Article

Serum 25-hydroxyvitamin D Levels and Youth-onset Type 2 Diabetes: A Two-sample Mendelian Randomization Study

Benjamin De La Barrera and Despoina Manousaki [*]

Article

Serum 25-hydroxyvitamin D Levels and youth-Onset Type 2 Diabetes: A Two-Sample Mendelian Randomization Study

Benjamin De La Barrera [1] and Despoina Manousaki [1,2,*]

[1] Research Center of the Sainte-Justine University Hospital, University of Montreal, Montreal, Quebec, Canada

[2] Departments of Pediatrics, Biochemistry and Molecular Medicine, University of Montreal, Montreal, Canada

* Correspondence: despina.manousaki@umontreal.ca; Tel.: +1-514-345-4931 (ext. 4735)

Abstract: Observational studies have linked vitamin D insufficiency to pediatric type 2 diabetes (T2D), but evidence from vitamin D supplementation trials is sparse. Given the rising prevalence of pediatric T2D in all ethnicities, determining a protective role of vitamin D has significant public health importance. We tested whether serum 25-hydroxyvitamin D (25OHD) levels are causally linked to youth-onset T2D risk using Mendelian randomization (MR). We selected 54 single nucleotide polymorphisms (SNPs) associated with 25OHD in a European genome-wide association study (GWAS) on 443,734 individuals and obtained their effects on pediatric T2D from the multi-ethnic PRODIGY GWAS (3,006 cases/6,061 controls). We applied inverse variance weighted (IVW) MR, and a series of MR methods to control for pleiotropy. We undertook sensitivity analyses in ethnic sub-cohorts of PRODIGY, and using SNPs in core vitamin D genes or ancestry-informed 25OHD SNPs. Multivariable MR accounted for mediating effects of body mass index. We found that a standard deviation increase in 25OHD in the logarithmic scale did not affect youth-onset T2D risk (IVW MR odds ratio (OR) = 1.04, 95% CI=0.96-1.13, P=0.35) in the multi-ethnic analysis, and sensitivity, ancestry-specific and multivariable MR analyses showed consistent results. Our study had limited power to detect small/moderate effects of 25OHD (OR of pediatric T2D < 1.39 to 2.1). In conclusion, 25OHD levels are unlikely to have large effects on risk of youth-onset T2D across different ethnicities.

Keywords: vitamin D; pediatric type 2 diabetes; Mendelian randomization; GWAS; causal inference

1. Introduction

Type 2 diabetes (T2D) is considered as a global epidemic in both adults and youth [1]. Previously a rare disease in children, T2D has an increasing prevalence in children and adolescents worldwide and in all ethnicities, even if the prevalence of obesity remains stable [2]. Notably, T2D accounts for up to 45% of cases of diabetes in youth in certain at-risk populations [3]. Furthermore, individuals who develop T2D in childhood and adolescence have a substantial risk for T2D-related complications at a young age [4]. Thus, prevention of young-onset T2D presents a particularly important yet difficult challenge. Primary prevention through lifestyle-based approaches aiming to reduce the burden of obesity is paramount, but the long-term success of such interventions is limited [5]. Therefore, identifying novel biomarkers as modifiable risk factors for primary prevention of T2D in youth is a public health priority.

Due to its potential immune-modulating and anti-inflammatory properties, vitamin D has been extensively studied as a biomarker for the autoimmune type 1 diabetes [6], but data on its role on youth-onset T2D are sparse. Case-control studies have demonstrated that adolescents with insulin resistance display lower levels of 25 hydroxyvitamin D (25OHD), the biomarker of Vitamin D in humans [7], but this does not necessarily apply to pediatric patients with prediabetes or T2D [8,9]. Various studies suggest that vitamin D supplementation can reverse the progress of prediabetes, a condition present in up to 20% of adolescents, to overt type 2 diabetes [10,11].While many large randomized controlled trials have investigated a potential effect of vitamin D in adult T2D [12],

mostly with negative results, the evidence on the benefit of vitamin D supplementation to prevent T2D in adolescents with insulin resistance is contradictory [13,14]. A meta-analysis of vitamin D randomized controlled trials (RCTs) in obese adolescents has shown an improvement of HOMA-IR, a marker of insulin resistance [15]. A recent study from Iran demonstrated the cost-effectiveness of population-wide vitamin D supplementation in youth to prevent T2D in adulthood [16]. Against this backdrop, better evidence is needed to understand if vitamin D insufficiency, which affects up to 40% of children and adolescents among different ethnicities [17-19], predisposes to youth-onset T2D.

One of the approaches which has been widely used to study causality between vitamin D and T2D in adults [12] is Mendelian randomization (MR). Under specific assumptions, MR uses single nucleotide polymorphisms (SNPs) as instruments for a modifiable exposure to study causal effects of this exposure on a disease-outcome [20]. The effects of these SNPs on both exposure and outcome are derived from large genome-wide association studies (GWAS). Due to the random allocation of genetic variants at conception, MR allows for causal inference by limiting bias due to confounding and reverse causation hampering observational studies, since these bias cannot affect the germline genetic architecture of an individual. As such, MR provides evidence of effects of a lifetime exposure to a genetically determined level of an exposure on a disease-outcome.

In this study, we aimed to test whether genetically altered 25OHD levels are causally associated with risk of youth-onset T2D in a mixed-ancestry and in ethnic-specific cohorts using MR. To do this, we leveraged data from the largest available European and ancestry-specific GWAS on 25OHD levels [21,22] and on the only available to date multi-ethnic GWAS on pediatric T2D [23].

2. Materials and Methods

Our methods and findings are reported according to the MR-STROBE checklist (Supplemental Material).

2.1. SNPs Associated with 25OHD Levels

In order to test the causality of 25OHD on risk of pediatric T2D within the MR framework, we first obtained conditionally independent SNPs associated with 25OHD in a European GWAS meta-analysis of the SUNLIGHT consortium with UK BIOBANK totaling 443,734 individuals [21]. In this GWAS, the mean 25OHD level, measured using the Diasorin assay, was 70 nmol/L (SD 34.7 nmol/L). The effect of these SNPs on risk of pediatric T2D were sought in the multi-ethnic PRODIGY GWAS cohort (n=3,006 youth cases [mean age 15.1 years] and 6,061 adult controls) [23]. The 3,006 cases of the PRODIGY combined 449 youth with T2D from the TODAY study with >2,000 adolescents with T2D from a TODAY ancillary genetics study, and 468 adolescents with T2D from SEARCH for Diabetes in Youth. The adult controls were retrieved from the T2D-GENES study.

We also undertook ethnic-specific analyses by extracting effects of the 25OHD SNPs in the three ethnic sub-cohorts of PRODIGY (Non-Hispanic Whites, n=664 cases/1,434 controls;, African-Americans, n= 1,068 cases/1,068 controls; and Hispanics, n= 1,274 cases/3,559 controls). Descriptives of the GWAS populations can be found in the respective GWAS publications [21,23]. For 25OHD-related SNPs not directly found in the ethnic-specific PRODIGY GWAS, we sought proxy SNPs (LD $r^2 > 0.7$) using ldlink [24] and its LDproxy function in matching populations from the 1000 genomes phase 3 panel. For our main MR analysis, we computed the Wald ratios of the SNP-IVs and meta-analyzed these individual Wald rations using the inverse-variance weighted (IVW) approach [25]. The results of our MR analyses are expressed as the odds ratio (OR) of pediatric T2D associated to one standard deviation (SD) increase in the log-transformed level of 25OHD (which represents a 40.9 nmol/L change in serum 25OHD level in a vitamin-D sufficient individual). MR P-values below 0.05 were considered significant. Figure 1 illustrates the flowchart of our analyses and Figure 2 depicts the MR direct acyclic graph of our study.
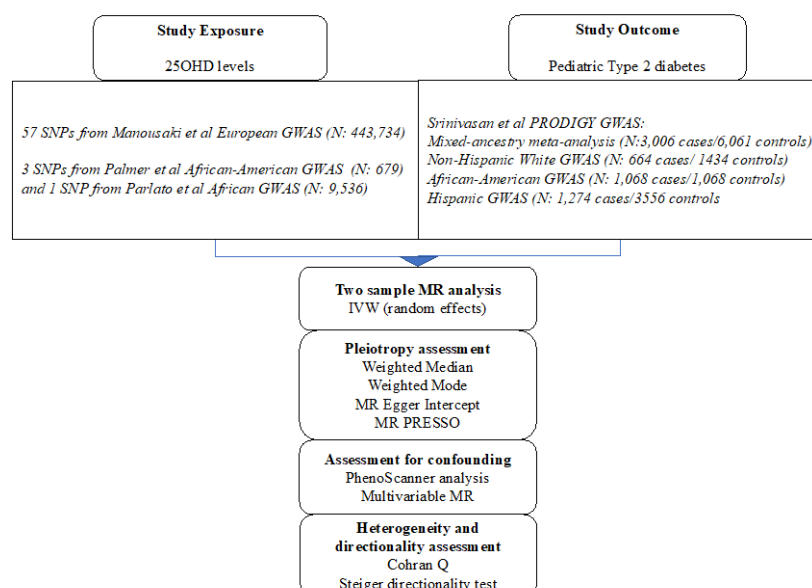
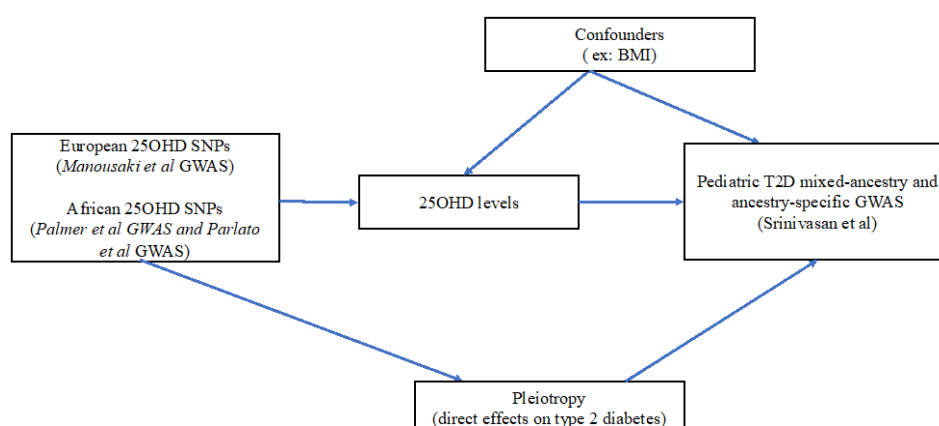**Figure 1.** Flowchart with the design of our MR study.

Figure 2. MR Direct Acyclic Graph (DAG) of our study.

2.2. MR Assumptions

In order to conduct any MR study, the variants used as instrumental variables (IVs) of an exposure must satisfy three assumptions. The first assumption (relevance assumption) requires that these SNPs should be strongly associated with the exposure, in this case, the 25OHD levels. This is ensured by using SNPs linked to 25OHD at a genome-wide significant level (p-value <5 x 10⁻⁸). We also calculated the F-statistic for the 25OHD SNP-IVs, as an additional measure of the strength of our MR instruments.

The second assumption (independence assumption) requires that the SNPs used as IVs should not be linked to confounders of the association between the exposure and outcome. For instance, in this study, a possible confounder can be the body mass index (BMI), since obesity is associated with lower 25OHD levels [26] and truncal adiposity is and established risk factor for both adult and pediatric T2D [1]. Another confounder of the association between 25OHD and T2D risk is ethnicity. For instance, African-Americans present lower 25OHD levels and increased risk of T2D [27]. We undertook multiple sensitivity analyses with different sets of 25OHD SNPs to ensure that our MR results are not biased by the above confounders. Additionally, we performed a multivariable MR analysis considering both 25OHD and pediatric BMI as exposures. The third MR assumption

(exclusion restriction assumption) requires that the SNP-IVs affect the outcome (here pediatric T2D) solely via the exposure (here 25OHD levels). Pleiotropy refers to a situation where this assumption is violated. In order to test for presence of pleiotropy, we conducted sensitivity analyses applying various pleiotropy-robust MR methods, each one with its own assumptions.

2.3. Sensitivity Analyses Addressing Bias Due to Confounding

Using the PhenoScanner database [28], we filtered the MR instruments for 25OHD for SNPs with previously reported GWAS association with confounders of the 25OHD-pediatric T2D association. A detailed description of this approach can be found in previous MR studies by our group [29,30]. Specifically, we conducted sensitivity MR analyses excluding 25OHD SNPs presenting genome-wide associations with BMI, body composition traits, and adult-onset T2D, since family history of adult-onset T2D is a strong risk factor for youth-onset T2D.

To further account for effects of BMI on our MR estimates, we conducted a multivariable MR (MVMR) analysis [31]. To do this, we queried the effects of 25OHD SNPs on childhood BMI from a large European GWAS meta-analysis by the EGG consortium (5,530 cases / 8,318 controls) [32].

Finally, to further account for confounding due to ancestry, we conducted a sensitivity analysis selecting three directly matching rare SNPs (rs14355701 in *TINK*, rs116950775 in *KIAA1644/LDOC1L* and rs111955953 in *FTMT*) from an African-American 25OHD GWAS on 697 individuals [22], which were combined to a common SNP in *GC* (rs4588) identified in a recent larger GWAS meta-analysis on 2,602 African Americans from the Southern Community Cohort Study and 6,934 African- or Caribbean-ancestry participants from the UK Biobank [33]. Effects (betas) of these SNPs on 25OHD were extracted from the above African 25OHD GWAS and their effects on pediatric T2D were identified in the African-American subset of PRODIGY.

2.4. Sensitivity Analyses Addressing Pleiotropy

We employed four pleiotropy-robust MR methods (weighted-median, weighted mode, MR-Egger, and MR-PRESSO) to investigate if pleiotropy could have biased the MR estimates of our main IVW MR analysis. Each one of the above methods has its own assumptions.

Specifically, the weighted-median method can generate valid estimates if more than 50% of the SNPs used as instruments are valid. This methods is based on the fact that estimates of SNPs without pleiotropic effects are usually found closer to the median, while those of pleiotropic SNPs would be found further from the median and introduce heterogeneity [34]. The weighted mode method is very similar to the weighted-median method, with the only difference is that it identifies SNPs without pleiotropic effects as the ones with estimates near the mode [35]. The MR-Egger method [36] detects potentially unbalanced directional pleiotropy. This method generates a regression slope, which signifies the causal estimate, and an intercept, which signifies the presence of directional pleiotropy if it differs from the null. To do this, MR-Egger allows for balanced pleiotropy, but requires that the association of each IV the exposure is not proportional to its pleiotropic effect (InSIDE hypothesis). Moreover, the Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO) method detects, excludes outlier SNP with potential pleiotropic effects, and test for significant distortion of the main MR estimate after outlier removal (global, outlier and distortion tests) [37]. As an additional approach to detect pleiotropy, we computed heterogeneity estimates for each 25OHD SNP in our IVW and MR-Egger analyses, using the Cohran-Q metric. Finally, we applied the Steiger directionality test [38] to investigate if the direction of the association (ie the assumption than exposure causes outcome) was valid in our MR analyses.

As a further strategy to address bias due to pleiotropy, we undertook MR analyses using four SNPs for 25OHD mapping in genes with a known role in vitamin D metabolism [39] (rs12785878 in *NADSYN1-DHCR7*, rs10741657 in *CYP2R1*, rs3755967 in *GC* and rs17216707 in *CYP24A1*) since SNPs in these genes are unlikely to be pleiotropic. Effects (betas) on 25OHD of these four SNPs were derived from the Jiang et al vitamin D GWAS [40].

The TwoSampleMR R package (version 0.5.6) [41], and its default parameters (LD-clumping r2=0.001) was used to retrieve 25OHD SNPs, harmonize them between the exposure and outcome

GWAS, and compute the various MR estimates [IVW, weighted median, MR-Egger and weighted mode] in our analyses. Random effects IVW was used, given the presence of heterogeneity in our main analyses. Scatter plots and forest plots of our MR studies were prepared using the TwoSampleMR R package. We used the MVMR R package for our multivariable MR analysis [42]. Our MR-PRESSO analyses were applied using the MR-PRESSO R package (version 1.0) [37].

2.5. Statistical Power Analysis

We computed the power in our main MR analyses, using the full set of 25OHD SNPs in the mixed-ancestry PRODIGY or the ethnic-specific PRODIGY sub-cohorts using an established power calculation method for MR [43]. Specifically, we computed the MR odds ratio (OR) for pediatric T2D for which our analyses obtained a power of 80%, setting the alpha level at 0.05, using the variance explained of 25OHD by its respective SNP-IVs, and the sample sizes of the entire PRODIGY and its ethnic sub-cohorts.

3. Results

3.1. Main MR Studies on the Effect of Serum 25OHD on Risk of Pediatric T2D across Different Ancestries

Among 138 independent genome-wide significant SNPs reported in the 25OHD GWAS, we first excluded 4 ambiguous ones (ie SNPs with nonconcordant alleles e.g., A/G vs. A/C). We retained 87 SNPs with minor allele frequency (MAF) >1%, since rare variants were unlikely to be found in the mixed ancestry meta-analysis and the ethnic-specific GWAS of PRODIGY. Following LD clumping ($r2 < 0.001$), the SNP number was reduced to 57. In the entire PRODIGY, and its Non-Hispanic White, African-American, and Hispanic sub-cohorts, we found 49, 49, 46 and 47 directly matching among the 57 SNPs (Tables S1-S4). We were able to identify five proxy SNPs with LD $r2 > 0.8$ in the Non-Hispanic White subset of the PRODIGY GWAS but not in the other MR analyses.

In our main MR studies, the 49, 54, 46 and 47 common 25OHD SNPs explained 3.1% of the variance in 25OHD levels in Non-Hispanic Whites and 2.5% of the variance in 25OHD in the mixed-ancestry meta-analysis, in African-Americans and in Hispanics respectively. Using these SNPs as instruments, we did not find any significant MR association between 25OHD levels and pediatric T2D risk in the mixed-ancestry PRODIGY cohort and its ethnic-specific subsets (Table 1, Tables S1-S4, Figures S1-S4). As shown in Table 1, our MR estimates remained largely consistent and close to the null across different MR methods, with the exception of the weighted median MR demonstrating a marginal effect of a SD increase in log-transformed 25OHD on pediatric T2D risk in the mixed-ancestry cohort (OR pediatric T2D 1.09, 95% CI 1.00-1.18, P = 0.049) and in Hispanics (OR 1.13, 95% CI 1.02-1.26, P=0.019). Nevertheless, these results were not supported by the MR-Egger, IVW and the weighted mode methods in the mixed-ancestry cohort and by the IVW and weighted mode methods in the Hispanic sub-cohort. Given the large consistency of our results across different methods and ancestries, we conclude that the above positive findings, while suggestive, cannot unequivocally prove a role of 25OHD in pediatric T2D.

Table 1. Results (in the form of MR OR for pediatric T2D per SD increase in log-transformed 25OHD) of the main MR analysis in the mixed-ancestry and ancestry-specific PRODIGY cohorts.

Method	N of SNPs	OR	95%_LL	95%_UL	P-val	Cochran-Q	Cochran-Q p-val	MR-Egger interecept	MR-Egger intercept p-val	MR-PRESSO global test RSS obs	MR-PRESSO global test p-val	MR-PRESSO distortion coeff.	MR-PRESSO distortion test p-val
Meta-analysis													
IVW	49	1.041	0.958	1.132	0.346	89.547	**1.27E-04**						
Weighted median	49	1.086	1.000	1.179	**0.049**								
Weighted mode	49	1.066	0.992	1.145	0.089								
MR-Egger	49	1.087	0.973	1.214	0.147	86.956	**1.75E-04**	-0.003	0.196				
MR-PRESSO	49	1.041	0.958	1.132	0.351					95.135	**0.001**	29.149	0.737
Non-Hispanic Whites													
IVW	54	0.948	0.840	1.070	0.390	66.782	0.10						
Weighted median	54	0.935	0.813	1.074	0.341								
Weighted mode	54	0.937	0.824	1.065	0.324								
MR-Egger	54	0.947	0.810	1.108	0.503	66.781	**0.08**	6.20E-05	0.984				
MR-PRESSO	54	0.948	0.840	1.070	0.394					68.270	0.129	NA	NA
African Americans													
IVW	46	1.103	0.899	1.353	0.347	61.003	0.056						
Weighted median	46	1.194	0.929	1.535	0.165								
Weighted mode	46	1.194	0.967	1.475	0.107								
MR-Egger	46	1.252	0.934	1.679	0.140	59.147	0.063	-0.005	0.246				
MR-PRESSO	46	1.103	0.899	1.353	0.394					63.122	0.076	NA	NA

Hispanics													
IVW	47	1.069	0.972	1.177	0.170	75.166	**0.004**						
Weighted median	47	1.133	1.021	1.258	**0.019**								
Weighted mode	47	1.089	1.000	1.185	0.055								
MR-Egger	47	1.142	1.007	1.295	**0.044**	71.316	**0.007**	-0.004	0.126				
MR-PRESSO	47	1.069	0.972	1.177	0.176					83.103	0.015	NA	NA

* asterisks denote significant p-values.

As shown in Table 1, the intercept of the MR-Egger showed no evidence of unbalanced horizontal pleiotropy in any of the MR studies, but there was significant heterogeneity among the MR instruments in all MR analyses except the one in African-Americans. All 25OHD SNP-IVs in our MR analyses had an F-statistic >10 (their average F-statistic was 234) (Tables S1-S4). The Steiger directionality test indicated that the correct causal direction was "TRUE" in all our main MR analyses, confirming that the assumption that altered 25OHD levels cause pediatric T2D (and not the inverse) was valid.

3.2. Sensitivity MR Analyses

As demonstrated in Tables S5 and S6, the results of our sensitivity analyses using four SNPs in core vitamin D genes (explaining 1.03% of the variance in 25OHD), as well as the four African-American 25OHD SNPs, did not indicate any causal effect of 25OHD on pediatric T2D either. Our sensitivity analysis excluding SNPs associated with BMI, body composition traits or adult T2D in the PhenoScanner database provided similar results as those of the main analysis (Table S7) with the exception of a significant estimate of the weighted median analyses in the mixed-ancestry and the Hispanic PRODIGY cohorts. Finally, the results of the MVMR testing for mediating effects of pediatric BMI were equally non-significant (Table S8).

3.3. MR Power Calculation

Our main MR study had 80% power to identify effects as small as an OR of 1.39 for pediatric T2D per 1 SD increase in log-transformed 25OHD in the mixed-ancestry PRODIGY cohort. The respective power estimates for the ancestry-specific MR analyses were an OR of 1.66 for Hispanics, 2.00 for Non-Hispanic Whites and 2.10 for African-Americans (Table S9).

4. Discussion

Using an unbiased MR approach, and leveraging data from the largest available GWAS datasets for 25OHD and pediatric T2D across different ancestries, we did not observe evidence for a large causal effect of 25OHD levels on risk of pediatric T2D. Also, a plethora of sensitivity analyses addressing pleiotropy and accounting for possible confounding effects of BMI provided consistent null results with those of the main IVW analyses. While our study is not powered to exclude small to moderate effects in certain ethnic groups, we can reasonably conclude that it is unlikely that variation of 25OHD levels within the normal distribution affects substantially an individual's predisposition to develop T2D early in life.

Our results suggest that the low 25OHD levels found in adolescents with insulin resistance in an observational study [7], are unlikely to be causal for their altered glucose metabolism, but they are rather driven by confounders, such as obesity. On the other hand, adolescents with obesity tend to

be less physically active are spend less time outdoors, and as such can be less exposed to sunlight and as have decreased vitamin D synthesis in the skin [9]. Moreover, they often have poor dietary habits and suboptimal vitamin D intake [26]. Another possible explanation for the observed associations between low 25OHD levels and risk of pediatric T2D is ancestry. In this respect, a study [8] showed that, although vitamin D deficiency or insufficiency appeared to affect a substantial proportion of youth with T2D, notably non-Europeans, their prevalence was comparable to that in youth without diabetes. Taken together, all the above mechanisms highlight the presence of various possible confounders in the association between vitamin D and pediatric T2D.

Our MR study has a few considerable limitations. First, we did not study causal effects of changes in the active form of vitamin D [1,25 di-hydroxyvitamin D or 1,25(OH)2D] due to lack of availability of GWAS data on this form of vitamin D allowing to extract IVs for 1,25(OH)2D. Another limitation of our study is that the 25OHD SNPs which were used as instruments explain only up to 3.1% of variance in 25OHD levels in the Non-Hispanic White analysis, and only 2.5% of its variance in the other MR analyses. This, combined to the limited sample size of the PRODIGY GWAS, and in particular of its ancestry-specific sub-GWAS, restricted the power of our study to identify small and moderate effects of 25OHD on pediatric T2D risk. We elected to use as instruments in our main mixed-ancestry and ancestry-specific MR analyses 25OHD SNPs identified in a large European vitamin D GWAS. This could have introduced bias due to the fact that effects of these SNPs on 25OHD levels in non-Europeans can differ substantially from those identified in a European GWAS. Nevertheless, there is no available 25OHD GWAS in Hispanics, and the results of our sensitivity analysis using African-American specific 25OHD SNPs were similar to those of the analysis using European 25OHD SNPs. Future large ancestry-informed GWAS on 25OHD and pediatric T2D will enable better power to interrogate their associations in various ethnic populations, including populations not studied in this work, using MR. Finally, the two-sample MR design of our study did not allow to undertake a stratified MR analysis to assess non-linear effects of 25OHD levels, since there are no available 25OHD data in PRODIGY participants. As such, effects of extremely low or high 25OHD levels on risk of pediatric T2D cannot be excluded.

5. Conclusions

In conclusion, our MR study is the first to interrogate the causal association between 25OHD levels and risk of T2D in youth across different ancestries. The absence of evidence for a large causal effect of 25OHD on risk of pediatric T2D can inform decisions on conducting clinical trials or public health interventions. Our results do not support vitamin D supplementation as a measure to prevent T2D in youth of any ethnic background, but we cannot exclude small to moderate causal effects. Targeting lifestyle habits, which affect both 25OHD levels and established risk factors for pediatric T2D, such as obesity, is likely a more promising strategy than vitamin D supplements to prevent T2D in youth without frank vitamin D deficiency.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org: MR-STROBE checklist; Table S1: SNPs used as instruments in the Non-Hispanic White MR analysis with their respective effects on 25OHD and pediatric T2D; Table S2: SNPs used as instruments in the African-American MR analysis with their respective effects on 25OHD and pediatric T2D; Table S3: SNPs used as instruments in the Hispanic MR analysis with their respective effects on 25OHD and pediatric T2D; Table S4: SNPs used as instruments in the mixed-ancestry MR analysis with their respective effects on 25OHD and pediatric T2D; Table S5: Results of the MR sensitivity analysis with 4 vitamin D core gene SNPs; Table S6: Results of the MR sensitivity analysis using 4 African 25OHD SNPs; Table S7: Results of the main MR analysis presented in Table 1 after filtering for SNPs associated with confounders using PhenoScanner; Table S8: Results of the MVMR testing effects of both 25OHD and pediatric BMI on pediatric T2D risk; Table S9: Results of the MR power analysis in the main MR analyses; Figure S1: Scatter plot and forest plot of the main MR analysis in the multi-ethnic PRODIGY cohort; Figure S2: Scatter plot and forest plot of the main MR analysis in Non-Hispanic White PRODIGY sub-cohort; Figure S3: Scatter plot and forest plot of the main MR analysis in the African-American PRODIGY sub-cohort; Figure S4: Scatter plot and forest plot of the main MR analysis in the Hispanic PRODIGY sub-cohort.

References

1. Reinehr, T. Type 2 diabetes mellitus in children and adolescents. *World J Diabetes* **2013**, *4*, 270–281. https://doi.org/10.4239/wjd.v4.i6.270.
2. May, A.L.; Kuklina, E.V.; Yoon, P.W. Prevalence of cardiovascular disease risk factors among US adolescents, 1999-2008. *Pediatrics* **2012**, *129*, 1035–1041. https://doi.org/10.1542/peds.2011-1082.
3. D'Adamo, E.; Caprio, S. Type 2 diabetes in youth: epidemiology and pathophysiology. *Diabetes Care* **2011**, *34 Suppl 2*, S161-165. https://doi.org/10.2337/dc11-s212.
4. Krakoff, J.; Hanson, R.L.; Kobes, S.; Knowler, W.C. Comparison of the effect of plasma glucose concentrations on microvascular disease between Pima Indian youths and adults. *Diabetes Care* **2001**, *24*, 1023–1028. https://doi.org/10.2337/diacare.24.6.1023.
5. CIHI], P.H.A.o.C.t.C.I.o.H.I.P. Obesity in Canada. **2011**.
6. Cadario, F. Vitamin D and omega-3 Polyunsaturated Fatty Acids towards a Personalized Nutrition of Youth Diabetes: A Narrative Lecture. *Nutrients* **2022**, *14*. https://doi.org/10.3390/nu14224887.
7. Peterson, C.A.; Tosh, A.K.; Belenchia, A.M. Vitamin D insufficiency and insulin resistance in obese adolescents. *Ther Adv Endocrinol Metab* **2014**, *5*, 166–189. https://doi.org/10.1177/2042018814547205.
8. Wood, J.R.; Connor, C.G.; Cheng, P.; Ruedy, K.J.; Tamborlane, W.V.; Klingensmith, G.; Schatz, D.; Gregg, B.; Cengiz, E.; Willi, S., et al. Vitamin D status in youth with type 1 and type 2 diabetes enrolled in the Pediatric Diabetes Consortium (PDC) is not worse than in youth without diabetes. *Pediatr Diabetes* **2016**, *17*, 584–591. https://doi.org/10.1111/pedi.12340.
9. de las Heras, J.; Rajakumar, K.; Lee, S.; Bacha, F.; Holick, M.F.; Arslanian, S.A. 25-Hydroxyvitamin D in obese youth across the spectrum of glucose tolerance from normal to prediabetes to type 2 diabetes. *Diabetes Care* **2013**, *36*, 2048–2053. https://doi.org/10.2337/dc12-1288.
10. Zhang, Y.; Tan, H.; Tang, J.; Li, J.; Chong, W.; Hai, Y.; Feng, Y.; Lunsford, L.D.; Xu, P.; Jia, D., et al. Effects of Vitamin D Supplementation on Prevention of Type 2 Diabetes in Patients With Prediabetes: A Systematic Review and Meta-analysis. *Diabetes Care* **2020**, *43*, 1650–1658. https://doi.org/10.2337/dc19-1708.
11. Dawson-Hughes, B.; Staten, M.A.; Knowler, W.C.; Nelson, J.; Vickery, E.M.; LeBlanc, E.S.; Neff, L.M.; Park, J.; Pittas, A.G.; Group, D.d.R. Intratrial Exposure to Vitamin D and New-Onset Diabetes Among Adults With Prediabetes: A Secondary Analysis From the Vitamin D and Type 2 Diabetes (D2d) Study. *Diabetes Care* **2020**, *43*, 2916–2922. https://doi.org/10.2337/dc20-1765.
12. Bouillon, R.; Manousaki, D.; Rosen, C.; Trajanoska, K.; Rivadeneira, F.; Richards, J.B. The health effects of vitamin D supplementation: evidence from human studies. *Nat Rev Endocrinol* **2022**, *18*, 96–110. https://doi.org/10.1038/s41574-021-00593-z.
13. Javed, A.; Vella, A.; Balagopal, P.B.; Fischer, P.R.; Weaver, A.L.; Piccinini, F.; Dalla Man, C.; Cobelli, C.; Giesler, P.D.; Laugen, J.M., et al. Cholecalciferol supplementation does not influence beta-cell function and insulin action in obese adolescents: a prospective double-blind randomized trial. *J Nutr* **2015**, *145*, 284–290. https://doi.org/10.3945/jn.114.202010.
14. Belenchia, A.M.; Tosh, A.K.; Hillman, L.S.; Peterson, C.A. Correcting vitamin D insufficiency improves insulin sensitivity in obese adolescents: a randomized controlled trial. *Am J Clin Nutr* **2013**, *97*, 774–781. https://doi.org/10.3945/ajcn.112.050013.
15. Gou, H.; Wang, Y.; Liu, Y.; Peng, C.; He, W.; Sun, X. Efficacy of vitamin D supplementation on child and adolescent overweight/obesity: a systematic review and meta-analysis of randomized controlled trials. *Eur J Pediatr* **2022**, 10.1007/s00431-022-04673-8. https://doi.org/10.1007/s00431-022-04673-8.
16. Zandieh, N.; Hemami, M.R.; Darvishi, A.; Hasheminejad, S.M.; Abdollahi, Z.; Zarei, M.; Heshmat, R. Economic evaluation of a national vitamin D supplementation program among Iranian adolescents for the prevention of adulthood type 2 diabetes mellitus. *BMC Complement Med Ther* **2022**, *22*, 1. https://doi.org/10.1186/s12906-021-03474-0.

17. Gordon, C.M.; DePeter, K.C.; Feldman, H.A.; Grace, E.; Emans, S.J. Prevalence of Vitamin D Deficiency Among Healthy Adolescents. *Archives of Pediatrics & Adolescent Medicine* **2004**, *158*, 531–537. https://doi.org/10.1001/archpedi.158.6.531.

18. Guo, Y.; Ke, H.-J.; Liu, Y.; Fu, M.; Ning, J.; Yu, L.; Xiao, Y.; Che, D.; Chen, X.-Y.; Deng, Y.-H., et al. Prevalence of vitamin D insufficiency among children in southern china: A cross-sectional survey. *Medicine* **2018**, *97*, e11030. https://doi.org/10.1097/md.0000000000011030.

19. Khadilkar, A.; Kajale, N.; Oza, C.; Oke, R.; Gondhalekar, K.; Patwardhan, V.; Khadilkar, V.; Mughal, Z.; Padidela, R. Vitamin D status and determinants in Indian children and adolescents: a multicentre study. *Sci Rep* **2022**, *12*, 16790. https://doi.org/10.1038/s41598-022-21279-0.

20. Lawlor, D.A.; Harbord, R.M.; Sterne, J.A.; Timpson, N.; Davey Smith, G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* **2008**, *27*, 1133–1163. https://doi.org/10.1002/sim.3034.

21. Manousaki, D.; Mitchell, R.; Dudding, T.; Haworth, S.; Harroud, A.; Forgetta, V.; Shah, R.L.; Luan, J.; Langenberg, C.; Timpson, N.J., et al. Genome-wide Association Study for Vitamin D Levels Reveals 69 Independent Loci. *Am J Hum Genet* **2020**, *106*, 327–337. https://doi.org/10.1016/j.ajhg.2020.01.017.

22. Palmer, N.D.; Lu, L.; Register, T.C.; Lenchik, L.; Carr, J.J.; Hicks, P.J.; Smith, S.C.; Xu, J.; Dimitrov, L.; Keaton, J., et al. Genome-wide association study of vitamin D concentrations and bone mineral density in the African American-Diabetes Heart Study. *PLoS One* **2021**, *16*, e0251423. https://doi.org/10.1371/journal.pone.0251423.

23. Srinivasan, S.; Chen, L.; Todd, J.; Divers, J.; Gidding, S.; Chernausek, S.; Gubitosi-Klug, R.A.; Kelsey, M.M.; Shah, R.; Black, M.H., et al. The First Genome-Wide Association Study for Type 2 Diabetes in Youth: The Progress in Diabetes Genetics in Youth (ProDiGY) Consortium. *Diabetes* **2021**, 10.2337/db20-0443. https://doi.org/10.2337/db20-0443.

24. Availabe online: https://ldlink.nci.nih.gov (accessed on

25. Burgess, S.; Butterworth, A.; Thompson, S.G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* **2013**, *37*, 658–665. https://doi.org/10.1002/gepi.21758.

26. Zakharova, I.; Klimov, L.; Kuryaninova, V.; Nikitina, I.; Malyavskaya, S.; Dolbnya, S.; Kasyanova, A.; Atanesyan, R.; Stoyan, M.; Todieva, A., et al. Vitamin D Insufficiency in Overweight and Obese Children and Adolescents. *Front Endocrinol (Lausanne)* **2019**, *10*, 103. https://doi.org/10.3389/fendo.2019.00103.

27. Schwartz, J.B.; Lai, J.; Lizaola, B.; Kane, L.; Markova, S.; Weyland, P.; Terrault, N.A.; Stotland, N.; Bikle, D. A comparison of measured and calculated free 25(OH) vitamin D levels in clinical populations. *J Clin Endocrinol Metab* **2014**, *99*, 1631–1637. https://doi.org/10.1210/jc.2013-3874.

28. Kamat, M.A.; Blackshaw, J.A.; Young, R.; Surendran, P.; Burgess, S.; Danesh, J.; Butterworth, A.S.; Staley, J.R. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* **2019**, *35*, 4851–4853. https://doi.org/10.1093/bioinformatics/btz469.

29. Adi, M.; Ghanbari, F.; Downie, M.L.; Hung, A.; Robinson-Cohen, C.; Manousaki, D. Effects of 25 hydroxyvitamin D levels on renal function: a bidirectional Mendelian randomization study. *J Clin Endocrinol Metab* **2022**, 10.1210/clinem/dgac724. https://doi.org/10.1210/clinem/dgac724.

30. Manousaki, D.; Harroud, A.; Mitchell, R.E.; Ross, S.; Forgetta, V.; Timpson, N.J.; Smith, G.D.; Polychronakos, C.; Richards, J.B. Vitamin D levels and risk of type 1 diabetes: A Mendelian randomization study. *PLoS Med* **2021**, *18*, e1003536. https://doi.org/10.1371/journal.pmed.1003536.

31. Burgess, S.; Thompson, S.G. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* **2015**, *181*, 251–260. https://doi.org/10.1093/aje/kwu283.

32. Bradfield, J.P.; Taal, H.R.; Timpson, N.J.; Scherag, A.; Lecoeur, C.; Warrington, N.M.; Hypponen, E.; Holst, C.; Valcarcel, B.; Thiering, E., et al. A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat Genet* **2012**, *44*, 526–531. https://doi.org/10.1038/ng.2247.

33. Lisa A. Parlato, R.W., Irene M. Ong, Jirong Long, Qiuyin Cai, Mark D. Steinwandel, William J. Blot, Wei Zheng, Shaneda Warren Andersen. Genome-wide association study (GWAS) of circulating vitamin D outcomes among individuals of African ancestry. *Am. J. Clin. Nutr.* **2022**. (https://doi.org/10.1016/j.ajcnut.2022.12.001.

34. Bowden, J.; Davey Smith, G.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* **2016**, *40*, 304–314. https://doi.org/10.1002/gepi.21965.

35. Hartwig, F.P.; Davey Smith, G.; Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* **2017**, *46*, 1985–1998. https://doi.org/10.1093/ije/dyx102.

36. Bowden, J.; Davey Smith, G.; Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* **2015**, *44*, 512–525. https://doi.org/10.1093/ije/dyv080.

37. 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**, *50*, 693–698. https://doi.org/10.1038/s41588-018-0099-7.

38. Hemani, G.; Tilling, K.; Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet* **2017**, *13*, e1007081. https://doi.org/10.1371/journal.pgen.1007081.

39. Mokry, L.E.; Ross, S.; Ahmad, O.S.; Forgetta, V.; Smith, G.D.; Goltzman, D.; Leong, A.; Greenwood, C.M.; Thanassoulis, G.; Richards, J.B. Vitamin D and Risk of Multiple Sclerosis: A Mendelian Randomization Study. *PLoS Med* **2015**, *12*, e1001866. https://doi.org/10.1371/journal.pmed.1001866.

40. Jiang, X.; O'Reilly, P.F.; Aschard, H.; Hsu, Y.H.; Richards, J.B.; Dupuis, J.; Ingelsson, E.; Karasik, D.; Pilz, S.; Berry, D., et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat Commun* **2018**, *9*, 260. https://doi.org/10.1038/s41467-017-02662-2.

41. Hemani, G.; Zheng, J.; Elsworth, B.; Wade, K.H.; Haberland, V.; Baird, D.; Laurin, C.; Burgess, S.; Bowden, J.; Langdon, R., et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **2018**, *7*. https://doi.org/10.7554/eLife.34408.

42. Sanderson, E.; Davey Smith, G.; Windmeijer, F.; Bowden, J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol* **2019**, *48*, 713–727. https://doi.org/10.1093/ije/dyy262.

43. Brion, M.J.; Shakhbazov, K.; Visscher, P.M. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* **2013**, *42*, 1497–1501. https://doi.org/10.1093/ije/dyt179.