh, M.R. Haussler, 1 Alpha,25-dihydroxycholecalciferol receptors in intestine. II. Temperature-dependent transfer of the hormone to chromatin via a specific cytosol receptor, J. Biol. Chem. 249 (1974) 1258–1262.
[23] P.F. Brumbaugh, M.R. Haussler, 1 Alpha,25-dihydroxycholecalciferol receptors in intestine. I. Association of 1 alpha,25-dihydroxycholecalciferol with intestinal mucosa chromatin, J. Biol. Chem. 249 (1974) 1251–1257.
[24] T. Sakaki, N. Kagawa, K. Yamamoto, K. Inouye, Metabolism of vitamin D3 by cytochromes P450, Front. Biosci. 10 (2005) 119–134.
[25] Z. Dastani, R. Li, B. Richards, Genetic regulation of vitamin D levels, Calcif. Tissue Int. 92 (2013) 106–117, https://doi.org/10.1007/s00223-012-9660-z.
[26] E.J. Benjamin, J. Dupuis, M.G. Larson, K.L. Lunetta, S.L. Booth, D.R. Govindaraju, S. Kathiresan, J.F. Keaney, M.J. Keyes, J.-P. Lin, J.B. Meigs, S.J. Robins, J. Rong, R. Schnabel, J.A. Vita, T.J. Wang, P.W.F. Wilson, P.A. Wolf, R.S. Vasan, Genome-wide association with select biomarker traits in the Framingham Heart Study, BMC Med. Genet. 8 (Suppl. 1) (2007) S11, https://doi.org/10.1186/1471-2350-8-S1-S11.
[27] J. Ahn, K. Yu, R. Stolzenberg-Solomon, K.C. Simon, M.L. McCullough, L. Gallicchio, E.J. Jacobs, A. Ascherio, K. Helzlsouer, K.B. Jacobs, Q. Li, S.J. Weinstein, M. Purdue, J. Virtamo, R. Horst, W. Wheeler, S. Chanock, D.J. Hunter, R.B. Hayes, P. Kraft, D. Albanes, Genome-wide association study of circulating vitamin D levels, Hum. Mol. Genet. 19 (2010) 2739–2745, https://doi.org/10.1093/hmg/ddq155.
[28] T.J. Wang, F. Zhang, J.B. Richards, B. Kestenbaum, J.B. van Meurs, D. Berry, D.P. Kiel, E.A. Streeten, C. Ohlsson, D.L. Koller, L. Peltonen, J.D. Cooper, P.F. O'Reilly, D.K. Houston, N.L. Glazer, L. Vandenput, M. Peacock, J. Shi, F. Rivadeneira, M.I. McCarthy, P. Anneli, I.H. de Boer, M. Mangino, B. Kato, D.J. Smyth, S.L. Booth, P.F. Jacques, G.L. Burke, M. Goodarzi, C.-L. Cheung, M. Wolf, K. Rice, D. Goltzman, N. Hidiroglou, M. Ladouceur, N.J. Wareham, L.J. Hocking, D. Hart, N.K. Arden, C. Cooper, S. Malik, W.D. Fraser, A.-L. Hartikainen, G. Zhai, H.M. Macdonald, N.G. Forouhi, R.J.F. Loos, D.M. Reid, A. Hakim, E. Dennison, Y. Liu, C. Power, H.E. Stevens, L. Jaana, R.S. Vasan, N. Soranzo, J. Bojunga, B.M. Psaty, M. Lorentzon, T. Foroud, T.B. Harris, A. Hofman, J.-O. Jansson, J.A. Cauley, A.G. Uitterlinden, Q. Gibson, M.-R. Järvelin, D. Karasik, D.S. Siscovick, M.J. Econs, S.B. Kritchevsky, J.C. Florez, J.A. Todd, J. Dupuis, E. Hyppönen, T.D. Spector, Common genetic determinants of vitamin D insufficiency: a genome-wide association study, Lancet 376 (2010) 180–188, https://doi.org/10.1016/S0140-6736(10)60588-0.
[29] X. Jiang, P.F. O'Reilly, H. Aschard, Y.-H. Hsu, J.B. Richards, J. Dupuis, E. Ingelsson, D. Karasik, S. Pilz, D. Berry, B. Kestenbaum, J. Zheng, J. Luan, E. Sofianopoulou, E.A. Streeten, D. Albanes, P.L. Lutsey, L. Yao, W. Tang, M.J. Econs, H. Wallaschofski, H. Völzke, A. Zhou, C. Power, M.I. McCarthy, E.D. Michos, E. Boerwinkle, S.J. Weinstein, N.D. Freedman, W.-Y. Huang, N.M. Van Schoor, N. van der Velde, L.C.P.G.M. de Groot, A. Enneman, L.A. Cupples, S.L. Booth, R.S. Vasan, C.-T. Liu, Y. Zhou, S. Ripatti, C. Ohlsson, L. Vandenput, M. Lorentzon, J.G. Eriksson, M.K. Shea, D.K. Houston, S.B. Kritchevsky, Y. Liu, K.K. Lohman, L. Ferrucci, M. Peacock, C. Gieger, M. Beekman, E. Slagboom, J. Deelen, D. van Heemst, M.E. Kleber, W. März, I.H. de Boer, A.C. Wood, J.I. Rotter, S.S. Rich, C. Robinson-Cohen, M. den Heijer, M.-R. Jarvelin, A. Cavadino, P.K. Joshi, J.F. Wilson, C. Hayward, L. Lind, K. Michaëlsson, S. Trompet, M.C. Zillikens, A.G. Uitterlinden, F. Rivadeneira, L. Broer, L. Zgaga, H. Campbell, E. Theodoratou, S.M. Farrington, M. Timofeeva, M.G. Dunlop, A.M. Valdes, E. Tikkanen, T. Lehtimäki, L.-P. Lyytikäinen, M. Kähönen, O.T. Raitakari, V. Mikkilä, M.A. Ikram, N. Sattar, J.W. Jukema, N.J. Wareham, C. Langenberg, N.G. Forouhi, T.E. Gundersen, K.-T. Khaw, A.S. Butterworth, J. Danesh, T. Spector, T.J. Wang, E. Hyppönen, P. Kraft, D.P. Kiel, Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels, Nat. Commun. 9 (2018) 260, , https://doi.org/10.1038/s41467-017-02662-2.
[30] C.D. Engelman, K.J. Meyers, J.T. Ziegler, K.D. Taylor, N.D. Palmer, S.M. Haffner, T.E. Fingerlin, L.E. Wagenknecht, J.I. Rotter, D.W. Bowden, C.D. Langefeld, J.M. Norris, Genome-wide association study of vitamin D concentrations in Hispanic Americans: the IRAS family study, J. Steroid Biochem. Mol. Biol. 122 (2010) 186–192, https://doi.org/10.1016/j.jsbmb.2010.06.013.
[31] D. Anderson, B.J. Holt, C.E. Pennell, P.G. Holt, P.H. Hart, J.M. Blackwell, Genome-wide association study of vitamin D levels in children: replication in the Western Australian Pregnancy Cohort (Raine) study, Genes Immun. 15 (2014) 578–583, https://doi.org/10.1038/gene.2014.52.
[32] B.R. Sapkota, R. Hopkins, A. Bjonnes, S. Ralhan, G.S. Wander, N.K. Mehra, J.R. Singh, P.R. Blackett, R. Saxena, D.K. Sanghera, Genome-wide association study of 25(OH) vitamin D concentrations in Punjabi Sikhs: results of the Asian Indian diabetic heart study, J. Steroid Biochem. Mol. Biol. 158 (2016) 149–156, https://doi.org/10.1016/j.jsbmb.2015.12.014.
[33] K.M. O'Brien, D.P. Sandler, M. Shi, Q.E. Harmon, J.A. Taylor, C.R. Weinberg, Genome-wide association study of serum 25-hydroxyvitamin D in US women, Front. Genet. 9 (2018) 67, , https://doi.org/10.3389/fgene.2018.00067.
[34] D. Manousaki, T. Dudding, S. Haworth, Y.-H. Hsu, C.-T. Liu, C. Medina-Gómez, T. Voortman, N. van der Velde, H. Melhus, C. Robinson-Cohen, D.L. Cousminer, M. Nethander, L. Vandenput, R. Noordam, V. Forgetta, C.M.T. Greenwood, M.L. Biggs, B.M. Psaty, J.I. Rotter, B.S. Zemel, J.A. Mitchell, B. Taylor, M. Lorentzon, M. Karlsson, V.V.W. Jaddoe, H. Tiemeier, N. Campos-Obando, O.H. Franco, A.G. Utterlinden, L. Broer, N.M. van Schoor, A.C. Ham, M.A. Ikram, D. Karasik, R. de Mutsert, F.R. Rosendaal, M. den Heijer, T.J. Wang, L. Lind, E.S. Orwoll, D.O. Mook-Kanamori, K. Michaëlsson, B. Kestenbaum, C. Ohlsson, D. Mellström, L.C.P.G.M. de Groot, S.F.A. Grant, D.P. Kiel, M.C. Zillikens, F. Rivadeneira, S. Sawcer, N.J. Timpson, J.B. Richards, Low-frequency synonymous coding variation in CYP2R1 has large effects on vitamin D levels and risk of multiple sclerosis, Am. J. Hum. Genet. 101 (2017) 227–238, https://doi.org/10.1016/j.ajhg.2017.06.014.
[35] N.J. Timpson, C.M.T. Greenwood, N. Soranzo, D.J. Lawson, J.B. Richards, Genetic architecture: the shape of the genetic contribution to human traits and disease, Nat. Rev. Genet. 19 (2018) 110–124, https://doi.org/10.1038/nrg.2017.101.
[36] J.E. Manson, S.S. Bassuk, I.-M. Lee, N.R. Cook, M.A. Albert, D. Gordon, E. Zaharris, J.G. Macfadyen, E. Danielson, J. Lin, S.M. Zhang, J.E. Buring, The VITamin D and OmegA-3 TriaL (VITAL): rationale and design of a large randomized controlled trial of vitamin D and marine omega-3 fatty acid supplements for the primary prevention of cancer and cardiovascular disease, Contemp. Clin. Trials 33 (2012) 159–171, https://doi.org/10.1016/j.cct.2011.09.009.
[37] R.E. Neale, B.K. Armstrong, C. Baxter, B. Duarte Romero, P. Ebeling, D.R. English, M.G. Kimlin, D.S.A. McLeod, R.L.O. Connell, J.C. van der Pols, A.J. Venn, P.M. Webb, D.C. Whiteman, L. Wockner, The D-health trial: a randomized trial of

vitamin D for prevention of mortality and cancer, Contemp. Clin. Trials 48 (2016) 83–90, https://doi.org/10.1016/j.cct.2016.04.005.
[38] G. Davey Smith, S. Ebrahim, 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int. J. Epidemiol. 32 (2003) 1–22, https://doi.org/10.1093/ije/dyg070.
[39] T.J. VanderWeele, E.J. Tchetgen Tchetgen, M. Cornelis, P. Kraft, Methodological challenges in Mendelian randomization, Epidemiology 25 (2014) 427–435, https://doi.org/10.1097/EDE.0000000000000081.
[40] J. Zheng, D. Baird, M.-C. Borges, J. Bowden, G. Hemani, P. Haycock, D.M. Evans, G.D. Smith, Recent developments in Mendelian randomization studies, Curr. Epidemiol. Rep. 4 (2017) 330–345, https://doi.org/10.1007/s40471-017-0128-6.
[41] S.-S. Li, L.-H. Gao, X.-Y. Zhang, J.-W. He, W.-Z. Fu, Y.-J. Liu, Y.-Q. Hu, Z.-L. Zhang, Genetically low vitamin D levels, bone mineral density, and bone metabolism markers: a Mendelian randomisation study, Sci. Rep. 6 (2016) 33202, , https://doi.org/10.1038/srep33202.
[42] S.C. Larsson, H. Melhus, K. Michaëlsson, Circulating serum 25-hydroxyvitamin D levels and bone mineral density: Mendelian randomization study, J. Bone Miner. Res. (2018), https://doi.org/10.1002/jbmr.3389.
[43] T. Dudding, S.J. Thomas, K. Duncan, D.A. Lawlor, N.J. Timpson, Re-examining the association between vitamin D and childhood caries, PLoS ONE 10 (2015) e0143769, , https://doi.org/10.1371/journal.pone.0143769.
[44] K. Trajanoska, J.A. Morris, L. Oei, H.-F. Zheng, D.M. Evans, D.P. Kiel, C. Ohlsson, J.B. Richards, F. Rivadeneira, GEFOS/GENOMOS consortium and the 23andMe research team, Assessment of the genetic and clinical determinants of fracture risk: genome wide association and Mendelian randomisation study, BMJ 362 (2018) k3225.
[45] J.-S. Ong, G. Cuellar-Partida, Y. Lu, Australian Ovarian Cancer Study, P.A. Fasching, A. Hein, S. Burghaus, M.W. Beckmann, D. Lambrechts, E. Van Nieuwenhuysen, I. Vergote, A. Vanderstichele, J. Anne Doherty, M. Anne Rossing, J. Chang-Claude, U. Eilber, A. Rudolph, S. Wang-Gohrke, M.T. Goodman, N. Bogdanova, T. Dörk, M. Dürst, P. Hillemanns, I.B. Runnebaum, N. Antonenkova, R. Butzow, A. Leminen, H. Nevanlinna, L.M. Pelttari, R.P. Edwards, J.L. Kelley, F. Modugno, K.B. Moysich, R.B. Ness, R. Cannioto, E. Høgdall, C.K. Høgdall, A. Jensen, G.G. Giles, F. Bruinsma, S.K. Kjaer, M.A. Hildebrandt, D. Liang, K.H. Lu, X. Wu, M. Bisogna, F. Dao, D.A. Levine, D.W. Cramer, K.L. Terry, S.S. Tworoger, M. Stampfer, S. Missmer, L. Bjorge, H.B. Salvesen, R.K. Kopperud, K. Bischof, K.K. Aben, L.A. Kiemeney, L.F. Massuger, A. Brooks-Wilson, S.H. Olson, V. McGuire, J.H. Rothstein, W. Sieh, A.S. Whittemore, L.S. Cook, N.D. Le, C.B. Gilks, J. Gronwald, A. Jakubowska, J. Lubiński, T. Kluz, H. Song, J.P. Tyrer, N. Wentzensen, L. Brinton, B. Trabert, J. Lissowska, J.R. McLaughlin, S.A. Narod, C. Phelan, H. Anton-Culver, A. Ziogas, D. Eccles, I. Campbell, S.A. Gayther, A. Gentry-Maharaj, U. Menon, S.J. Ramus, A.H. Wu, A. Dansonka-Mieszkowska, J. Kupryjanczyk, A. Timorek, L. Szafron, J.M. Cunningham, B.L. Fridley, S.J. Winham, E.V. Bandera, E.M. Poole, T.K. Morgan, H.A. Risch, E.L. Goode, J.M. Schildkraut, C.L. Pearce, A. Berchuck, P.D. Pharoah, G. Chenevix-Trench, P. Gharahkhani, R.E. Neale, P.M. Webb, S. MacGregor, Association of vitamin D levels and risk of ovarian cancer: a Mendelian randomization study, Int. J. Epidemiol. 45 (2016) 1619–1630, https://doi.org/10.1093/ije/dyw207.
[46] S. Afzal, P. Brøndum-Jacobsen, S.E. Bojesen, B.G. Nordestgaard, Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts, BMJ 349 (2014) g6330.
[47] E. Theodoratou, T. Palmer, L. Zgaga, S.M. Farrington, P. McKeigue, F.V.N. Din, A. Tenesa, G. Davey-Smith, M.G. Dunlop, H. Campbell, Instrumental variable estimation of the causal effect of plasma 25-hydroxy-vitamin D on colorectal cancer risk: a Mendelian randomization analysis, PLoS ONE 7 (2012) e37662, , https://doi.org/10.1371/journal.pone.0037662.
[48] O. Trummer, U. Langsenlehner, S. Krenn-Pilko, T.R. Pieber, B. Obermayer-Pietsch, A. Gerger, W. Renner, T. Langsenlehner, Vitamin D and prostate cancer prognosis: a Mendelian randomization study, World J. Urol. 34 (2016) 607–611, https://doi.org/10.1007/s00345-015-1646-9.
[49] S. Wang, D. Huo, S. Kupfer, D. Alleyne, T.O. Ogundiran, O. Ojengbede, W. Zheng, K.L. Nathanson, B. Nemesure, S. Ambs, O.I. Olopade, Y. Zheng, Genetic variation in the vitamin D related pathway and breast cancer risk in women of African ancestry in the root consortium, Int. J. Cancer 142 (2018) 36–43, https://doi.org/10.1002/ijc.31038.
[50] V.I. Dimitrakopoulou, K.K. Tsilidis, P.C. Haycock, N.L. Dimou, K. Al-Dabhani, R.M. Martin, S.J. Lewis, M.J. Gunter, A. Mondul, I.M. Shui, E. Theodoratou, K. Nimptsch, S. Lindström, D. Albanes, T. Kühn, T.J. Key, R.C. Travis, K.S. Vimaleswaran, GECCO Consortium, PRACTICAL Consortium, GAME-ON Network (CORECT, DRIVE, ELLIPSE, FOCI-OCAC, TRICL-ILCCO), P. Kraft, B.L. Pierce, J.M. Schildkraut, Circulating vitamin D concentration and risk of seven cancers: Mendelian randomisation study, BMJ 359 (2017) j4761.
[51] U.C. Winsløw, B.G. Nordestgaard, S. Afzal, High plasma 25-hydroxyvitamin D and high risk of nonmelanoma skin cancer: a Mendelian randomization study of 97 849 individuals, Br. J. Dermatol. (2017), https://doi.org/10.1111/bjd.16127.
[52] T. Dudding, M. Johansson, S.J. Thomas, P. Brennan, R.M. Martin, N.J. Timpson, Assessing the causal association between 25-hydroxyvitamin D and the risk of oral and oropharyngeal cancer using Mendelian randomization, Int. J. Cancer (2018), https://doi.org/10.1002/ijc.31377.
[53] H. Takahashi, A.J. Cornish, A. Sud, P.J. Law, B. Kinnersley, Q.T. Ostrom, K. Labreche, J.E. Eckel-Passow, G.N. Armstrong, E.B. Claus, D. Ll'yasova, J. Schildkraut, J.S. Barnholtz-Sloan, S.H. Olson, J.L. Bernstein, R.K. Lai, M.J. Schoemaker, M. Simon, P. Hoffmann, M.M. Nöthen, K.-H. Jöckel, S. Chanock, P. Rajaraman, C. Johansen, R.B. Jenkins, B.S. Melin, M.R. Wrensch, M. Sanson, M.L. Bondy, C. Turnbull, R.S. Houlston, Mendelian randomisation study of the relationship between vitamin D and risk of glioma, Sci. Rep. 8 (2018) 2339, , https://doi.org/10.1038/s41598-018-20844-w.
[54] K.S. Vimaleswaran, A. Cavadino, D.J. Berry, LifeLines Cohort Study investigators, R. Jorde, A.K. Dieffenbach, C. Lu, A.C. Alves, H.J.L. Heerspink, E. Tikkanen, J. Eriksson, A. Wong, M. Mangino, K.A. Jablonski, I.M. Nolte, D.K. Houston, T.S. Ahluwalia, P.J. van der Most, D. Pasko, L. Zgaga, E. Thiering, V. Vitart, R.M. Fraser, J.E. Huffman, R.A. de Boer, B. Schöttker, K.-U. Saum, M.I. McCarthy, J. Dupuis, K.-H. Herzig, S. Sebert, A. Pouta, J. Laitinen, M.E. Kleber, G. Navis, M. Lorentzon, K. Jameson, N. Arden, J.A. Cooper, J. Acharya, R. Hardy, O. Raitakari, S. Ripatti, L.K. Billings, J. Lahti, C. Osmond, B.W. Penninx, L. Rejnmark, K.K. Lohman, L. Paternoster, R.P. Stolk, D.G. Hernandez, L. Byberg, E. Hagström, H. Melhus, E. Ingelsson, D. Mellström, O. Ljunggren, I. Tzoulaki, S. McLachlan, E. Theodoratou, C.M.T. Tiesler, A. Jula, P. Navarro, A.F. Wright, O. Polasek, International Consortium for Blood Pressure (ICBP), Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium, Global Blood Pressure Genetics (Global BPGen) consortium, Caroline Hayward, J.F. Wilson, I. Rudan, V. Salomaa, J. Heinrich, H. Campbell, J.F. Price, M. Karlsson, L. Lind, K. Michaëlsson, S. Bandinelli, T.M. Frayling, C.A. Hartman, T.I.A. Sørensen, S.B. Kritchevsky, B.L. Langdahl, J.G. Eriksson, J.C. Florez, T.D. Spector, T. Lehtimäki, D. Kuh, S.E. Humphries, C. Cooper, C. Ohlsson, W. März, M.H. de Borst, M. Kumari, M. Kivimaki, T.J. Wang, C. Power, H. Brenner, G. Grimnes, P. van der Harst, H. Snieder, A.D. Hingorani, S. Pilz, J.C. Whittaker, M.-R. Järvelin, E. Hyppönen, Association of vitamin D status with arterial blood pressure and hypertension risk: a Mendelian randomisation study, Lancet Diabetes Endocrinol. 2 (2014) 719–729, https://doi.org/10.1016/S2213-8587(14)70113-5.
[55] E.M. Ooi, S. Afzal, B.G. Nordestgaard, Elevated remnant cholesterol in 25-hydroxyvitamin D deficiency in the general population: Mendelian randomization study, Circ. Cardiovasc. Genet. 7 (2014) 650–658, https://doi.org/10.1161/CIRCGENETICS.113.000416.
[56] S. Afzal, P. Brøndum-Jacobsen, S.E. Bojesen, B.G. Nordestgaard, Vitamin D concentration, obesity, and risk of diabetes: a Mendelian randomisation study, Lancet Diabetes Endocrinol. 2 (2014) 298–306, https://doi.org/10.1016/S2213-8587(13)70200-6.
[57] Z. Ye, S.J. Sharp, S. Burgess, R.A. Scott, F. Imamura, InterAct Consortium, C. Langenberg, N.J. Wareham, N.G. Forouhi, Association between circulating 25-hydroxyvitamin D and incident type 2 diabetes: a Mendelian randomisation study, Lancet Diabetes Endocrinol. 3 (2015) 35–42, https://doi.org/10.1016/S2213-8587(14)70184-6.
[58] K.S. Vimaleswaran, D.J. Berry, C. Lu, E. Tikkanen, S. Pilz, L.T. Hiraki, J.D. Cooper, Z. Dastani, R. Li, D.K. Houston, A.R. Wood, K. Michaëlsson, L. Vandenput, L. Zgaga, L.M. Yerges-Armstrong, M.I. McCarthy, J. Dupuis, M. Kaakinen, M.E. Kleber, K. Jameson, N. Arden, O. Raitakari, J. Viikari, K.K. Lohman, L. Ferrucci, H. Melhus, E. Ingelsson, L. Byberg, L. Lind, M. Lorentzon, V. Salomaa, H. Campbell, M. Dunlop, B.D. Mitchell, K.-H. Herzig, A. Pouta, A.-L. Hartikainen, Genetic Investigation of Anthropometric Traits-GIANT Consortium, E.A. Streeten, E. Theodoratou, A. Jula, N.J. Wareham, C. Ohlsson, T.M. Frayling, S.B. Kritchevsky, T.D. Spector, J.B. Richards, T. Lehtimäki, W.H. Ouwehand, P. Kraft, C. Cooper, W. März, C. Power, R.J.F. Loos, T.J. Wang, M.-R. Järvelin, J.C. Whittaker, A.D. Hingorani, E. Hyppönen, Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts, PLoS Med. 10 (2013) e1001383, , https://doi.org/10.1371/journal.pmed.1001383.
[59] P. Brøndum-Jacobsen, M. Benn, S. Afzal, B.G. Nordestgaard, No evidence that genetically reduced 25-hydroxyvitamin D is associated with increased risk of ischaemic heart disease or myocardial infarction: a Mendelian randomization study, Int. J. Epidemiol. 44 (2015) 651–661, https://doi.org/10.1093/ije/dyv078.
[60] G. Cuellar-Partida, K.M. Williams, S. Yazar, J.A. Guggenheim, A.W. Hewitt, C. Williams, J.J. Wang, P.-F. Kho, S.M. Saw, C.-Y. Cheng, T.Y. Wong, T. Aung, T.L. Young, J.W.L. Tideman, J.B. Jonas, Consortium for Refractive Error and Myopia (CREAM), P. Mitchell, R. Wojciechowski, D. Stambolian, P. Hysi, C.J. Hammond, D.A. Mackey, R.M. Lucas, S. MacGregor, Genetically low vitamin D concentrations and myopic refractive error: a Mendelian randomization study, Int. J. Epidemiol. 46 (2017) 1882–1890, https://doi.org/10.1093/ije/dyx068.
[61] D. Manousaki, L.E. Mokry, S. Ross, D. Goltzman, J.B. Richards, Mendelian randomization studies do not support a role for vitamin D in coronary artery disease, Circ. Cardiovasc. Genet. 9 (2016) 349–356, https://doi.org/10.1161/CIRCGENETICS.116.001396.
[62] L.L.N. Husemoen, T. Skaaby, T. Martinussen, T. Jørgensen, B.H. Thuesen, C. Kistorp, J. Jeppesen, J.P. Thyssen, M. Meldgaard, P.B. Szecsi, M. Fenger, A. Linneberg, Investigating the causal effect of vitamin D on serum adiponectin using a Mendelian randomization approach, Eur. J. Clin. Nutr. 68 (2014) 189–195, https://doi.org/10.1038/ejcn.2013.233.
[63] M.C. Liefaard, S. Ligthart, A. Vitezova, A. Hofman, A.G. Uitterlinden, J.C. Kiefte-de Jong, O.H. Franco, M.C. Zillikens, A. Dehghan, Vitamin D and C-reactive protein: a Mendelian randomization study, PLoS ONE 10 (2015) e0131740, , https://doi.org/10.1371/journal.pone.0131740.
[64] N. Wang, C. Chen, L. Zhao, Y. Chen, B. Han, F. Xia, J. Cheng, Q. Li, Y. Lu, Vitamin D and nonalcoholic fatty liver disease: bi-directional Mendelian randomization analysis, EBioMedicine 28 (2018) 187–193, https://doi.org/10.1016/j.ebiom.2017.12.027.
[65] L.E. Mokry, S. Ross, O.S. Ahmad, V. Forgetta, G.D. Smith, D. Goltzman, A. Leong, C.M.T. Greenwood, G. Thanassoulis, J.B. Richards, Vitamin D and risk of multiple sclerosis: a Mendelian randomization study, PLoS Med. 12 (2015) e1001866, , https://doi.org/10.1371/journal.pmed.1001866.
[66] B. Rhead, M. Bäärnhielm, M. Gianfrancesco, A. Mok, X. Shao, H. Quach, L. Shen, C. Schaefer, J. Link, A. Gyllenberg, A.K. Hedström, T. Olsson, J. Hillert, I. Kockum, M.M. Glymour, L. Alfredsson, L.F. Barcellos, Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk, Neurol. Genet. 2 (2016) e97, , https://doi.org/10.1212/NXG.0000000000000097.
[67] M.A. Gianfrancesco, P. Stridh, B. Rhead, X. Shao, E. Xu, J.S. Graves, T. Chitnis, A. Waldman, T. Lotze, T. Schreiner, A. Belman, B. Greenberg, B. Weinstock-Guttman, G. Aaen, J.M. Tillema, J. Hart, S. Caillier, J. Ness, Y. Harris, J. Rubin, M. Candee, L. Krupp, M. Gorman, L. Benson, M. Rodriguez, S. Mar, I. Kahn, J. Rose, S. Roalstad, T.C. Casper, L. Shen, H. Quach, D. Quach, J. Hillert, M. Bäärnhielm, A. Hedstrom, T. Olsson, I. Kockum, L. Alfredsson, C. Metayer, C. Schaefer, L.F. Barcellos, E. Waubant, Network of Pediatric Multiple Sclerosis Centers,

Evidence for a causal relationship between low vitamin D, high BMI, and pediatric-onset MS, Neurology 88 (2017) 1623–1629, https://doi.org/10.1212/WNL.0000000000003849.
[68] L.E. Mokry, S. Ross, J.A. Morris, D. Manousaki, V. Forgetta, J.B. Richards, Genetically decreased vitamin D and risk of Alzheimer disease, Neurology 87 (2016) 2567–2574, https://doi.org/10.1212/WNL.0000000000003430.
[69] S.C. Larsson, M. Traylor, H.S. Markus, K. Michaëlsson, Serum parathyroid hormone, 25-hydroxyvitamin D, and risk of Alzheimer's disease: a Mendelian randomization study, Nutrients 10 (2018), https://doi.org/10.3390/nu10091243.
[70] E.B. Hysinger, J.D. Roizen, F.D. Mentch, L. Vazquez, J.J. Connolly, J.P. Bradfield, B. Almoguera, P.M. Sleiman, J.L. Allen, M.A. Levine, H. Hakonarson, Mendelian randomization analysis demonstrates that low vitamin D is unlikely causative for pediatric asthma, J. Allergy Clin. Immunol. 138 (2016) 1747–1749.e4, https://doi.org/10.1016/j.jaci.2016.06.056.
[71] D. Manousaki, L. Paternoster, M. Standl, M.F. Moffatt, M. Farrall, E. Bouzigon, D.P. Strachan, F. Demenais, M. Lathrop, W.O.C.M. Cookson, J.B. Richards, Vitamin D levels and susceptibility to asthma, elevated immunoglobulin E levels, and atopic dermatitis: a Mendelian randomization study, PLoS Med. 14 (2017) e1002294, , https://doi.org/10.1371/journal.pmed.1002294.
[72] Y. Mao, Y. Zhan, Y. Huang, Vitamin D and asthma: a Mendelian randomization study, Ann Allergy Asthma Immunol 119 (2017) 95–97.e1, https://doi.org/10.1016/j.anai.2017.05.018.
[73] S. Viatte, A. Yarwood, K. McAllister, S. Al-Mudhaffer, B. Fu, E. Flynn, D.P.M. Symmons, A. Young, A. Barton, The role of genetic polymorphisms regulating vitamin D levels in rheumatoid arthritis outcome: a Mendelian randomisation approach, Ann. Rheum. Dis. 73 (2014) 1430–1433, https://doi.org/10.1136/annrheumdis-2013-204972.
[74] A. Yarwood, S. Viatte, D. Plant, A.W. Morgan, J. Isaacs, A.G. Wilson, K. Hyrich, S. Eyre, A. Barton, Biologics in Rheumatoid Arthritis Genetics and Genomics Study Syndicate (BRAGGSS), Testing the role of vitamin D in response to antitumour necrosis factor α therapy in a UK cohort: a Mendelian randomisation approach, Ann. Rheum. Dis. 73 (2014) 938–940, https://doi.org/10.1136/annrheumdis-2013-204443.
[75] S.C. Larsson, A.B. Singleton, M.A. Nalls, J.B. Richards, International Parkinson's Disease Genomics Consortium (IPDGC), No clear support for a role for vitamin D in Parkinson's disease: a Mendelian randomization study, Mov. Disord. 32 (2017) 1249–1252, https://doi.org/10.1002/mds.27069.
[76] R. Jorde, E.B. Mathiesen, S. Rogne, T. Wilsgaard, M. Kjærgaard, G. Grimnes, H. Schirmer, Vitamin D and cognitive function: the Tromsø Study, J. Neurol. Sci. 355 (2015) 155–161, https://doi.org/10.1016/j.jns.2015.06.009.
[77] A.M. Kueider, T. Tanaka, Y. An, M.H. Kitner-Triolo, E. Palchamy, L. Ferrucci, M. Thambisetty, State- and trait-dependent associations of vitamin-D with brain function during aging, Neurobiol. Aging 39 (2016) 38–45, https://doi.org/10.1016/j.neurobiolaging.2015.11.002.
[78] J. Maddock, A. Zhou, A. Cavadino, E. Kuźma, Y. Bao, M.C. Smart, K.-U. Saum, B. Schöttker, J. Engmann, M. Kjærgaard, V. Karhunen, Y. Zhan, T. Lehtimäki, S.P. Rovio, L. Byberg, J. Lahti, P. Marques-Vidal, A. Sen, L. Perna, H. Schirmer, A. Singh-Manoux, J. Auvinen, N. Hutri-Kähönen, M. Kähönen, L. Kilander, K. Räikkönen, H. Melhus, E. Ingelsson, I. Guessous, K.E. Petrovic, H. Schmidt, R. Schmidt, P. Vollenweider, L. Lind, J.G. Eriksson, K. Michaëlsson, O.T. Raitakari, S. Hägg, N.L. Pedersen, K.-H. Herzig, M.-R. Järvelin, J. Veijola, M. Kivimaki, R. Jorde, H. Brenner, M. Kumari, C. Power, D.J. Llewellyn, E. Hyppönen, Vitamin D and cognitive function: a Mendelian randomisation study, Sci. Rep. 7 (2017) 13230, , https://doi.org/10.1038/s41598-017-13189-3.
[79] B.K. Bulik-Sullivan, P.-R. Loh, H.K. Finucane, S. Ripke, J. Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, N. Patterson, M.J. Daly, A.L. Price, B.M. Neale, LD Score regression distinguishes confounding from polygenicity in genome-wide association studies, Nat. Genet. 47 (2015) 291–295, https://doi.org/10.1038/ng.3211.
[80] H.K. Finucane, B. Bulik-Sullivan, A. Gusev, G. Trynka, Y. Reshef, P.-R. Loh, V. Anttila, H. Xu, C. Zang, K. Farh, S. Ripke, F.R. Day, ReproGen Consortium, Schizophrenia Working Group of the Psychiatric Genomics Consortium, RACI Consortium, S. Purcell, E. Stahl, S. Lindstrom, J.R.B. Perry, Y. Okada, S. Raychaudhuri, M.J. Daly, N. Patterson, B.M. Neale, A.L. Price, Partitioning heritability by functional annotation using genome-wide association summary statistics, Nat. Genet. 47 (2015) 1228–1235, https://doi.org/10.1038/ng.3404.
[81] J. Zheng, A.M. Erzurumluoglu, B.L. Elsworth, J.P. Kemp, L. Howe, P.C. Haycock, G. Hemani, K. Tansey, C. Laurin, Early Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium, B.S. Pourcain, N.M. Warrington, H.K. Finucane, A.L. Price, B.K. Bulik-Sullivan, V. Anttila, L. Paternoster, T.R. Gaunt, D.M. Evans, B.M. Neale, LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis, Bioinformatics 33 (2017) 272–279, https://doi.org/10.1093/bioinformatics/btw613.
[82] H. Aschard, A perspective on interaction effects in genetic association studies, Genet. Epidemiol. 40 (2016) 678–688, https://doi.org/10.1002/gepi.21989.