An atlas on causal role of circulating vitamin D concentrations in human complex traits and diseases

Xia Jiang[1,2*], Tian Ge[3], Chia-Yen Chen[3-5*]

1. Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, 677 Huntington avenue, Boston, MA 02115, USA
2. Unit of Cardiovascular disease, Institute of Environmental Medicine, Karolinska Institutet, 13 Nobels väg, Stockholm 17177, Sweden
3. Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA
4. Analytic and Translational Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA
5. Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

Correspondence to:

Dr. Chia-Yen Chen
Psychiatric and Neurodevelopmental Genetics Unit
Center for Genomic Medicine
Massachusetts General Hospital
185 Cambridge Street
Boston, MA 02114, USA
Email: chiayenc@gmail.com
Tel: 617-643-5405

Dr. Xia Jiang
Unit of Cardiovascular disease
Institute of Environmental Medicine
Karolinska Institutet
13 Nobels väg, Stockholm 17177, Sweden
Email: xiajiang@hsph.harvard.edu

Abstract

Background: Vitamin D has been linked to a variety of diseases in observational studies. However, the causal role of vitamin D in human complex traits and diseases being unclear and under much debate. In this study, we used Mendelian randomization (MR) with genetic variants (as instrumental variables) to examine the causal role of vitamin D on various human complex traits and diseases.

Methods: We performed MR analysis using genome-wide significant 25(OH)D SNPs from the latest vitamin D GWAS and 45 large-scale meta-GWASs on various outcomes (average sample size=137 112 per GWAS) to determine the causal role of 25(OH)D. We applied MR-Egger regression, MR-PRESSO, and weighted median approach to estimate the causal effect of 25(OH)D while examining and controlling for potential biases from horizontal pleiotropy.

Results: We found limited evidence in support for causal effects of 25(OH)D on obesity-related traits, autoimmune diseases, cardiometabolic traits, neurological and psychiatric disorders. Sensitivity analysis using additional vitamin D instruments (association $P<10^{-5}$) did not show any additional evidence for causal effect of vitamin D. Notably, we identified horizontal pleiotropy for 20% of the trait pairs analyzed here.

Conclusions: Despite the largely augmented sample size and substantially improved statistical power with the most recent, largest vitamin D GWAS, our MR analysis did not convincingly support a causal effect of circulating 25(OH)D on complex traits and diseases examined here. Our results can inform ongoing and future randomized clinical trials of vitamin D supplementation. Future studies are warranted to prioritize the most promising target diseases for vitamin D intervention.

Introduction

Vitamin D and its major circulating form, 25-hydroxyvitamin D [25(OH)D], is a fat soluble, steroid pre-hormone that plays an essential role in human health. Vitamin D deficiency has long been observed to be associated with rickets and osteomalacia.[1] In addition, a variety of common diseases such as cancer, autoimmune inflammatory diseases, cardiovascular conditions and diabetes, have been reported to be linked with vitamin D in observational studies.[2-6] However, measurement error, confounding or reverse causality hinders causal interpretation of the results from observational studies. Although the effect of circulating vitamin D on disease risk can be demonstrated by traditional randomized controlled trials (RCTs), large-scale RCTs of vitamin D supplementation are not readily available due to their high cost and long duration.[7,8] A recently completed RCT by VITAL (ClinicalTrials.gov: NCT01169259) involving a total of 25 871 participants with a median follow-up period of 5.3 years did not find lowered incidence of invasive cancer or cardiovascular events for those who took vitamin D_3 at a dose of 2000 IU per day compared the placebo group.[9] In addition, there are ongoing RCTs aiming at other diseases and conditions, such as depression (DepFuD; ClinicalTrials.gov: NCT02521012) and cognitive change (ClinicalTrials.gov: NCT03613116). To better inform the field on the chance of success for such vitamin D trials, a better understanding in the causal role of vitamin D in these diseases and traits, preferably through less costly observational study, is necessary.

The genetic basis of circulating 25(OH)D has been demonstrated by several large-scale genome-wide association studies (GWASs), which have identified associations of *GC, NADSYN1/DHCR7, CYP2R1, CYP24A1, SEC23A, AMDHD1* with serum levels of vitamin D.[10,11] These genetic discoveries in 25(OH)D may help disentangle the causal relationships between vitamin D and other traits through Mendelian randomization (MR), an approach that uses genetic variants as

instrumental variables (IVs) for assessing the causal effect of an exposure on an outcome.[12] Under certain assumptions – that the IVs are strongly associated with the exposure, but not associated with any confounders of the exposure-outcome relationship, nor affect the outcome via other pathways than through the exposure – an unbiased estimate of the causal effect can be estimated using the observed IV-exposure and IV-outcome associations.[13]

To date, a large number of MR studies examining the effect of circulating vitamin D on various diseases have been conducted. However, the causal role of vitamin D beyond its importance for bone health remains unclear and is under much debate, as nearly all outcomes assessed by MR studies of vitamin D have been negative. The only replicated results from MR studies are higher vitamin D level decreases the risk of adult/pediatric multiple sclerosis[14-16] and Alzheimer's disease;[17,18] whereas findings from MR studies for blood pressure, ovarian cancer, and all-cause mortality are yet to be replicated.[19-22] Overall, results from MR studies for vitamin D are not consistent with conventional observational studies, which have associated vitamin D with a variety of human diseases and traits. It is also worth noting that among almost all hitherto available MR studies for vitamin D are based on genetic variants (SNPs rs2282679, rs12785878, rs6013897 and rs10741657) identified by the SUNLIGHT meta-GWAS published in 2010.[10] The recently updated vitamin D meta-GWAS identified two new genetic loci associated with circulating 25(OH)D level and improved SNP-vitamin D association estimates with larger sample size, which provided an exceptional opportunity to re-evaluate the casual effect of 25(OH)D with an improved instrument strength and accuracy of estimation. Moreover, the rapidly accumulating publicly available GWAS summary statistics for various diseases and traits allowed us to evaluate the effect of vitamin D on a wide range of outcomes using two-sample MR.

Here, we conducted a two-sample MR analysis (where IV-exposure and IV-outcome associations are estimated in two samples) to examine the effect of 25(OH)D on complex human traits and diseases using 45 GWASs. Five genetic variants associated with plasma 25(OH)D concentration were used as IVs. Summary statistics for the IV-exposure associations were extracted from the largest vitamin D GWAS involving 73 699 individuals. Summary statistics for the IV-outcome associations were extracted from publicly available GWASs of 45 traits involving 6170 075 individuals (average 137 112 individuals per GWAS).

Methods

Data for IV-exposure

We retrieved summary data for the associations between six SNPs and circulating 25(OH)D concentration from the SUNLIGHT meta-GWAS involving 79 366 discovery samples and 42 757 replication samples of European ancestry. Genome-wide analyses were performed within each cohort following a uniform plan. Specifically, an additive genetic model using linear regression on natural-log transformed 25(OH)D were fitted to each SNP, and a fixed-effects inverse variance weighted meta-analysis across the contributing cohorts was performed, with control for population structure within each cohort, and consistent quality control thresholds across cohorts (minor allele frequency (MAF)>0.05, imputation info score>0.8, Hardy-Weinberg equilibrium (HWE)>1×10^{-6}). A minimum of 10 000 individuals was required to contribute to each reported SNP-phenotype association.[11]

Among the six SNPs, four were previously identified as being robustly associated with vitamin D (rs3755967 at *GC*, rs12785878 at *NADSYN1/DHCR7*, rs10741657 at *CYP2R1*, rs17216707 at *CYP24A1*); and the other two were newly identified and replicated (rs10745742 at *AMDHD1*, rs8018720 at *SEC23A*). We excluded strand ambiguous SNPs (A/T and G/C SNPs) from the analysis to harmonize between exposure and outcome GWASs and used five SNPs as IVs. In addition to the leading index SNPs, we further performed P-value informed pruning of SNPs across the genome based on their linkage disequilibrium (LD) patterns (r^2<0.0001 with clumping window size of 1Mb). We selected all LD-clumped SNPs associated with 25(OH)D concentration at P<1×10^{-5}. This resulted in an additional 21 independent index SNPs which may increase the power of our analysis by incorporating more instruments. Excluding strand ambiguous SNPs from this

extended SNP set yielded a total of 19 SNPs as IVs. Details of the IVs are presented in **Supplementary Table 1**.

Data for IV-outcome

We retrieved summary data for the associations between index SNPs and 45 complex traits and diseases from publicly available GWASs conducted mainly in European ancestry populations (3 GWASs were conducted in a mixed population including both European and Asian samples; **Supplementary Table 2**). These traits span a wide range of phenotypes including anthropometric traits, autoimmune inflammatory diseases, cardiovascular events and metabolomic traits, and neurological and psychiatric disorders. For each trait, we retrieved the appropriate variant annotations (SNP rs ID, chromosome, position, reference and alternate allele) and summary statistics (effect size, standard error, P-value, allele frequency [if available] and sample size of the study [if available]) for each index SNP. The sample sizes of these GWASs range from 9 954 to 766 345.

Statistical analysis

MR uses SNPs as IVs to estimate the effects of risk factors on outcomes while controlling for potential confounding, by leveraging the random allocation of SNP alleles at conception. Three assumptions need to be satisfied to ensure a valid IV. The first is the relevance assumption, that IVs should be strongly associated with the exposure; the second assumption requires no association between IVs and confounders of the exposure-outcome relationship; and the third is the exclusion restriction assumption, indicating that genetic variants should affect the outcome only through exposure. If all MR assumptions are met, a causal effect can be estimated based on the observed IV-exposure and IV-outcome associations.

We conducted two-sample MR to test for the potential causal relationship between circulating 25(OH)D and various traits. We applied a number of MR methods to estimate causal effects including an inverse variance weighted meta-analysis (IVW) approach,[23] MR-PRESSO,[24] MR-Egger,[25] and a weighted median approach.[26] While IVW does not account for potential horizontal pleiotropy, i.e., a genetic variant affects the outcome via a separate biological pathway from the exposure under investigation (violation of exclusion restriction assumption), MR-PRESSO, MR-Egger and weighted median approaches control potential horizontal pleiotropy under different model assumptions. In addition, we performed MR-Egger regression, modified Q test, modified Q' tests, and MR-PRESSO to detect estimation bias due to horizontal pleiotropy. MR-PRESSO implements a global test to evaluate the presence of horizontal pleiotropy, and an outlier test to detect specific outliers.[24] In MR-Egger, significant difference of an intercept from zero suggests the existence of average directional horizontal pleiotropy.[25] Modified Q and Q' tests are traditionally used to identify over-dispersion and have been applied in the context of MR to detect outliers caused by horizontal pleiotropy.[27]

Lastly, we estimated the power of detecting the causal effects of 25(OH)D on various traits in our study. We followed the simulation framework outlined in Verbanck et al.,[24] while assuming five IVs having the same effect sizes and standard errors for exposure, and allele frequencies as the five SNPs used in our real data analysis. We showed the power to detect significant causal effects at different sample sizes for the outcome GWAS (ranging from 10 000 to 150 000; matched with real outcome GWAS sample sizes), and different true causal effect sizes (ranging from 0 to 1) at a type-I error rate of 0.05.

Results

As shown in **Table 1**, **Table 2** and **Fig 1**, in the primary analysis where five index SNPs were used as IVs, vitamin D did not show any causal effects on the outcomes examined here. However, several traits were identified to be potentially causally affected by 25(OH)D using both IVW and the adjusted MR-PRESSO approaches, and without apparent signs of horizontal pleiotropy. These traits include height (β=0.12, 95%CI: 0.01-0.23), hemoglobin A1c (HbA1c; β=0.04, 95%CI: 0.02-0.07), insomnia complaints (odds ratio=1.23, 95%CI: 1.03-1.48), age at menopause (β=–0.42, 95%CI: –0.08- –0.03). However, only HbA1c remained significant in the MR-Egger regression approach (β=0.05, 95%CI: 0.01-0.09), and none of these putative causal pairs survived multiple testing correction. We visualized the effect sizes of IV-exposure vs. IV-outcome for the four nominally significant pairs in scatter plot (**Fig 2**). The corresponding odds ratios and 95%CI of binary traits are presented in **Supplementary Table 3**.

We further found that 25(OH)D levels significantly affected two traits using the MR-Egger regression approach: the Crohn's disease (β=0.45, 95%CI: 0.08-0.82) and triglycerides (β=0.12, 95%CI: 0.02-0.21). However, these two traits showed evidence for horizontal pleiotropy ($P_{pleiotropy}$=0.011 and 0.008 from MR-Egger intercept test for Crohn's disease and triglycerides, respectively) and were not significant in either IVW (P_{IVW}=0.98 and 0.76) or the adjusted MR-PRESSO approach ($P_{MR\text{-}PRESSO}$=0.98 and 0.76).

We also performed a sensitivity analysis using an expanded set of IVs (independent SNPs associated with 25(OH)D levels at the P-value threshold of 1×10^{-5}), as this approach was used in previous MR studies when fewer genome-wide significant SNPs were available as IVs.[28,29] As shown in **Supplementary Fig 1**, **Supplementary Table 4** and **Supplementary Table 5**, when incorporating additional IVs, only insomnia complaints remained nominally significant (P<0.05) using both the IVW and the adjusted MR-PRESSO approaches without evidence of horizontal

pleiotropy. Although height and HbA1c showed some degrees of horizontal pleiotropy, both were nominally significant in the adjusted MR-PRESSO approach ($P_{MR\text{-}PRESSO}$=0.02 and 0.01 for height and HbA1c, respectively).

For a majority of outcomes that we studied here (26 out of 45), their GWASs included more than 110 000 individuals. Under this sample size, our current study had >80% power to detect a causal effect (β) of 0.2 at the P-value threshold of 0.05. We also presented power estimation for a range of GWAS sample sizes of the outcome (**Supplementary Table 6**).

Discussion

In this study, we used two-sample MR methods with 25(OH)D-associated SNPs and 45 large-scale meta-GWASs covering a wide spectrum of human complex traits and diseases to determine the causal role of 25(OH)D.

In general, our results, which were based an increased number of IVs and larger GWAS sample sizes, showed limited evidence for an causal effect for 25(OH)D on the traits or diseases investigated here. An earlier MR study which aggregated information from 21 adult cohorts with up to 42 024 participants and explored the causal relationship between vitamin D status and obesity, did not identify any effect of genetically instrumented 25(OH)D on BMI (P=0.57).[30] In our current analysis, we expanded this outcome category by incorporating body fat percentage, hip circumference, waist circumference, waist-to-hip ratio (both BMI adjusted and unadjusted), in addition to BMI. We used the latest GWAS summary statistics of these outcomes with substantially augmented sample sizes, ranging from 76 137 in body fat percentage to 322 154 in BMI. We did not find any evidence in support for a causal role of 25(OH)D in obesity-related traits (all P-values>0.2). Likewise, earlier MR studies have reported null findings for most of the autoimmune diseases including inflammatory bowel disease [Crohn's disease (P=0.67), ulcerative colitis (P=0.42)],[31] eczema (P=0.27),[32] lupus (P=0.79) and rheumatoid arthritis (P=0.66).[33] We results using summary statistics from greatly enlarged GWAS were in line with previous findings in autoimmune diseases. Although a causal role of 25(OH)D has been considered more solid with multiple sclerosis, where three independent MR studies have consistently identified an association between genetically predicated 25(OH)D and the disease,[14-16] we unfortunately could not replicate this finding due to limited data accessibility. Among the cardiometabolic traits, previous MR studies did not identify any causal role of 25(OH)D with coronary artery disease,[34] myocardial

infarction,[35] total cholesterol levels including triglycerides and low-density lipoprotein (LDL) cholesterol,[36] type 2 diabetes,[37,38] fasting insulin and adiponectin,[39] all of which, were further validated by our results. There are, however, some discrepancies. For example, an earlier study incorporated 31 435 individuals and used four index SNPs from two 25(OH)D-associated genes found no associations between 25(OH)D-lowering alleles and either non-fasting remnant cholesterol or LDL-cholesterol, but identified that a 50% decrease in 25(OH)D levels was genetically associated with 6.0% (P=0.001) lower HDL-cholesterol levels.[36] The effect of 25(OH)D on HDL-cholesterol was no longer significant in our current analysis where the summary statistics of a more recent HDL GWAS involving 187 167 individuals (a 5-fold increase in sample size) were used. Finally, putative causal relationships between 25(OH)D and neurological or psychiatric traits have rarely been investigated. So far, only a few MR studies explored the role of 25(OH)D in cognitive function,[40] major depression,[41] and Alzheimer's disease;[17,18] and nominally significant findings have only been reported in Alzheimer's disease but not the other two traits. It might worth noting that previous MR studies conducted in Alzheimer's disease used either IV-exposure association estimates from the smaller 25(OH)D GWAS (SUNLIGHT consortium 2010)[10] or correlated IVs, which affected the validity of the results. In our updated analysis, the causal relationship between 25(OH)D and Alzheimer's disease attenuated (P=0.053). Extending the outcome list by incorporating additional neurological or psychiatric traits did not identify any significant results. These results are important in informing the ongoing RCT of vitamin D on depression (ClinicalTrials.gov: NCT02521012) and cognitive change (ClinicalTrials.gov: NCT03613116).

While null findings for the outcomes examined here would the main conclusion of our study, we found nominally significant results for 25(OH)D to affect height (P=0.026), glycated hemoglobin

(HbA1c) (P=0.001), age at menopause (P=0.03) and insomnia (P=0.02). Note that one previous MR study did not identify any association between 25(OH)D and glycemic traits.[42] These results should be interpreted with caution given the large number of statistical tests we performed, and the fact that none of them survived multiple testing correction. Further investigations are warranted to elucidate these suggestive findings.

To ensure the validity of MR results, several important assumptions need to be satisfied. First of all, the relevance assumption, that IVs should be strongly associated with the exposure, is naturally guaranteed by selecting 25(OH)D-associated independent SNPs with genome-wide significance as IVs. Secondly, none of the IVs used in our analysis were cited by the NHGRI-EBI Catalog of published GWASs as associated with potential confounders, such as BMI, smoking or alcohol consumption at $P=1\times10^{-5}$ level. This could also be partly reflected by negligible associations between 25(OH)D and various traits from the current and past MR studies, many of which, act as potential confounders for other disease outcomes (e.g., obesity related traits). Finally, the exclusion restriction assumption requires IVs to affect the outcome only through the exposure (no horizontal pleiotropy). We employed a number of methods to control for horizontal pleiotropy. While MR-Egger regression with only a few SNPs can be underpowered to identify pleiotropy, we also used MR-PRESSO global test and modified Q and Q' tests test for horizontal pleiotropy. We also increased the number of IVs by incorporating independent SNPs associated with 25(OH)D at $P=1\times10^{-5}$ level. No causal relationships appeared when using additional instruments after multiple testing correction. We expect that some of these ambiguous causal links will become evident in the future when larger GWASs become available. However, we also note that our MR analysis is well-powered to detect moderate causal effects of 25(OH)D (>80% power with a causal effect $\beta>0.2$), as shown in our power calculation.

Our MR analysis showed limited evidence on any effects of vitamin D on the traits and diseases examined here. This is inconsistent with many observational studies which may be affected by various confounding biases. Our results may inform costly RCTs that would be paid for by publicly funded agencies and help to prioritize the most promising target diseases for vitamin D intervention.

Acknowledgements

We thank the various genome-wide association consortia for generously sharing the genome-wide association summary statistics. This work was supported by the Swedish Research Council [Vetenskapsrådet International Postdoc grant] (X.J.) and the National Institute on Aging at the National Institutes of Health [grant number K99AG054573] (T.G.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

X.J. and C-Y.C. contributed to study conception, data analysis, interpretation of the results and drafting of the manuscript. T.G. contributed to interpretation of the results and critical revision of the manuscript.

Competing interests

The authors declare no competing interests.

Supplementary information

Supplementary Figure 1 and Supplementary Tables 1 to 6 are provided in separate PDF and Excel files.

Supplementary Figure 1. Estimated causal effects of 25(OH)D on various traits and diseases based on the MR-PRESSO approach SNPs associated with 25(OH)D at p-value < 10^{-5}

Supplementary Table 1. List of genetic instrumental variables (SNPs)

Supplementary Table 2. Referencse for publicly available GWAS summary statistics used in this study

Supplementary Table 3. Causal effect estimates converted to odds ratio for binary traits

Supplementary Table 4. Causal effect estimates of 25(OH)D with 45 complex traits and diseases using SNPs associated with 25(OH)D at p-value < 10^{-5}

Supplementary Table 5. Test for potential biases in MR analysis using SNPs associated with 25(OH)D at p-value < 10^{-5}

Supplementary Table 6. Power calculation for hypothesis tests of detecting causal effects

References

1. Doppelt SH. Vitamin D, rickets, and osteomalacia. *Orthop Clin North Am*. 1984 Oct;**15**(4):671–686.
2. Scaranti M, Júnior G de C, Hoff AO. Vitamin D and cancer: does it really matter? *Curr Opin Oncol*. 2016 May;**28**(3):205–209.
3. Altieri B, Muscogiuri G, Barrea L, et al. Does vitamin D play a role in autoimmune endocrine disorders? A proof of concept. *Rev Endocr Metab Disord*. 2017 Sep;**18**(3):335–346.
4. Gois PHF, Ferreira D, Olenski S, Seguro AC. Vitamin D and Infectious Diseases: Simple Bystander or Contributing Factor? *Nutrients*. 2017 Jun 24;**9**(7).
5. Al Mheid I, Quyyumi AA. Vitamin D and Cardiovascular Disease: Controversy Unresolved. *J Am Coll Cardiol*. 2017 Jul 4;**70**(1):89–100.
6. Issa CM. Vitamin D and Type 2 Diabetes Mellitus. *Adv Exp Med Biol*. 2017;**996**:193–205.
7. Manson JE, Bassuk SS, Lee I-M, et al. 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*. 2012 Jan;**33**(1):159–171.
8. Neale RE, Armstrong BK, Baxter C, et al. The D-Health Trial: A randomized trial of vitamin D for prevention of mortality and cancer. *Contemp Clin Trials*. 2016;**48**:83–90.
9. Manson JE, Cook NR, Lee I-M, et al. Vitamin D Supplements and Prevention of Cancer and Cardiovascular Disease. *N Engl J Med*. 2018 Nov 10;
10. Wang TJ, Zhang F, Richards JB, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet Lond Engl*. 2010 Jul 17;**376**(9736):180–188.
11. Jiang X, O'Reilly PF, Aschard H, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat Commun*. 2018 Jan 17;**9**(1):260.
12. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003 Feb 1;**32**(1):1–22.
13. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. *Epidemiol Camb Mass*. 2014 May;**25**(3):427–435.
14. Mokry LE, Ross S, Ahmad OS, et al. Vitamin D and Risk of Multiple Sclerosis: A Mendelian Randomization Study. *PLoS Med*. 2015 Aug;**12**(8):e1001866.
15. Rhead B, Bäärnhielm M, Gianfrancesco M, et al. Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk. *Neurol Genet*. 2016 Oct;**2**(5):e97.
16. Gianfrancesco MA, Stridh P, Rhead B, et al. Evidence for a causal relationship between low vitamin D, high BMI, and pediatric-onset MS. *Neurology*. 2017 Apr 25;**88**(17):1623–1629.
17. Mokry LE, Ross S, Morris JA, Manousaki D, Forgetta V, Richards JB. Genetically decreased vitamin D and risk of Alzheimer disease. *Neurology*. 2016 Dec 13;**87**(24):2567–2574.
18. Larsson SC, Traylor M, Markus HS, Michaëlsson K. Serum Parathyroid Hormone, 25-Hydroxyvitamin D, and Risk of Alzheimer's Disease: A Mendelian Randomization Study. *Nutrients*. 2018 Sep 6;**10**(9).
19. Ong J-S, Cuellar-Partida G, Lu Y, et al. Association of vitamin D levels and risk of ovarian cancer: a Mendelian randomization study. *Int J Epidemiol*. 2016;**45**(5):1619–1630.

20. Afzal S, Brøndum-Jacobsen P, Bojesen SE, Nordestgaard BG. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. *BMJ*. 2014 Nov 18;**349**:g6330.
21. Dimitrakopoulou VI, Tsilidis KK, Haycock PC, et al. Circulating vitamin D concentration and risk of seven cancers: Mendelian randomisation study. *BMJ*. 2017 31;**359**:j4761.
22. Vimaleswaran KS, Cavadino A, Berry DJ, et al. Association of vitamin D status with arterial blood pressure and hypertension risk: a mendelian randomisation study. *Lancet Diabetes Endocrinol*. 2014 Sep;**2**(9):719–729.
23. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG, EPIC- InterAct Consortium. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol*. 2015 Jul;**30**(7):543–552.
24. Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018 May;**50**(5):693–698.
25. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015 Apr;**44**(2):512–525.
26. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016 May;**40**(4):304–314.
27. Bowden J, Del Greco M F, Minelli C, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int J Epidemiol* [Internet]. [cited 2019 Jan 3]; Available from: https://academic.oup.com/ije/advance-article/doi/10.1093/ije/dyy258/5251908
28. Gage SH, Jones HJ, Burgess S, et al. Assessing causality in associations between cannabis use and schizophrenia risk: a two-sample Mendelian randomization study. *Psychol Med*. 2017 Apr;**47**(5):971–980.
29. Hartwig FP, Borges MC, Horta BL, Bowden J, Davey Smith G. Inflammatory Biomarkers and Risk of Schizophrenia: A 2-Sample Mendelian Randomization Study. *JAMA Psychiatry*. 2017 01;**74**(12):1226–1233.
30. Vimaleswaran KS, Berry DJ, Lu C, et al. Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. *PLoS Med*. 2013;**10**(2):e1001383.
31. Lund-Nielsen J, Vedel-Krogh S, Kobylecki CJ, Brynskov J, Afzal S, Nordestgaard BG. Vitamin D and Inflammatory Bowel Disease: Mendelian Randomization Analyses in the Copenhagen Studies and UK Biobank. *J Clin Endocrinol Metab*. 2018 Sep 1;**103**(9):3267–3277.
32. Manousaki D, Paternoster L, Standl M, et al. Vitamin D levels and susceptibility to asthma, elevated immunoglobulin E levels, and atopic dermatitis: A Mendelian randomization study. *PLoS Med*. 2017 May;**14**(5):e1002294.
33. Bae S-C, Lee YH. Vitamin D level and risk of systemic lupus erythematosus and rheumatoid arthritis: a Mendelian randomization. *Clin Rheumatol*. 2018 Sep;**37**(9):2415–2421.
34. Manousaki D, Mokry LE, Ross S, Goltzman D, Richards JB. Mendelian Randomization Studies Do Not Support a Role for Vitamin D in Coronary Artery Disease. *Circ Cardiovasc Genet*. 2016 Aug;**9**(4):349–356.

35. Brøndum-Jacobsen P, Benn M, Afzal S, Nordestgaard BG. No evidence that genetically reduced 25-hydroxyvitamin D is associated with increased risk of ischaemic heart disease or myocardial infarction: a Mendelian randomization study. *Int J Epidemiol*. 2015 Apr;**44**(2):651–661.
36. Ooi EM, Afzal S, Nordestgaard BG. Elevated remnant cholesterol in 25-hydroxyvitamin D deficiency in the general population: Mendelian randomization study. *Circ Cardiovasc Genet*. 2014 Oct;**7**(5):650–658.
37. Afzal S, Brøndum-Jacobsen P, Bojesen SE, Nordestgaard BG. Vitamin D concentration, obesity, and risk of diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol*. 2014 Apr;**2**(4):298–306.
38. Lu L, Bennett DA, Millwood IY, et al. Association of vitamin D with risk of type 2 diabetes: A Mendelian randomisation study in European and Chinese adults. *PLoS Med*. 2018 May;**15**(5):e1002566.
39. Husemoen LLN, Skaaby T, Martinussen T, et al. Investigating the causal effect of vitamin D on serum adiponectin using a Mendelian randomization approach. *Eur J Clin Nutr*. 2014 Feb;**68**(2):189–195.
40. Maddock J, Zhou A, Cavadino A, et al. Vitamin D and cognitive function: A Mendelian randomisation study. *Sci Rep*. 2017 Oct 16;**7**(1):13230.
41. Michaëlsson K, Melhus H, Larsson SC. Serum 25-Hydroxyvitamin D Concentrations and Major Depression: A Mendelian Randomization Study. *Nutrients*. 2018 Dec 15;**10**(12).
42. Ye Z, Sharp SJ, Burgess S, et al. Association between circulating 25-hydroxyvitamin D and incident type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol*. 2015 Jan;**3**(1):35–42.

Figure 1. Estimated causal effects of 25(OH)D on various traits and diseases based on the MR-PRESSO approach using genome-wide significant SNPs of 25(OH)D.

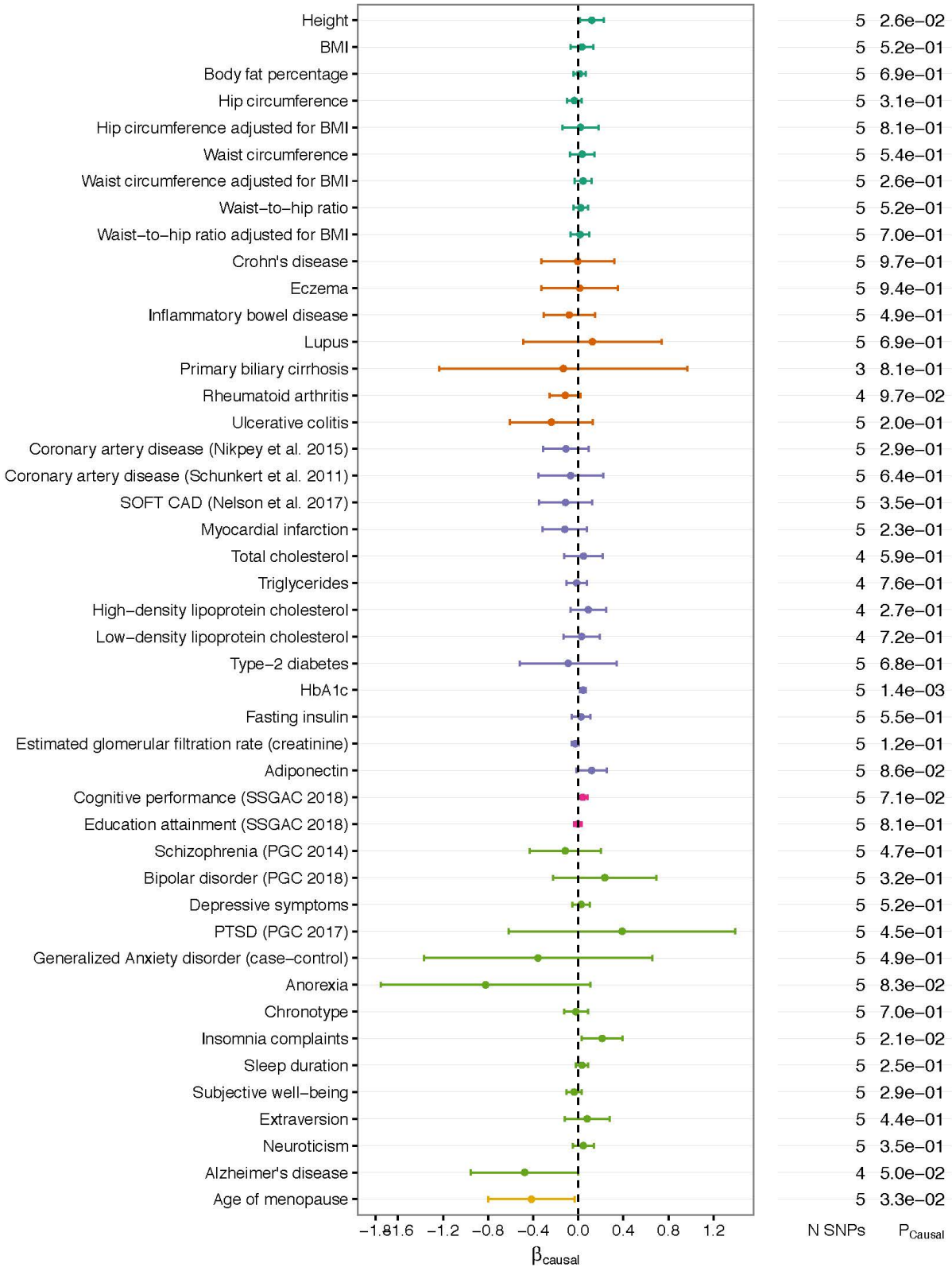

Figure 2. Scatter plots of effect sizes of IV-exposure vs. IV-outcome for nominally significant trait pairs in inverse-variance weighted meta-analysis. (a) height; (b) HbA1c; (c) insomnia complaints; (d) age of menopause

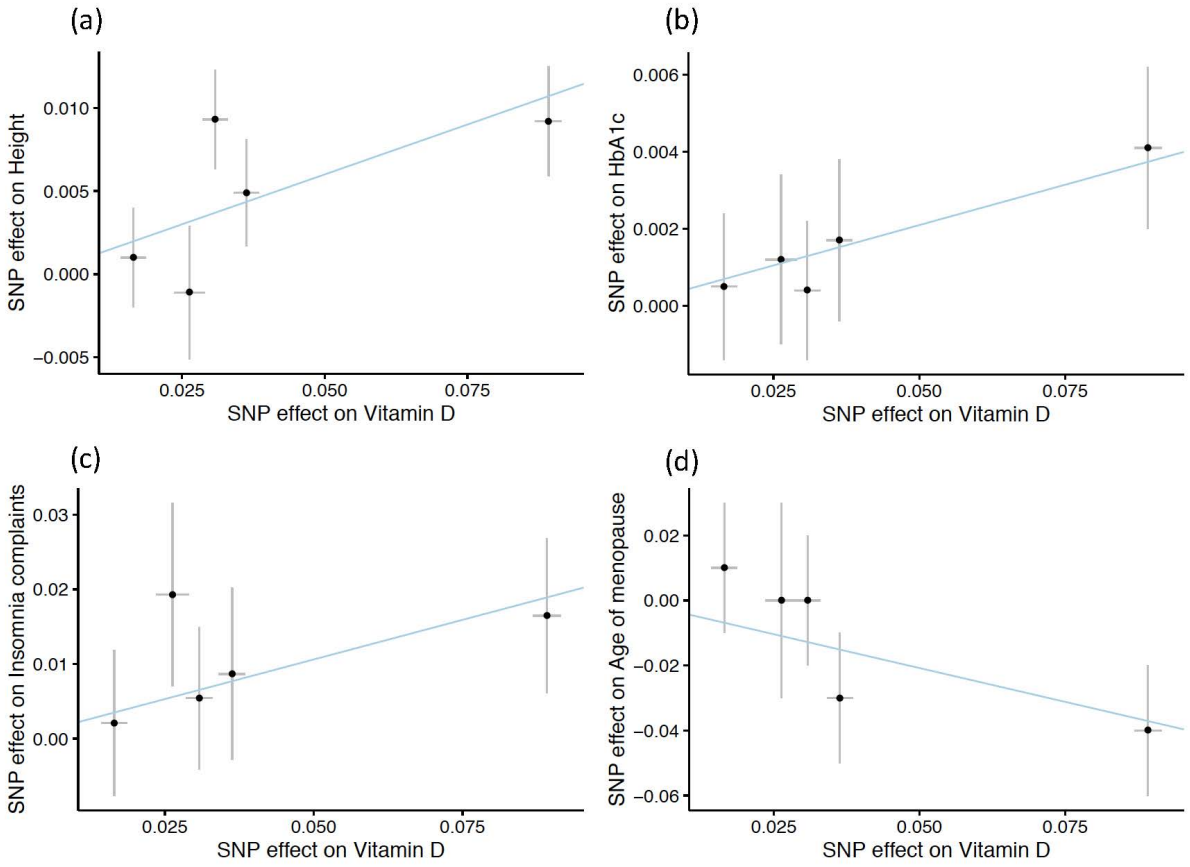

Table 1. Causal effect estimates of 25(OH)D with complex traits and diseases using genome-wide significant SNPs of 25(OH)D. We highlighted all nominally significant results, including 4 vitamin D-outcome pairs from IVW meta-analysis depicted in scatter plots in Figure 2. Note that none of these nominally significant results remained significant after multiple comparison correction.

Outcome	IVW meta-analysis			MR-Egger			MR-PRESSO adjusted			Median estimate		
	BETA	95% CI	P	BETA	95% CI	P	BETA	95% CI	P	BETA	95% CI	P
Height	**0.12**	**(0.01, 0.23)**	**0.03**	0.10	(-0.07, 0.27)	0.26	**0.12**	**(0.01, 0.23)**	**0.03**	**0.11**	**(0.01, 0.21)**	**0.03**
BMI	0.03	(-0.07, 0.13)	0.52	0.04	(-0.17, 0.26)	0.71	0.03	(-0.07, 0.13)	0.52	0.02	(-0.07, 0.11)	0.68
Body fat percentage	0.01	(-0.04, 0.07)	0.69	0.00	(-0.12, 0.11)	0.94	0.01	(-0.04, 0.07)	0.69	0.00	(-0.12, 0.13)	0.96
Hip circumference	0.02	(-0.15, 0.03)	0.85	-0.04	(-0.40, 0.32)	0.85	-0.03	(-0.10, 0.03)	0.31	-0.02	(-0.12, 0.09)	0.75
Hip circumference adjusted for BMI	0.02	(-0.14, 0.18)	0.82	0.02	(-0.33, 0.37)	0.92	0.02	(-0.14, 0.18)	0.82	0.01	(-0.10, 0.13)	0.80
Waist circumference	0.03	(-0.08, 0.14)	0.54	0.00	(-0.23, 0.22)	0.97	0.03	(-0.08, 0.14)	0.54	0.01	(-0.09, 0.12)	0.83
Waist circumference adjusted for BMI	0.04	(-0.03, 0.12)	0.26	0.01	(-0.13, 0.15)	0.85	0.04	(-0.03, 0.12)	0.26	0.03	(-0.07, 0.13)	0.54
Waist-to-hip ratio	0.02	(-0.04, 0.09)	0.52	-0.02	(-0.15, 0.11)	0.80	0.02	(-0.04, 0.09)	0.52	0.02	(-0.07, 0.12)	0.62
Waist-to-hip ratio adjusted for BMI	0.02	(-0.07, 0.10)	0.70	0.01	(-0.17, 0.19)	0.93	0.02	(-0.07, 0.10)	0.70	0.01	(-0.09, 0.12)	0.82
Crohn's disease	-0.01	(-0.33, 0.32)	0.98	**0.45**	**(0.08, 0.82)**	**0.02**	-0.01	(-0.33, 0.32)	0.98	-0.02	(-0.57, 0.53)	0.95
Eczema	0.01	(-0.33, 0.35)	0.94	0.40	(-0.18, 0.97)	0.18	0.01	(-0.33, 0.35)	0.94	0.08	(-0.35, 0.51)	0.72
Inflammatory bowel disease	-0.08	(-0.31, 0.15)	0.49	0.12	(-0.27, 0.51)	0.55	-0.08	(-0.31, 0.15)	0.49	-0.03	(-0.44, 0.39)	0.90
Lupus	0.48	(-0.73, 0.74)	0.44	0.33	(-2.20, 2.86)	0.80	0.13	(-0.49, 0.74)	0.69	0.27	(-0.44, 0.98)	0.46
Primary biliary cirrhosis	-0.13	(-1.23, 0.97)	0.81	-1.49	(-3.18, 0.19)	0.08	-0.13	(-1.23, 0.97)	0.81	-0.14	(-0.95, 0.66)	0.73
Rheumatoid arthritis	-0.12	(-0.25, 0.02)	0.10	-0.14	(-0.45, 0.18)	0.40	-0.12	(-0.25, 0.02)	0.10	-0.11	(-0.58, 0.37)	0.66
Ulcerative colitis	-0.24	(-0.61, 0.13)	0.20	-0.13	(-0.82, 0.57)	0.72	-0.24	(-0.61, 0.13)	0.20	-0.20	(-0.73, 0.33)	0.47
Coronary artery disease (Nikpey et al. 2015)	-0.11	(-0.31, 0.09)	0.29	-0.20	(-0.62, 0.21)	0.34	-0.11	(-0.31, 0.09)	0.29	-0.11	(-0.34, 0.12)	0.34
Coronary artery disease (Schunkert et al. 2011)	-0.07	(-0.36, 0.22)	0.64	-0.07	(-0.69, 0.55)	0.82	-0.07	(-0.36, 0.22)	0.64	-0.08	(-0.40, 0.25)	0.63
SOFT CAD (Nelson et al. 2017)	-0.11	(-0.35, 0.12)	0.35	-0.14	(-0.63, 0.35)	0.56	-0.11	(-0.35, 0.12)	0.35	-0.13	(-0.34, 0.09)	0.25
Myocardial infarction	-0.12	(-0.31, 0.08)	0.23	-0.19	(-0.59, 0.21)	0.35	-0.12	(-0.31, 0.08)	0.23	-0.12	(-0.37, 0.14)	0.36

Total cholesterol	0.05	(-0.12, 0.22)	0.59	-0.10	(-0.45, 0.26)	0.60	0.05	(-0.12, 0.22)	0.59	-0.02	(-0.14, 0.11)	0.79
Triglycerides	-0.01	(-0.11, 0.08)	0.76	**0.12**	**(0.02, 0.21)**	**0.01**	-0.01	(-0.11, 0.08)	0.76	-0.01	(-0.12, 0.10)	0.90
High-density lipoprotein cholesterol	0.09	(-0.07, 0.25)	0.27	-0.07	(-0.31, 0.17)	0.56	0.09	(-0.07, 0.25)	0.27	0.06	(-0.07, 0.19)	0.35
Low-density lipoprotein cholesterol	0.03	(-0.13, 0.19)	0.72	-0.14	(-0.45, 0.18)	0.39	0.03	(-0.13, 0.19)	0.72	-0.02	(-0.15, 0.11)	0.79
Type-2 diabetes	-0.09	(-0.52, 0.34)	0.68	0.48	(-0.06, 1.02)	0.08	-0.09	(-0.52, 0.34)	0.68	-0.03	(-0.34, 0.28)	0.86
HbA1c	**0.04**	**(0.02, 0.07)**	**0.001**	**0.05**	**(0.01, 0.09)**	**0.01**	**0.04**	**(0.02, 0.07)**	**0.001**	0.05	(-0.01, 0.10)	0.09
Fasting insulin	0.03	(-0.06, 0.11)	0.55	0.04	(-0.14, 0.21)	0.68	0.03	(-0.06, 0.11)	0.55	0.04	(-0.05, 0.13)	0.41
Estimated glomerular filtration rate (creatinine)	-0.04	(-0.12, 0.01)	0.29	0.01	(-0.12, 0.14)	0.89	-0.03	(-0.06, 0.01)	0.12	-0.03	(-0.06, 0.00)	0.06
Adiponectin	0.12	(-0.02, 0.25)	0.09	0.09	(-0.15, 0.33)	0.47	0.12	(-0.02, 0.25)	0.09	0.11	(-0.02, 0.24)	0.09
Cognitive performance (SSGAC 2018)	0.04	(0.00, 0.08)	0.07	0.04	(-0.04, 0.12)	0.34	0.04	(0.00, 0.08)	0.07	0.04	(-0.03, 0.12)	0.25
Education attainment (SSGAC 2018)	0.00	(-0.04, 0.03)	0.81	-0.02	(-0.08, 0.05)	0.62	0.00	(-0.04, 0.03)	0.81	-0.01	(-0.05, 0.03)	0.67
Schizophrenia (PGC 2014)	-0.12	(-0.43, 0.20)	0.47	0.03	(-0.59, 0.65)	0.92	-0.12	(-0.43, 0.20)	0.47	-0.07	(-0.34, 0.21)	0.64
Bipolar disorder (PGC 2018)	0.08	(-0.50, 0.69)	0.78	0.29	(-0.96, 1.53)	0.65	0.23	(-0.22, 0.69)	0.32	0.22	(-0.18, 0.62)	0.27
Depressive symptoms	0.03	(-0.05, 0.10)	0.52	0.01	(-0.15, 0.18)	0.87	0.03	(-0.05, 0.10)	0.52	0.03	(-0.07, 0.12)	0.59
PTSD (PGC 2017)	0.39	(-0.62, 1.39)	0.45	-0.30	(-2.14, 1.53)	0.75	0.39	(-0.62, 1.39)	0.45	0.16	(-0.79, 1.12)	0.74
Generalized anxiety disorder (case-control)	-0.36	(-1.37, 0.65)	0.49	0.28	(-1.62, 2.17)	0.78	-0.36	(-1.37, 0.65)	0.49	-0.40	(-1.12, 0.33)	0.28
Anorexia	-0.82	(-1.75, 0.11)	0.08	-1.65	(-3.27, -0.02)	0.05	-0.82	(-1.75, 0.11)	0.08	**-0.96**	**(-1.90, -0.02)**	**0.04**
Chronotype	-0.02	(-0.13, 0.08)	0.70	0.11	(-0.04, 0.26)	0.15	-0.02	(-0.13, 0.08)	0.70	0.00	(-0.10, 0.10)	0.99
Insomnia complaints	**0.21**	**(0.03, 0.39)**	**0.02**	0.15	(-0.12, 0.42)	0.28	**0.21**	**(0.03, 0.39)**	**0.02**	0.19	(-0.07, 0.44)	0.15
Sleep duration	0.03	(-0.02, 0.09)	0.25	0.02	(-0.09, 0.13)	0.75	0.03	(-0.02, 0.09)	0.25	0.03	(-0.07, 0.13)	0.53
Subjective well-being	-0.02	(-0.15, 0.03)	0.78	-0.14	(-0.38, 0.10)	0.26	-0.04	(-0.11, 0.03)	0.29	-0.03	(-0.10, 0.04)	0.37
Extraversion	0.08	(-0.12, 0.28)	0.44	-0.19	(-0.38, 0.01)	0.07	0.08	(-0.12, 0.28)	0.44	0.03	(-0.10, 0.17)	0.62
Neuroticism	0.04	(-0.05, 0.14)	0.35	0.07	(-0.11, 0.26)	0.45	0.04	(-0.05, 0.14)	0.35	0.04	(-0.05, 0.13)	0.41
Alzheimer's disease	-0.48	(-0.95, 0.00)	0.05	-0.20	(-0.93, 0.52)	0.58	-0.48	(-0.95, 0.00)	0.05	-0.42	(-0.91, 0.08)	0.10
Age of menopause	**-0.42**	**(-0.80, -0.03)**	**0.03**	-0.64	(-1.33, 0.06)	0.07	**-0.42**	**(-0.80, -0.03)**	**0.03**	-0.44	(-0.96, 0.08)	0.10

Table 2. Test for potential biases in MR analysis using genome-wide significant SNPs for 25(OH)D. We highlighted all nominally significant results. Note that only Estimated glomerular filtration rate (creatinine) showed significant pleiotropic bias after multiple comparison correction.

Outcome	MR-Egger intercept test p-value	MR-PRESSO global test p-value	N of outliers	Modified Q test p-value	Modified Q' test p-value
Height	0.75	0.39	0	0.30	0.31
BMI	0.93	0.40	0	0.21	0.22
Body fat percentage	0.75	0.92	0	0.93	0.93
Hip circumference	0.74	0.07	1	**0.01**	**0.02**
Hip circumference adjusted for BMI	1.00	0.12	0	**0.03**	**0.03**
Waist circumference	0.69	0.38	0	0.26	0.29
Waist circumference adjusted for BMI	0.63	0.79	0	0.71	0.74
Waist-to-hip ratio	0.50	0.85	0	0.75	0.80
Waist-to-hip ratio adjusted for BMI	0.92	0.62	0	0.54	0.54
Crohn's disease	**0.01**	0.39	0	0.79	1.00
Eczema	0.14	0.16	0	0.51	0.85
Inflammatory bowel disease	0.27	0.69	0	0.84	0.92
Lupus	0.89	**0.04**	1	**0.01**	**0.01**
Primary biliary cirrhosis	0.08	**0.05**	0	0.15	0.97
Rheumatoid arthritis	0.88	0.99	0	0.98	0.98
Ulcerative colitis	0.70	0.83	0	0.74	0.76
Coronary artery disease (Nikpey et al. 2015)	0.58	0.66	0	0.49	0.55
Coronary artery disease (Schunkert et al. 2011)	0.99	0.53	0	0.47	0.47
SOFT CAD (Nelson et al. 2017)	0.88	0.30	0	0.19	0.19
Myocardial infarction	0.66	0.78	0	0.65	0.69
Total cholesterol	0.40	**0.04**	0	0.09	0.25
Triglycerides	**0.01**	0.24	0	0.52	1.00
High-density lipoprotein cholesterol	0.21	0.05	0	0.16	0.58
Low-density lipoprotein cholesterol	0.27	0.05	0	0.15	0.48
Type-2 diabetes	**0.04**	**0.01**	0	**0.04**	0.77
HbA1c	0.32	0.99	0	0.99	1.00
Fasting insulin	0.88	0.51	0	0.42	0.42
Estimated glomerular filtration rate (creatinine)	0.41	**<0.001**	1	**2.52E-10**	**6.05E-08**
Adiponectin	0.77	0.23	0	0.29	0.31
Cognitive performance (SSGAC 2018)	0.98	0.91	0	0.89	0.89
Education attainment (SSGAC 2018)	0.66	0.57	0	0.58	0.62
Schizophrenia (PGC 2014)	0.59	0.28	0	0.16	0.21
Bipolar disorder (PGC 2018)	0.71	**0.04**	1	**0.01**	**0.02**
Depressive symptoms	0.86	0.51	0	0.46	0.47
PTSD (PGC 2017)	0.41	0.08	0	0.25	0.39
Generalized Anxiety disorder (case-control)	0.46	0.13	0	**0.04**	0.09
Anorexia	0.15	0.29	0	0.24	0.67

Chronotype	0.07	**0.03**	0	0.21	0.80
Insomnia complaints	0.55	0.86	0	0.85	0.88
Sleep duration	0.74	0.93	0	0.85	0.86
Subjective well-being	0.25	**0.02**	1	**0.0016**	**0.04**
Extraversion	**0.02**	**0.01**	0	**0.04**	0.88
Neuroticism	0.70	0.46	0	0.29	0.32
Alzheimer's disease	0.43	0.50	0	0.48	0.65
Age of menopause	0.34	0.85	0	0.77	0.86