

# Genetically increased circulating 25(OH)D level reduces the risk of type 2 diabetes in subjects with deficiency of vitamin D

## A large-scale Mendelian randomization study

Yingying Xu, MM<sup>a,\*</sup>, Yuan Zhou, MS<sup>b</sup>, Jingjing Liu, MM<sup>a</sup>, Chenfang Wang, MM<sup>a</sup>, Zhongjie Qu, MM<sup>a</sup>, Zhili Wei, BD<sup>c</sup>, Dan Zhou, PhD<sup>d</sup>

### Abstract

Observational studies have reported that Vitamin D deficiency and the risk type 2 diabetes are associated, but the causation is unclear. Mendelian randomization (MR) involving genetic variants as instrument variables (IVs) overcomes the reverse-casualty and unmeasured confounding. However, with limited sample size and IVs, previous MR studies showed inconsistent results. Leveraging by a largely increased sample size for both stages, we aim to provide an updated and precise estimate for the causality between Vitamin D and type 2 diabetes.

A 2-sample multi-IVs MR was performed. IVs for circulating 25-hydroxyvitamin D (25(OH)D) were obtained from a genome-wide association study from UK biobank involving 329,247 subjects of European ancestry. The causal effect of 25(OH)D and type 2 diabetes was estimated using traditional inverse variance weighting and MR pleiotropy residual sum and outlier (MR-PRESSO) framework which provides a robust estimate by systematically filtering out IVs identified with potential pleiotropy effects.

A higher genetically instrumented 25(OH)D was causally linked to reduced risk of type 2 diabetes risk by MR-PRESSO [odds ratio (OR) per standard deviation (SD)=0.950, 95% confidence interval (CI)=0.913–0.988,  $P=.010$ ] after removing 13 (13/193) invalid IVs. In addition, we confirmed the causal role Vitamin D using 2 synthesis-related single-nucleotide polymorphisms (SNPs) which are consistent with previous MR studies [OR per SD=0.894, 95% CI=0.816–0.979,  $P=.016$ ].

With a largely improved sample size, our results confirmed that genetically increased 25(OH)D concentration reduced the risk of type 2 diabetes and provided a more precise estimate for the effect size. The updated result empowers the role of Vitamin D and provides nontrivial evidence for interventional studies.

**Abbreviations:** 25(OH)D = 25(OH)D, 95% CI = 95% confidential interval, DIAGRAM = DIAbetes Genetics Replication And Meta-analysis, GWAS = genome-wide association study, IVs = instrument variables, IWV = inverse variance weighted, LD = linkage disequilibrium, MR = Mendelian randomization, MR-PRESSO= Mendelian randomization pleiotropy residual sum and outlier, OR = odds ratio, RCT = randomized controlled trial, SD = standard deviation, SNP = single-nucleotide polymorphism, T2D = type 2 diabetes.

**Keywords:** Mendelian randomization, type 2 diabetes, vitamin D

Editor: Antonio Palazón-Bru.

The authors used only publicly available data and methods as specified in the manuscript.

The GWAS summary results for the exposure and outcome we used are publicly available. The GWAS results for 25(OH)D were downloaded from Neale Lab ([www.nealelab.is/uk-biobank/](http://www.nealelab.is/uk-biobank/)). The type 2 diabetes GWAS summary statistics were obtained from the DIAGRAM consortium ([www.diagram-consortium.org/](http://www.diagram-consortium.org/)). All the GWAS results were used with no attempt to deidentify individual subjects.

This work was supported by Zhejiang Chinese Medical Science Foundation (2015ZB075).

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

The datasets generated during and/or analyzed during the current study are publicly available.

<sup>a</sup> Department of Endocrine Rheumatism and nephrology, The Third Affiliated Hospital of Zhejiang Chinese Medical University, <sup>b</sup> Department of Pathology, Zhejiang University School of Medicine, <sup>c</sup> Zhejiang Chinese Medical University, Hangzhou, Zhejiang, China, <sup>d</sup> Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN.

\* Correspondence: Yingying Xu, Department of Endocrine Rheumatism and Nephrology, The Third Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, Zhejiang 310005, China (e-mail: [xyingying9797@163.com](mailto:xyingying9797@163.com)).

Copyright © 2020 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Xu Y, Zhou Y, Liu J, Wang C, Qu Z, Wei Z, Zhou D. Genetically increased circulating 25(OH)D level reduces the risk of type 2 diabetes in subjects with deficiency of vitamin D: a large scale Mendelian randomization study. *Medicine* 2020;99:51(e23672).

Received: 13 February 2020 / Received in final form: 3 September 2020 / Accepted: 10 November 2020

<http://dx.doi.org/10.1097/MD.00000000000023672>

## 1. Introduction

Type 2 diabetes (T2D) is a global health concern in the 21st century. The incidence of type 2 diabetes has increased substantially around the world in recent decades.<sup>[1–3]</sup> A large number of observational studies have reported that Vitamin D deficiency (another highly prevalent issue across the globe) and the onset and progression of type 2 diabetes are associated.<sup>[3–5]</sup> However, even with reasonable design, the associations observed from these studies could be ambiguous due to reverse causation, selection bias, and unmeasured confounding effects. On the other side, although randomized controlled trial (RCT) could evidence causality, the high cost in terms of time and money would limit it in practice, especially for weak causation inferencing which needs a large number of participants.

Mendelian randomization (MR) is a rapidly growing study for causality inference that uses genetic variants as instrumental variables (IVs) to test the causal effect of a risk factor on an outcome.<sup>[6,7]</sup> Although depending on observational data, it overcomes the major challenges of observational study by introducing IVs. On the one hand, the reversed causality will be completely avoided because of the nonreversible association from the germline variant to phenotype. On the other hand, as genotypes are assigned randomly during meiosis, by assuming that mate choice is not associated with genotype, the genotype distribution should be unrelated to the confounding factors. In this way, MR can be considered as a “natural” RCT. Therefore, MR is a cost-effective way to step in the causal relevance of Vitamin D levels for the risk of type 2 diabetes. So far, some MR had been done but ended up with inconsistent results.<sup>[8–10]</sup> Ye et al<sup>[9]</sup> showed that no obvious effect was observed, but another MR study reported that the lower Vitamin D levels significantly increased the risk of type 2 diabetes.<sup>[10]</sup> We noticed that some of the effects are marginally significant which indicated that a larger and more precise MR is needed to confirm the association. Thanks to the availability of the huge genome-wide association study (GWAS) for Vitamin D level from the UK biobank, we should be able to have more powerful IVs for inference. Meanwhile, the statistical methods for MR continuing improved in recent years.<sup>[11–13]</sup> The 2-sample MR largely boosted the field because only summary-statistic are needed.<sup>[14,15]</sup> MR-PRESSO, a powerful 2-sample MR framework with horizontal pleiotropy control, has 3 components including detection of horizontal pleiotropy, correction of horizontal pleiotropy *via* outlier removal, and testing of significant distortion in the causal estimates before and after outlier removal.<sup>[16]</sup>

In the current study, we performed a multi-IVs 2-sample MR to examine the causal association between Vitamin D levels and the risk of type 2 diabetes *via* MR-PRESSO based on the most recently released large Vitamin D GWAS results and the latest DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) T2D dataset.

## 2. Methods

The main steps of our study are presented in Figure 1. Instrumental variables were identified by Vitamin D-associated variants (genome-widely significant) after linkage disequilibrium (LD)-clumping. Two-sample MR was performed to test the causation between Vitamin D level and the risk of type 2 diabetes. MR-PRESSO was used to remove potential outliers that violate the basic assumption of MR.

### 2.1. GWAS for the exposure

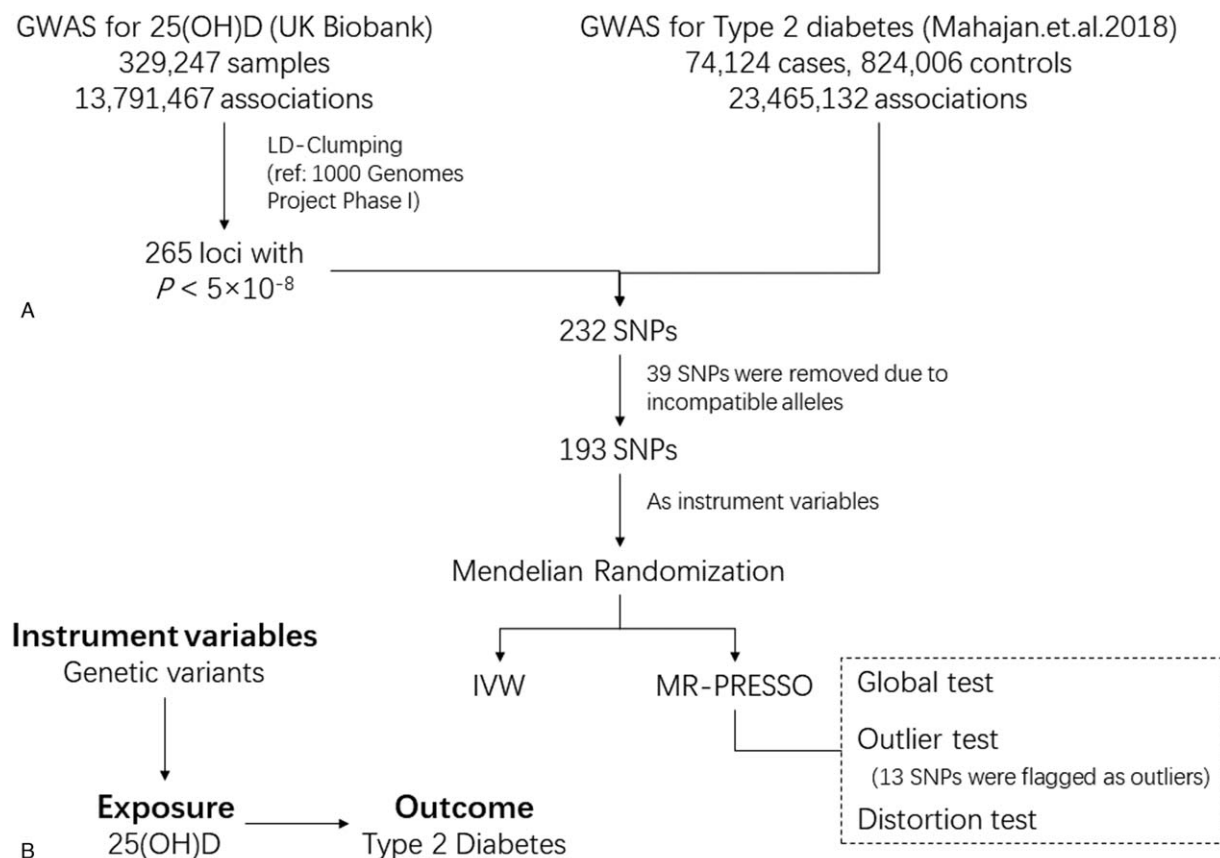
IVs were acquired from the GWAS of 25-hydroxyvitamin D (25(OH)D) from UK Biobank. UK Biobank, a large population-based open access health resource, recruit more than 500,000 volunteer participants from across the United Kingdom between 2006 and 2010.<sup>[17,18]</sup> Participants were collected phenotypic and genotypic detail for genotyping, biomarker, and other disease-related analyses.<sup>[19,20]</sup> The serum 25(OH)D levels (nmol/L) measurement was followed by CLIA and the UK Biobank’s quality control document accompanying the biomarker measurements.<sup>[19]</sup> A set of individuals who were genetically determined (K-means clustering) to be White-European and who also self-identified as White were included in this study. The participants aged between 40 and 69 (at biobank recruitment) who lived within 25 miles of one of the 22 assessment centers in England, Wales, and Scotland.<sup>[19]</sup> Age, age<sup>2</sup>, sex, age\*sex, sex\*age<sup>2</sup>, and PCs 1–20 were considered as covariates in the GWAS. To estimate whether “vitamin D intake,” a potential confounding factor, would affect the results. We performed a GWAS for “vitamin D intake.” The trait “vitamin D intake” was measured by asking “Do you regularly take Vitamin D?”. The same covariates were adjusted. The summary-statistic results were downloaded from Neale Lab (URL <http://www.nealelab.is/uk-biobank/>).

### 2.2. GWAS for the outcome

The latest type 2 diabetes GWAS summary statistics were downloaded from DIAGRAM Consortium (URL <https://www.diagram-consortium.org/downloads.html>). This meta-analysis collected GWAS results for 32 studies for 898,130 individuals (74,124 cases and 824,006 controls) of European ancestry.<sup>[21]</sup> The mean age range of the substudies is 32 to 75 years. Type 1 diabetes patients, ethnic outliers, cryptic first-degree relatives, and duplicates were excluded from the study.<sup>[21]</sup> Imputation for the 31 component studies was done using the Haplotype Reference Consortium reference panel, and the deCODE GWAS was imputed using a population-specific reference panel. To improve the quality of the genotype scaffold in each study, a harmonized protocol had been developed to minimize the heterogeneity (more details can be found in the original paper<sup>[21]</sup>). The population structure and relatedness were identified by principal components or a mixed model with random effects for kinship from a genetic-relationship matrix.<sup>[21]</sup> Gender, age, array, first 6 principal components, and study-specific covariates (more details can be found in the original paper) were adjusted.<sup>[21]</sup> In this study, we used the BMI-unadjusted association results to capture the potential mediation effect (not necessarily true) of BMI between Vitamin D and T2D, and to be consistent with the GWAS of 25(OH)D which did not include BMI as a covariate.

### 2.3. MR analysis

For each of Vitamin D associated ( $P < 5 \times 10^{-8}$ ) loci identified in the previous step, the single-nucleotide polymorphism (SNP) with the lowest  $P$  value was selected using the clumping function in TwoSampleMR Package (linkage disequilibrium threshold  $r^2 = 0.1$ , window size = 250 kb). The 1000 Genome Phase I CEU data was used as a reference to calculate linkage disequilibrium between the SNPs. The IV strength indicator  $F$ -statistic was



**Figure 1.** The flowchart of the study. A, A 2-sample Mendelian randomization (MR) was performed. The GWAS summary statistics of SNP-exposure (25(OH)D) and SNP-outcome (type 2 diabetes) were obtained from UK biobank and DIAGRAM, respectively. The causal effect of 25(OH)D on type 2 diabetes was estimated by MR-IVW and MR-PRESSO. B, The triangular plot of the MR. GWAS = genome-wide association study, IVW = inverse variance weighted, SNP = single-nucleotide polymorphism.

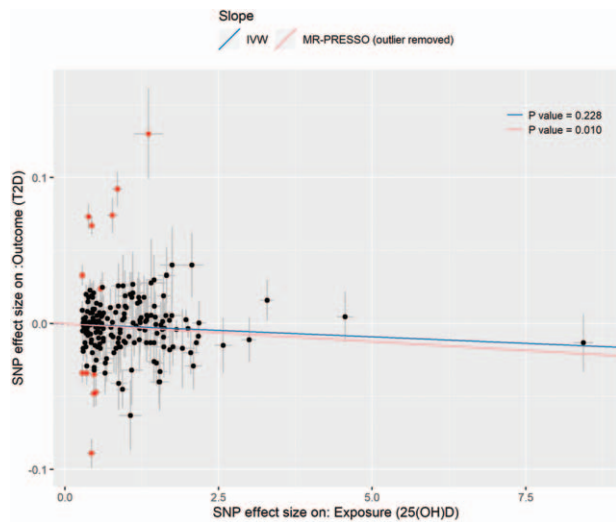
calculated as  $F = \frac{R^2(n-1-k)}{(1-R^2)k}$ .  $R^2$ ,  $n$ , and  $k$  represent the proportion of variance explained by the IV, the first stage sample size, and the number of IVs, respectively. The  $R^2$  was estimated by  $R^2 = \left(\beta \frac{2f(1-f)}{\sigma_Y}\right)^2$  where  $\beta$ ,  $f$ , and  $\sigma_Y$  respectively denote the coefficient, the minor allele frequency of IV, and the standard deviation of the outcome.

To match the effect allele of each SNP between the exposure and the outcome, harmonizing was performed using the `harmonise_data` function from the `TwoSampleMR` Package.<sup>[22]</sup> Then, 2-sample multi-IVs MR analysis was carried out using MR-PRESSO. The global IV validity was tested to check if any horizontal pleiotropy effect was detected which violates the major assumption of MR. Then, an outlier test was performed for each variant by comparing the observed residual sum of squares (RSS) to the expected RSS which yields an empirical  $P$  value. A distortion test was conducted to obtain a corrected slope. We also accessed the causal association using inverse-variance weighted (IVW)<sup>[23]</sup> as a pleiotropy unadjusted MR. The leave-one-out sensitivity method was conducted to identify influential individual SNPs under the conventional IVW model. In addition, we tested the causation with 25(OH)D concentration associated-SNPs (*DHCR7*-rs12785878, *CYP2R1*-rs10741657, *GC/DBP*-rs2282679, and *CYP24A1*-rs6013897) as IVs which are well established and commonly used in other studies.

All statistical analyses were conducted using R statistical software (version 3.6.1). The threshold of significance for analyses in this study was  $P < .05$ .

### 3. Results

Of all the UK biobank subjects, 329,247 participants were genetically identified as European with nonmissing 25(OH)D measurement and covariates. The mean and standard deviation of 25(OH)D level are 48.58 and 21.14 (nmol/L), respectively (Supplementary Figure 1, <http://links.lww.com/MD/F388>). 13,791,467 SNPs passed the postimputation quality control. The Manhattan plot of the GWAS was shown in Supplementary Figure 2, <http://links.lww.com/MD/F389>. Among the 265 genome-wide significant loci after LD-clumping, 232 were successfully harmonized with GWAS results of type 2 diabetes. Thirty-nine out of 232 SNPs that had incompatible alleles were removed (for example, a SNP has A/G alleles for the exposure and A/C alleles for the outcome). Finally, 193 SNPs remained as candidate IVs. The allele summary-statistic results of each SNP can be found in Supplementary Table 1, <http://links.lww.com/MD/F394>. To estimate whether “vitamin D intake,” a potential confounding factor would affect the results, we tested the associations between all of the 193 SNPs and the “vitamin D intake” by a GWAS using UK biobank samples (N of regular user: 13,687; N of none regular user: 345,558). None of them



**Figure 2.** The scatter plot of the valid and invalid IVs for 25[OH]D level (nmol/L) for MR. The x-axis and the y-axis denote the SNP effect sizes on the exposure (25[OH]D) and the outcome (type 2 diabetes), respectively. Each of the 193 IVs is plotted with error bars. IVs identified as outliers are flagged in red. The IVW estimates and the pleiotropy-adjusted MR-PRESSO estimate were shown.

were significantly associated with vitamin D intake ( $P_{FDR} < .05$ ).

By applying the global test from MR-PRESSO, significant horizontal pleiotropy ( $P = 2.56 \times 10^{-4}$ ) was indicated. Thirteen SNPs were identified as outliers and were removed in the downstream analysis (highlighted in Fig. 2). Of the 180 valid IVs, the  $F$ -statistic is 40.93 ( $>10$ ), suggesting that it is unlikely to be affected by weak instrument bias. The outliers located in the upper left (a bit more) and lower left of the scatter plot which resulted in a clockwise rotation of the new zero-crossing regression line. The rotated regression line with a steeper slope indicated that higher 25(OH)D concentration significantly reduced the risk of type 2 diabetes (OR per SD: 0.950, 95% CI: 0.913–0.988,  $P = .010$ ) (Fig. 2). However, no significant result was observed using IVW analysis ( $P = .228$ ). The scatter plot with the regression line from MR-PRESSO and IVW analyses was presented in (Fig. 2). In parallel, we performed the MR using quantile normalized 25(OH)D level as exposure. Results were consistent (Supplementary Figure 3, <http://links.lww.com/MD/F390> and Supplementary Figure 4, <http://links.lww.com/MD/F391>). The leave-one-out sensitivity analysis under the conventional IVW method was performed to evaluate the causal effects of 25(OH)D level on the risk of T2D using the summary-level data from GWASs (Supplementary Figure 5, <http://links.lww.com/MD/F392> and Supplementary Figure 6, <http://links.lww.com/MD/F393>). NO heterogeneous SNPs that largely influence the causal estimates for 25(OH)D on T2D were identified.

In addition, we performed MR leveraging four commonly used IVs for Vitamin D. As expected, the 4 SNPs showed robustly associated in UK Biobank samples (marked in Supplementary Figure 2, <http://links.lww.com/MD/F389>). *CYP24A1*-rs6013897 were filtered out due to incompatible alleles between UK biobank and DIAGRAM GWAS results. The IVW estimate did not support the causal role. However, by using the 2 synthesis-related SNPs (*DHCR7*-rs12785878, *CYP2R1*-rs10741657) as instru-

mental variables, we found the result (IVW OR per SD: 0.894, 95% CI: 0.816–0.979,  $P = .016$ ) was consistent with our MR using genome-wide noncandidate IVs.

#### 4. Discussion

MR is a fast-developing genetic epidemiology method beefed up by the expanding open-access GWAS dataset and novel statistical approaches. MR studies sit at the interface of RCT and observational studies. The reduced influence of confounding and reverse causality provides more reliable evidence to guide interventional studies and provides critical information when an RCT may not be feasible (or to reach a reasonable sample size when the effect size is small).<sup>[24,25]</sup> In this study, we performed a large-scale 2-sample Mendelian randomization analysis to detect the causal effect from genetically regulated Vitamin D level on type 2 diabetes in European populations. We found that genetically predicted 25(OH)D level was inversely associated with type 2 diabetes risk using 180 valid IVs. This result is unbiased and could be generalized to European populations. Further studies in other populations are needed.

To our knowledge, the horizontal pleiotropy effect is considered as the major weakness in MR practice as it might be easily violated. MR-PRESSO, one of the outlier-robust approach, best suited when horizontal pleiotropy occurs in  $< 50\%$  of instruments. A more precise estimate could be observed by MR-PRESSO after removing invalid SNPs labeled as outliers. IVW, without such a feature, would result in a biased estimate when the pleiotropy effect exists. With the motivation to compare the results from previous analyses, we also conducted IVW estimates using the only synthesis-related SNPs. In this case, because of the limited number of IVs, MR-PRESSO was not applicable. (The challenge of pleiotropy estimate with limited IVs might be solved by MRGxE,<sup>[26]</sup> a novel MR approach that involves a covariate which modifies the SNP effect on exposure.)

In agreement with our findings, a meta-analysis using 2 synthesis-related SNPs as IVs demonstrated that a higher plasma 25(OH)D concentration was associated with a lower risk of type 2 diabetes. However, some other MR studies have reported nonsignificant results with relatively small sample sizes, while MR studies always require a large number of samples to generate proper IVs and hold the power for causal inference. Taking advantage of the large-scale UK Biobank dataset, we dramatically increased the sample size from a few thousand (previous studies) to more than 329k (current study) for the GWAS of 25(OH)D levels. Together with the nearly 1 million type 2 diabetes GWAS samples, we provided the largest MR with a more precise estimate. The effect size is small but significant, and informative. Nevertheless, the observed association of the current MR study could not be directly interpreted as causality without further functional studies.

Some researches had studied the mechanisms for the effects of Vitamin D on type 2 diabetes.<sup>[5,27]</sup> Human pancreatic cell lines and rats experiments suggested that elevated Vitamin D levels improved insulin status and regulated calcium flux through the  $\beta$  cell.<sup>[5,28]</sup> Calcium is essential for insulin production and secretion.<sup>[5,28]</sup> These studies showed increasing Vitamin D levels might be a potential solution to decrease the risk of type 2 diabetes through strengthening  $\beta$  cell function. Another potential mechanism of the causation is mediated by affecting insulin resistance.<sup>[29]</sup> Vitamin D may reduce excessive insulin release by reducing the insulin resistance in the surrounding tissues.<sup>[30]</sup> The

study also showed that Vitamin D supplementation will decrease parathyroid hormone levels which will enhance insulin sensitivity, lipolysis, and lipogenesis controlling.<sup>[31]</sup> Furthermore, Vitamin D-binding protein plays a role in insulin resistance.<sup>[32]</sup> The role of Vitamin D on insulin resistance level was validated by a recent RCT which found that Vitamin D supplementation administered for 12 months in healthy men maintained insulin levels and HOMA-IR values relative to the increase in the placebo group.<sup>[33]</sup> It indicated that, in the short term, Vitamin D plays a role to maintain insulin sensitivity. As we know, insulin resistance starts in the early stage of developing T2D, that is T2D is a long-term phenotypic consequence of insulin resistant.<sup>[34]</sup> It could (at least partially) explain the negative result of the Vitamin D intervention on T2D prevention in prediabetes subjects from a recent RCT.<sup>[35]</sup> If the major role of Vitamin D is to maintain insulin sensitivity, it is reasonable to observe a negative result with only 2.5 years (median) follow-up time in prediabetes subjects. Nevertheless, the hazard ratio (HR) for vitamin D as compared with placebo was promising (HR=0.88, 95% CI, 0.75–1.04).<sup>[35]</sup> According to our study, the preventive effect of Vitamin D on the risk of T2D is small. It is possible that the current large-scale RCT is still underpowered, especially within a limited follow-up time. Compared to RCT, MR study is more cost-effective since the germline variants predisposed the exposure (Vitamin D) in the beginning of life and the effect would last during lifetime. That means the participants in the MR study had been randomly assigned, according to the Mendel's second law, to groups with different "treatment" of Vitamin D levels for decades.

The current study has some inevitable limitations. We were not being able to exclude the subjects with Vitamin D supplementation since we only have access to the GWAS-summary-statistics. However, the Vitamin D-mediated effect of genetic variants on the behavior of Vitamin D supplementation intake is very small. Once the correlation of exposure (genetic variants) and supplementation intake is negligible, the supplementation intake will not be considered as a confounder between SNP and 25(OH)D levels. We were not able to adjust the season of 25(OH)D measurement, physical activities, and other outcome-related traits. As we mentioned above, because the correlations between 25(OH)D-associated SNPs and these "potential confounders" are not existing (season) or negligible (physical activities), these "potential confounders" will not be considered as confounders. Without adjusting these outcome-related traits, the residual variance will be high. It will result in a lower power to detect the association. However, as we discussed above, the estimate will not be biased. Another limitation is that our result only estimated the general effect of 25(OH)D, the effect size may be sensitive to season. Although we only included European ancestry population for this study, the results may be affected due to the potential heterogeneity of the 2 cohorts.

## 5. Conclusions

With much larger GWAS datasets for both stages, we provided reliable evidence that the SNPs associated with the levels of 25(OH)D were relevant to the SNP associated with T2D in European ancestry population. It indicated that genetically increased 25(OH)D level is associated with lower susceptibility to type 2 diabetes in a population with low Vitamin D levels. It is an update with the latest data which provides valuable evidence to guide interventional studies in public health.

## Acknowledgments

The authors are grateful to all the study subjects for their kind participation.

## Author contributions

**Conceptualization:** Yingying Xu and Dan Zhou.

**Data curation:** Yuan Zhou and Chenfang Wang.

**Formal analysis:** Yuan Zhou, Dan Zhou, Jingjing Liu.

**Funding acquisition:** Yingying Xu.

**Investigation:** Yingying Xu, Zhongjie Qu, and Zhili Wei.

**Methodology:** Yingying Xu.

**Project administration:** Yingying Xu.

**Supervision:** Yingying Xu.

**Writing – original draft:** Yingying Xu, Dan Zhou and Yuan Zhou.

**Writing – review & editing:** Yingying Xu, Dan Zhou and Yuan Zhou.

## References

- Forouhi NG, Wareham NJ. Epidemiology of diabetes. *Medicine* 2014;42:698–702.
- de Sousa-Uva M, Antunes L, Nunes B, et al. Trends in diabetes incidence from 1992 to 2015 and projections for 2024: a Portuguese General Practitioner's Network study. *Primary Care Diabetes* 2016;10:329–33.
- Wu Y, Ding Y, Tanaka Y, et al. Risk factors contributing to type 2 diabetes and recent advances in the treatment and prevention. *Int J Med Sci* 2014;11:1185–200.
- Mitri J, Muraru MD, Pittas AG. Vitamin D and type 2 diabetes: a systematic review. *Eur J Clin Nutr* 2011;65:1005–15.
- Pittas AG, Lau J, Hu FB, et al. The role of vitamin D and calcium in type 2 diabetes. A systematic review and meta-analysis. *J Clin Endocrinol Metab* 2007;92:2017–29.
- Lawlor DA, Harbord RM, Sterne JA, et al. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;27:1133–63.
- Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;32:1–22.
- Afzal S, Brondum-Jacobsen P, Bojesen SE, et al. Vitamin D concentration, obesity, and risk of diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol* 2014;2:298–306.
- 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;3:35–42.
- 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;15:e1002566.
- Boef AG, Dekkers OM, le Cessie S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. *Int J Epidemiol* 2015;44:496–511.
- Sekula P, Del Greco MF, Pattaro C, et al. Mendelian randomization as an approach to assess causality using observational data. *J Am Soc Nephrol* 2016;27:3253–65.
- Zheng J, Baird D, Borges MC, et al. Recent developments in Mendelian randomization studies. *Curr Epidemiol Rep* 2017;4:330–45.
- Lawlor DA. Commentary: two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol* 2016;45:908–15.
- Hartwig FP, Davies NM, Hemani G, et al. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol* 2016;45:1717–26.
- Verbanck M, Chen CY, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 2018;50:693–8.
- Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
- Cox N. UK Biobank shares the promise of big data. *Nature* 2018;562:194–5.

- [19] Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–9.
- [20] Bahcall OG. UK Biobank—a new era in genomic medicine. *Nat Rev Genet* 2018;19:737.
- [21] Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;50:1505–13.
- [22] Hemani G, Zheng J, Elsworth B, et al. The MR-base platform supports systematic causal inference across the human phenome. *Elife* 2018;7:e34408.
- [23] Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;37:658–65.
- [24] Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;362:k601.
- [25] Gulseth HL, Wium C, Angel K, et al. Effects of vitamin D supplementation on insulin sensitivity and insulin secretion in subjects with type 2 diabetes and vitamin D deficiency: a randomized controlled trial. *Diabetes Care* 2017;40:872–8.
- [26] Spiller W, Slichter D, Bowden J, et al. Detecting and correcting for bias in Mendelian randomization analyses using Gene-by-Environment interactions. *Int J Epidemiol* 2019;48:702–12.
- [27] Halban PA, Polonsky KS, Bowden DW, et al. beta-cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. *Diabetes Care* 2014;37:1751–8.
- [28] Clark SA, Stumpf WE, Sar M. Effect of 1,25 dihydroxyvitamin D3 on insulin secretion. *Diabetes* 1981;30:382–6.
- [29] Chiu KC, Chu A, Go VLW, et al. Hypovitaminosis D is associated with insulin resistance and  $\beta$  cell dysfunction. *Am J Clin Nutr* 2004;79:820–5.
- [30] Kavadar G, Demircioğlu DT, Özgönenel L, et al. The relationship between vitamin D status, physical activity and insulin resistance in overweight and obese subjects. *Bosn J Basic Med Sci* 2015;15:62–6.
- [31] Soares M, Pathak K, Calton E. Calcium and vitamin D in the regulation of energy balance: where do we stand? *Int J Molecular Sci* 2014;15:4938–45.
- [32] Hirai M, Suzuki S, Hinokio Y, et al. Variations in vitamin D-binding protein (group-specific component protein) are associated with fasting plasma insulin levels in Japanese with normal glucose tolerance. *J Clin Endocrinol Metab* 2000;85:1951–3.
- [33] Tepper S, Shahar D, Geva D, et al. Differences in homeostatic model assessment (HOMA) values and insulin levels after vitamin D supplementation in healthy men: a double-blind randomized controlled trial. *Diabetes Obesity Metabolism* 2016;18:633–7.
- [34] Lillioja S, Mott DM, Spraul M, et al. Insulin resistance and insulin secretory dysfunction as precursors of non-insulin-dependent diabetes mellitus: prospective studies of Pima Indians. *New Engl J Med* 1993;329:1988–92.
- [35] Pittas AG, Dawson-Hughes B, Sheehan P, et al. Vitamin D supplementation and prevention of type 2 diabetes. *New Engl J Med* 2019;381:520–30.