Vitamin D and the preclinical Alzheimer cognitive composite cognition (PACC) score: a two-sample mendelian randomization study

Yann Ilboudo¹,², Satoshi Yoshiji¹,²,³,⁴,⁵, Tianyuan Lu¹,⁶,¹⁰, Guillaume Butler-Laporte⁸, Sirui Zhou¹,³, J. Brent Richards¹,²,⁸,⁹,¹⁰,¹¹*

Affiliations:
1 Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada
2 Department of Medicine, McGill University, Montreal, Quebec, Canada
3 Department of Human Genetics, McGill University, Montréal, Québec, Canada
4 Kyoto-McGill International Collaborative Program in Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan
5 Japan Society for the Promotion of Science, Japan
6 Quantitative Life Sciences Program, McGill University, Montréal, Québec, Canada
7 Department of Human Genetics, McGill University, Montreal, Quebec, Canada
8 Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec, Canada
9 Department of Twin Research, King’s College London, London, U.K.
10 Prime Sciences, Montréal, Québec, Canada
11 Department of Twin Research, King’s College London, London, United Kingdom

Correspondence:
J. Brent Richards, brent.richards@mcgill.ca
Professor of Medicine, McGill University, Senior Lecturer, King’s College London (Honorary), Pavilion H-413, Jewish General Hospital, 3755 Côte-Ste-Catherine Montréal, Québec, H3T 1E2, Canada.
Tel: +1-514-340-8222    Fax: +1-514-340-7529
ABSTRACT

**Objective**: Observational studies on cognition found that vitamin D supplementation is associated with improved cognition. Further, recent Mendelian randomization (MR) studies have shown that increased vitamin D levels, as measured by 25-hydroxyvitamin D [25(OH)D], may protect against Alzheimer’s disease. Thus, it is possible that 25OHD may protect against Alzheimer’s disease by improving cognition.

**Methods**: To evaluate this hypothesis, we began by performing an observational study by testing the association of 25OHD levels with 5 cognitive outcomes related to memory and executive function, adjusting for age, sex, BMI, smoking status and nutritional risk index in 26,787 older individuals (median age 62). Since such observational analyses can be biased by confounding, we next performed two-sample mendelian randomization (MR) analyses testing the effect of 25(OH)D on cognitive outcomes, since this method helps to protect against bias from confounding.

**Results**: Observational studies suggested a strong protective association between 25(OH)D and cognitive measures. However, in MR analyses, we found no estimated effect of 25(OH)D on these outcomes.

**Conclusion**: These findings suggest that associations between 25(OH)D levels and cognition are likely to be biased by confounding and that the relationship between 25(OH)D levels and Alzheimer’s disease is not due to the effect of 25(OH)D on cognition.
INTRODUCTION

According to the Canadian census bureau, the number of 85 years old and older will triple over the next 25 years to 2.5 million\(^1\). The rate of dementia-related diseases increases exponentially for individuals 90 years and older\(^2\). Therefore, identifying therapeutic targets and enhancing our understanding of age-related neurodegeneration is necessary.

Some studies have demonstrated an increased prevalence of low 25-hydroxyvitamin D [25(OH)D] levels in the elderly population\(^3,4\). Further, cognitive decline and loss of autonomy in the elderly, which are key drivers of Alzheimer’s disease (AD) have been observed in individuals with hypovitaminosis D\(^5\). There is a wide range of pathways through which vitamin D may influence brain health, including inflammation, thrombosis, and growth factors\(^6,7\). Several observational studies have suggested a role of vitamin D in modulating neuronal development, increasing senescence, reducing sensory and social acuity\(^8,9,10\). Meta-analyses of the association of vitamin D and stroke incidence and dementia risk found a dose-response relationship between vitamin D and these complications\(^11,12\). Since vitamin D levels can be increased with relative safety using available supplements, increasing vitamin D levels could potentially help to protect against cognitive decline.

The preclinical Alzheimer’s cognitive composite cognition score\(^13\) (PACC) tests multiple brain function parameters. The meta-analysis from which the score was derived demonstrated a strong association of the score with the gene encoding for the apolipoprotein E \(\varepsilon 4\) polymorphism\(^14\). These genetic variants are the most significant risk factors for Alzheimer’s disease, which suggests that the PACC score is a helpful score to measure Alzheimer’s disease-related cognitive impairment, since it shares biological determinants with Alzheimer’s disease.

Understanding the causal effect of vitamin D on cognitive function would therefore be important to better understand widespread findings from observational study data demonstrating an association of 25(OH)D with cognitive function. However, such studies
may be biased by confounding and previous work has shown that several observed associations between vitamin D levels and disease outcomes cannot be corrected by vitamin D supplementation in large-scale randomized trials\textsuperscript{15,16,17,18,19,20,21,22}. Mendelian randomization (MR) is a genetic epidemiology causal inference method that uses instrumental variables to reduce bias from confounding and reverse causation\textsuperscript{23,24}. Contrasting the existing observational studies mentioned earlier, a one-sample MR study in a meta-analysis on the impact of vitamin D on cognition found no evidence for 25(OH)D levels as a causal factor for cognitive performance\textsuperscript{25}. Additionally, different ways to measure cognition (including memory and global cognition scores) amongst the different cohorts made the cognition measurements heterogeneous. Finally, the study employed just two SNPs (\textit{CYP2R1, DHCR7}) associated with 25(OH)D, thus reducing the variance explained in 25(OH)D levels and, consequently statistical power. On the other hand, recent MR studies have demonstrated that increased 25(OH)D levels protect against Alzheimer’s disease\textsuperscript{26}. Since cognition is an important risk factor for Alzheimer’s disease, it would be important to better understand if increased 25(OH)D levels were protective for cognition and if these results depended upon age.

Therefore, to better understand the relationship between 25(OH)D and cognition (determined using the PACC score), we undertook a large-scale observational study, testing the association of 25(OH)D with cognitive performance in 26,787 individuals in the Canadian Longitudinal Study of Aging (CLSA). Next, we used MR to better understand if these associations were influenced by confounding using genome-wide significant SNPs from the largest (N=443,734) GWAS of 25OHD levels\textsuperscript{27} as genetic instruments. We then queried their effect on the PACC score in 17,166 samples from the Canadian longitudinal study on aging (CLSA). The results provide insights into the effect of 25(OH)D levels of cognition in humans.
METHODS

Observational associations

Using linear regression, we tested the association of log transformed 25(OH)D with each individual cognition score (AFT, AFT2, MAT, REYI, REYII, COLTIME) and with the composite PACC score. Since 25(OH)D is confounded by various factors (anthropometric variables, lifestyle habits, and more), we identified covariates available in the CLSA cohort to add to our regression model to partially control for this potential bias. Cognition scores were adjusted for age, sex, smoking status (answered on a questionnaire with possible answers: Yes (I currently smoke)), No (I don't smoke and I never have), Former (I don't smoke now but I have in the past), high nutritional risk indicator (answered on a questionnaire with possible answers: Not at high nutritional risk, and High nutritional risk), and BMI classification for adults aged 18 and over (the classification includes underweight, normal weight, overweight, obese – Class I, obese – Class II, obese – Class III). Cognition scores were inverse normal transformed. All measures were at baseline and included at most 25,357 individuals. To control for multiple testing, we considered Bonferroni significant those associations with $P < 7.1 \times 10^{-3}$ ($0.05/7$ (where 7 is the number of cognitive scores tested)).

Instrumental variables for 25(OH)D

Instrumental variables were defined as independent genome-wide significant SNPs ($P < 6.6 \times 10^{-9}$) for 25(OH)D levels in 443,734 individuals. Identification of independent loci excluded high collinearity ($R^2 > 0.9$) SNPs within a 20,000 kilobase window. Details on the conditionally independent approach, imputation quality control, missingness rates excluded individuals and software tools employed in the GWAS are provided in great detail in our original paper. The approach yielded 138 independent SNPs. In order to ensure that we are able to recover all the variants from the outcome GWAS, we restricted our study to SNPs with a MAF > 5%. When a SNP was not present in the results for the GWAS of the PACC, AFT, COLTIME, REYI, REYII, and MAT score, we alternatively looked up a proxy SNP in linkage disequilibrium, with an $R^2 \geq 0.8$ using the 1000G European reference panel (Supplementary Table 2). Therefore, we limited our analysis to 80 SNPs for which we estimated the variance explained in 25(OH)D. We computed
the variance explained using the following formula: \( r^2 = 2\beta^2 f (1 - f) \),

where \( \beta \) and \( f \) denote the effect of the SNP on 25OHD level and the MAF, respectively.

Which we then used to calculate the \( F \)-statistic (formula \( F: r^2 x (N-2)/(1-r^2) \)) where \( r^2 \) is the variance explained and \( N \) denotes the sample size.

### Study outcomes

The PACC score is a composite score which synthesizes tests on episodic memory, timed executive function, and global cognition. The composite measure can precisely recognize early signs of cognitive changes prior to mild cognitive impairment (MCI)\(^\text{13}\). In terms of the definition of the individual scores, The Rey auditory verbal learning test – trial 1 and five-minute delayed recall (respectively) REYI, and REYII score represents the number of words correctly recalled in 90 seconds. The first one assesses verbal episodic memory, whereas the second one assesses delayed recall. The mental alteration test (MAT) score assesses the number of correct consecutive numeric and alphabetical alternations in 30 seconds. The time to say the ink colour of each colour name (COLTIME measured in seconds) ascertains the executive and the set switching cognitive domains. Finally, animal fluency (AFT & AFT2) assesses the different animals recited in 60 seconds. **Table 1** summarizes the individual cognitive scores in CLSA.

To ascertain the association between 25(OH)D levels and the PACC\(^\text{13}\) score, we retrieved the effects of these 80 variants on the PACC, REYI, REYII, MAT, AFT and AFT2 scores, which is a GWAS of the aforementioned score in the CLSA cohort described elsewhere\(^\text{30}\). Since 92.6% of CLSA participants are of Europeans ancestry\(^\text{30}\), our analysis focused on individuals with European ancestry only with PACC score data (\( N = 18,706 \)). Population structure was identified through principal component analysis as detailed elsewhere\(^\text{30}\). We employed FastGWA as implemented in GCTA\(^\text{31}\) and PLINK2\(^\text{32}\) to perform association tests, which we adjusted for sex, age, and the first 20 principal components. We then stratified the population by age testing the entire cohort and individuals at least 60 years of age (\( N = 9,489 \)) to compare the effect of age on the cognitive scores (**Supplementary Table1**).
Mendelian randomization

Units
The causal associations were evaluated using odds ratios (ORs) for the susceptibility to multiple sclerosis, which are expressed according to a standard deviation (SD) increase in 1SD of genetically predicted natural log of 25(OH)D (nmol/l). For the cognition scores, one standard deviation increase in cognitive scores are expressed according to one standard deviation increase in genetically predicted natural log of 25(OH)D (nmol/l).

Two-sample MR
We performed the two-sample mendelian randomization analyses using multiplicative random effect inverse-variance weighted (referred to as IVW(re) ) to estimate the effect of a 1 SD increase in standardized natural log-transformed 25(OH)D on 1 a SD of PACC, AFT, REYI, REYII, MAT, COLTIME scores and odds of multiple sclerosis. We additionally stratified our analysis by age (individuals above 60 years of age versus all individuals) when estimating the effect on cognitive scores. Our analyses we performed with the package TwoSampleMR (v 0.5.6) in R(v4.2.0). We considered significant MR results with $P$-value $< 6.3 \times 10^{-3}$ (Bonferroni threshold accounting for multiple testing for the seven cognition scores and the risk of multiple sclerosis: 0.05/8).

MR Assumptions
Three assumptions are required in MR analyses. The first MR assumption requires that the genetic instrument must be strongly associated with the exposure. We selected conditionally independent genome-wide significant common ($P$-value $< 6.6 \times 10^{-9}$) variants to satisfy this assumption. Of 138 conditionally independent variants, 80 remained (Supplementary Table1).

The second MR assumption stipulates that genetic instruments should not be correlated to confounders linking exposure and outcome.

The third MR assumption requires the instruments to influence the outcome only via their effects on the exposure (the exclusion restriction assumption), i.e. that horizontal...
pleiotropy does not influence the results. We therefore investigated for the presence of directional horizontal pleiotropy in our study. First, we calculated MR Egger regression slope and intercept, in addition to IVW median and IVW mode. MR-Egger is a weighted linear regression which allows the estimation of an intercept as a measure of the average pleiotropic effect and produces a slope coefficient that is robust to directional horizontal pleiotropy. MR-Egger relaxes the exclusion restriction assumption and is valid under the Instrument Strength Independent of Direct Effect (InSIDE) assumption wherein associations of the genetic variants with the 25(OH)D are independent of direct effects of the genetic variants on the cognitive score. Additionally, we performed a weighted median analysis which weights MR estimates by the effect size/standard error. This approach leverages the fact that causal estimates from non-pleiotropic instrumental variable are more probable to converge toward the median. This approach provides reliable causal estimates when less than 50% of the total weight is derived from variants with pleiotropic effects. Additionally, we assessed Cochran’s Q test. The test yields an $I^2$ index which gives an estimate of the amount of heterogeneity amongst IVs ($I^2$ index $\leq$ 25%, is considered low heterogeneity, and a $I^2$ index $> 50%$ is considered high heterogeneity). Finally, we employed MR-PRESSO global test, which can detect the presence of pleiotropic outlier SNPs. By undertaking this variety of sensitivity analyses, and comparing results using these approaches, each one with different underlying assumptions, we could assess, to some degree, whether our findings were biased by horizontal pleiotropy. We considered results as significant when $P$ was $< 0.05$ for all horizontal pleiotropy tests (MR-Egger intercept test, Cochran’s Q test, MR-PRESSO global test).

This study was conducted in accordance with the STROBE-MR guideline (STROBE-MR checklist is provided in Supplementary Material).
RESULTS

Observational associations
All the associations (Table 2) between log transformed 25(OH)D and the seven cognitive scores (the individual scores and the composite score) adjusted for age, sex, smoking status, nutritional index, and BMI classification were Bonferroni significant ($P$-value < 7.1 x 10^{-3}). While the direction of effect between 25(OH)D and the cognition scores was positive, we observed a negative effect between the vitamin and the executive and set switching tests measuring how much time it takes to name ink colours (COLTIME). Thus increased 25(OH)D was consistently associated with improved cognition in observational associations.

Instrumental variables for 25(OH)D
We used 80 common (MAF > 5%) independent SNPs (S1 Table) explaining 3.8% of the variance in 25(OH)D levels, with an $F$-statistic of 214.75 as instruments in our MR study. 5 SNPs were absent from the PACC score GWAS and were replaced by proxies in high LD ($r^2 > 0.8$). TwoSampleMR identified rs10832289, rs11182550, and rs13336347 as palindromic and were therefore removed from further analyses.

GWAS of PACC score
We performed a genome-wide study of the PACC score in European individuals from the CLSA (N=17,166). When we restricted our analysis to individuals aged 60 and older, we found a genome-wide significant association rs429358 ($P$=1.8 x 10^{-8}) at apolipoprotein E (APOE) (Figure 3 C&D). This is a well-characterized locus for dementia and Alzheimer's disease^{35,36,37}.

PACC score and REYI, REYI, AFT, AFT2, MAT, COLTIME score outcome
For the PACC score, MR showed that the genetically predicted increase per 1 SD in natural log of 25(OH)D had no effect on PACC score estimated in the general population ($\text{Beta}_{\text{PACC \cdot All}} = -0.005, 95\% \text{ CI} -0.058 \sim 0.048, P = 0.85$). Results were consistent in individuals over 60 ($\text{Beta}_{\text{PACC \cdot Age \geq 60}} = 0.0009, 95\% \text{ CI} -0.07 \sim 0.07, P = 0.98$) (Figure 1 & Table2).
For REYI, REYII, MAT, and COLTIME, the MR results were consistent with those from the PACC score. There was no discernible effect of genetically predicted increase per 1 SD in the natural log of 25(OH)D on any of these scores (Table 2).

For MR for the AFT, and AFT2 outcomes, we did observe a nominally significant association, which did not survive Bonferroni correction, in the population over 60 years of age. For AFT, the genetically predicted increase per 1 SD in natural log of 25(OH)D has no observed effect on the overall population (Beta\textsubscript{AFT - All} = 0.083, 95% CI -0.0057 – 0.17, \(P = 0.07\)), but did show an effect in individuals over 60 (Beta\textsubscript{AFT2 - Age >60} = 0.13, 95% CI 0.02 – 0.25, \(P = 0.02\)). The trend was consistent for AFT2, Beta\textsubscript{AFT2 - All} = 0.069, 95% CI -0.022 – 0.16, \(P = 0.14\), but there was an uncertain effect in individuals over 60, Beta\textsubscript{AFT2 - Age >60} = 0.12, 95% CI 0.0061 – 0.23, \(P = 0.049\).

**Sensitivity analysis**

We computed MR-PRESSO, Egger’s regression, and Cochran’s Q test to assess the robustness of our results. Our sensitivity analyses compiled in Table 3 for the PACC score and the other cognitive scores show that, by and large, there is low to moderate heterogeneity since \(I^2\) values are below 25%. The only moderate heterogeneity value (\(I^2 = 47\%; I^2_{value} = 0.005\)) was for the REYI score in individuals above 60 years. In fact, the MR-PRESSO global test, which can detect horizontal pleiotropic outlier SNPs, found a pleiotropic SNP for the Rey verbal learning test I. Assessing at the distortion of the causal estimate before and after outlier correction, the MR results remained the same. We found that MR-Egger intercept test, Cochran’s Q test, MR-PRESSO global test and distortion test did not detect horizontal pleiotropy (P-values > 0.05). We therefore did not identify evidence of directional horizontal pleiotropy in the MR-results.
DISCUSSION

In this study, we carried out an observational study followed by a two-sample MR to ascertain the effect of 25(OH)D supplementation on cognitive outcome test scores. First, we computed the linear regression adjusting for age, sex, BMI, smoking status, and nutritional risk index, between 25(OH)D and cognition test scores. Higher 25(OH)D levels were associated with improvements in all cognition measures. We then used MR, which reduces bias from reverse causation and confounding, thus enabling us to test for causal effects. The results of the MR study did not support a causal effect between a genetically determined change in 25(OH)D levels and the PACC score and the cognitive score. These results suggest that observational associations between 25(OH)D levels and improved cognition are likely due to confounding.

Vitamin D supplementation has been proposed for treatment of many common diseases. While supplementation could be beneficial in certain contexts (i.e. prevention of multiple sclerosis, hypovitaminosis D and Alzheimer's disease), results from observational studies need to be considered carefully. For instance, an observational study published in 2014 with a diverse and small cohort (N=382; 41.4% individuals from European ancestry, 29.6% individuals from African American ancestry, 25.1% individuals from Latino/Hispanic ancestry) showed that low vitamin D level is associated with an increased rate of cognitive decline. However, this was not supported by an MR study in a cohort of 153,187 individuals of European ancestry. The study looked at the effect of only two 25(OH)D canonical variants rs12785878 (vitamin D-decreasing allele, G), located near gene coding 7-dehydrocholesterol reductase (DHCR7), and rs12794714 (vitamin D-decreasing allele, A) near 25-hydroxylase (CYP2R1) on global cognition score (Beta: -0.00; 95% CI: (-0.01, 0.01); P-value = 1.0), and memory cognitive score (Beta: -0.00; 95% CI: (-0.01, 0.00); P-value = 0.6). Both the previously published observational studies and MR studies on the effect of 25(OH)D on cognitive scores are congruent with our study.

How 25(OH)D may protect against Alzheimer's disease in human is unknown. MR has shown a protective effect of 25(OH)D on Alzheimer's disease in humans. Additionally, previous and novel evidence demonstrated that cognition is protective for
Alzheimer’s disease\textsuperscript{40, 41}. Therefore, we hypothesized that 25(OH)D may protect against Alzheimer’s disease in humans by increasing cognition. However, our results suggest that the observed MR effects of 25(OH)D on Alzheimer’s disease are explained by improved cognition.

We note several strengths in our study. We leveraged the strengths of the MR design to lessen bias from confounders and reverse causation, therefore enabling more reliable estimate causal effects provided adherence to MR assumptions. To maximize statistical power, we utilized cohorts with large sample sizes.

Our study also contains several limitations. Since MR depends mainly on three causal inference assumptions (relevance, independence, and exclusion restriction) any violation of these could invalidate our results. To address these assumptions, we performed several sensitivity analyses investigating horizontal pleiotropy and heterogeneity (MR-Egger, MR-PRESSO, Cochran’s Q test). Our causal estimate between 25(OH)D and REYI score in older individuals was considered to have a moderate amount of heterogeneity. After the removal of outlying SNP through MR-PRESSO, the causal estimate remained the same. This suggests that the MR results are not likely to be strongly biased by directional horizontal pleiotropy. Further, our study focuses on European ancestry individuals, therefore, conclusions and interpretation of the results are not generalizable to other ancestries. This once again emphasizes the need for large-scale studies assessing the genetic determinants of disease in individuals of non-European ancestry.

In conclusion, our MR study found no evidence for a causal effect between 25(OH)D concentrations and cognitive performance in European ancestry individuals.
FUNDING STATEMENT

The Richards research group is supported by the Canadian Institutes of Health Research (CIHR: 365825; 409511, 100558, 169303), the McGill Interdisciplinary Initiative in Infection and Immunity (MI4), the Lady Davis Institute of the Jewish General Hospital, the Jewish General Hospital Foundation, the Canadian Foundation for Innovation, the NIH Foundation, Cancer Research UK, Genome Québec, the Public Health Agency of Canada, McGill University, Cancer Research UK [grant number C18281/A29019] and the Fonds de Recherche Québec Santé (FRQS). JBR is supported by a FRQS Mérite Clinical Research Scholarship. Support from Calcul Québec and Compute Canada is acknowledged. TwinsUK is funded by the Welcome Trust, Medical Research Council, European Union, the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust in partnership with King’s College London. These funding agencies had no role in the design, implementation or interpretation of this study.

CONFLICT OF INTEREST

JBR’s institution has received investigator-initiated grant funding from Eli Lilly, GlaxoSmithKline and Biogen for projects unrelated to this research. He is the CEO of 5 Prime Sciences (www.5primesciences.com), which provides research services for biotech, pharma and venture capital companies for projects unrelated to this research.

ACKNOWLEDGMENTS

We thank all the participants, the UK Biobank coordinating team, and the CLSA (Project# 1906026 and 190208) coordinating team for their approval, valuable time, generosity, and contributions to the data collection.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.
DATA AVAILABILITY

All GWAS summary statistics used in this study are made publicly available through the GWAS catalog. 25(OH)D and PACC score GWAS summary statistics (All ages and individuals 60 and older) can be downloaded through the following accession: 25(OH)D: GCST010144; PACC_GWAS: GCST90239732; PACC_GWAS_old: GCST90239733. Further inquiry can be directed to the corresponding author.
References


4. Cédric A, Olivier B. Vitamin D in Older Adults: The Need to Specify Standard Values with Respect to Cognition. Published online April 15, 2014. doi:10.3389/fnagi.2014.00072


Table 1. CLSA cognitive measurements used in this study

<table>
<thead>
<tr>
<th>Cognitive Function</th>
<th>Cognitive Domain</th>
<th>CLSA Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>Verbal Episodic Memory</td>
<td>The Rey Auditory Verbal Learning Test I (REYI)</td>
<td>Number of words (or variants) correctly recalled in 90 seconds – Immediate Recall</td>
</tr>
<tr>
<td>Memory</td>
<td>Delayed Recall</td>
<td>The Rey Auditory Verbal Learning Test II (REYII)</td>
<td>Number of words (or variants) correctly recalled in 90 seconds – Delayed Recall</td>
</tr>
<tr>
<td>Executive Function</td>
<td>Executive and Set Switching</td>
<td>Mental Alteration Test (MAT)</td>
<td>Number of correct consecutive numeric and alphabetical alternations in 30 seconds</td>
</tr>
<tr>
<td>Executive Function</td>
<td>Executive and Set Switching</td>
<td>Stroop (Victoria version) (COLTIME)</td>
<td>Time to say ink colour of each colour name (in seconds)</td>
</tr>
<tr>
<td>Executive Function</td>
<td>Verbal Fluency</td>
<td>Animal Fluency (AFT1)</td>
<td>Number of different animals recited in 60 seconds</td>
</tr>
<tr>
<td>Executive Function</td>
<td>Verbal Fluency</td>
<td>Animal Fluency (AFT2)</td>
<td>Number of different animals recited in 60 seconds</td>
</tr>
</tbody>
</table>
Table 2. MR estimates of log(25(OH)D) on cognitive scores in CLSA. Abbreviations: IVW(re): inverse variance weighted multiplicative random effect. Wmedian: weighted median.
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Beta (95% CI)</th>
<th>P-value</th>
<th>Outcome</th>
<th>Beta (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PACC - Age &gt; 60</td>
<td>$9.3 \times 10^{-4}$ (-0.07-0.072) IVW(re)</td>
<td>0.98</td>
<td>PACC - All age</td>
<td>$-0.005$ (-0.058-0.048) IVW(re)</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.98</td>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>$0.044$ (-0.056-0.14) Wmedian</td>
<td>0.39</td>
<td></td>
<td>$0.0038$ (-0.065-0.072) Wmedian</td>
<td>0.91</td>
</tr>
<tr>
<td>AFT - Age &gt; 60</td>
<td>$0.13$ (0.02-0.25) IVW(re)</td>
<td>0.022</td>
<td>AFT - All age</td>
<td>$0.083$ (-0.0057-0.17) IVW(re)</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td>$0.12$ (-0.04-0.28) Egger</td>
<td>0.15</td>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>$0.13$ (-0.024-0.29) Wmedian</td>
<td>0.097</td>
<td></td>
<td>$0.069$ (-0.046-0.18) Wmedian</td>
<td>0.24</td>
</tr>
<tr>
<td>AFT2 - Age &gt; 60</td>
<td>$0.12$ (0.00061-0.23) IVW(re)</td>
<td>0.049</td>
<td>AFT2 - All age</td>
<td>$0.069$ (-0.022-0.16) IVW(re)</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>$0.088$ (-0.071-0.25) Egger</td>
<td>0.28</td>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>$0.1$ (-0.054-0.26) Wmedian</td>
<td>0.2</td>
<td></td>
<td>$0.058$ (-0.06-0.17) Wmedian</td>
<td>0.34</td>
</tr>
<tr>
<td>COLTIME - Age &gt; 60</td>
<td>$-0.047$ (-0.16-0.063) IVW(re)</td>
<td>0.4</td>
<td>COLTIME - All age</td>
<td>$-0.041$ (-0.13-0.046) IVW(re)</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.12</td>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>$-0.17$ (-0.33--0.011) Wmedian</td>
<td>0.036</td>
<td></td>
<td>$-0.063$ (-0.18-0.055) Wmedian</td>
<td>0.29</td>
</tr>
<tr>
<td>MAT - Age &gt; 60</td>
<td>$0.0046$ (-0.1-0.11) IVW(re)</td>
<td>0.93</td>
<td>MAT - All age</td>
<td>$0.013$ (-0.064-0.09) IVW(re)</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.91</td>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>$0.029$ (-0.13-0.19) Wmedian</td>
<td>0.71</td>
<td></td>
<td>$-0.0085$ (-0.13-0.11) Wmedian</td>
<td>0.89</td>
</tr>
<tr>
<td>REYI - Age &gt; 60</td>
<td>$0.021$ (-0.11-0.15) IVW(re)</td>
<td>0.75</td>
<td>REYI - All age</td>
<td>$0.018$ (-0.071-0.11) IVW(re)</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.7</td>
<td></td>
<td>$9.6 \times 10^{-4}$ (-0.097-0.099) Egger</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>$0.022$ (-0.13-0.17) Wmedian</td>
<td>0.78</td>
<td></td>
<td>$0.015$ (-0.098-0.13) Wmedian</td>
<td>0.8</td>
</tr>
<tr>
<td>REYII - Age &gt; 60</td>
<td>$-0.02$ (-0.14-0.096) IVW(re)</td>
<td>0.74</td>
<td>REYII - All age</td>
<td>$0.0076$ (-0.076-0.092) IVW(re)</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>$0.028$ (-0.13-0.19) Egger</td>
<td>0.73</td>
<td></td>
<td>$0.056$ (-0.059-0.17) Egger</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>$0.12$ (-0.044-0.29) Wmedian</td>
<td>0.15</td>
<td></td>
<td>$0.058$ (-0.06-0.18) Wmedian</td>
<td>0.34</td>
</tr>
</tbody>
</table>
Table 3. Observational associations between log(25(OH)D) and cognitive scores in CLSA.

Associations were adjusted for age, sex, smoking status, BMI classification, and high nutritional risk indicator using linear regression. Scores are in 1 standard deviation, and vitamin D is Log 25(OH)D. Abbreviations: AFT: Number of different animals recited in 60 seconds; REYI: Number of words (or variants) correctly recalled in 90 seconds (immediate recall); REYII: Number of words recalled in 90 seconds (delayed recall); MAT: Number of correct consecutive numeric and alphabetical alternations in 30 seconds; COLTIME: Time to say ink color of each color name (in seconds); N: sample size; Beta (95%CI): effect size of Log 25(OH)D on standard deviation of each cognitive score. 95% confidence intervals are displayed in parantheses.

<table>
<thead>
<tr>
<th>Cognitive trait</th>
<th>N</th>
<th>Beta (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PACC</td>
<td>23,499</td>
<td>0.05 (0.05 – 0.11)</td>
<td>1.8 x 10^{-10}</td>
</tr>
<tr>
<td>AFT1</td>
<td>24,954</td>
<td>0.08 (0.05 – 0.11)</td>
<td>2.8 x 10^{-9}</td>
</tr>
<tr>
<td>AFT2</td>
<td>24,954</td>
<td>0.076 (0.29 – 0.65)</td>
<td>3.9 x 10^{-8}</td>
</tr>
<tr>
<td>REYI</td>
<td>24,709</td>
<td>0.06 (0.04 -0.09)</td>
<td>1.5 x 10^{-6}</td>
</tr>
<tr>
<td>REYII</td>
<td>24,698</td>
<td>0.049 (0.023- 0.075)</td>
<td>2.4 x 10^{-4}</td>
</tr>
<tr>
<td>MAT</td>
<td>24,348</td>
<td>0.04 (0.01 - 0.068)</td>
<td>4.9 x 10^{-3}</td>
</tr>
<tr>
<td>COLTIME</td>
<td>25,357</td>
<td>-0.068 (-0.09 - -0.044)</td>
<td>5.9 x 10^{-8}</td>
</tr>
</tbody>
</table>
### Table 4. MR sensitivity analysis

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Cochran Q</th>
<th>MR-Egger</th>
<th>MRPRESSO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$I^2$</td>
<td>$I^2$ p-value</td>
<td>Egger intercept</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>PACC - Age &gt; 60</td>
<td>19.7</td>
<td>0.12</td>
<td>-1.80 x 10^{-6}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>PACC- All age</td>
<td>19.7</td>
<td>0.11</td>
<td>-6.00 x 10^{-4}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>AFT2_score - Age &gt; 60</td>
<td>13.2</td>
<td>0.21</td>
<td>1.5 x 10^{-3}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>AFT2_score - All age</td>
<td>27.6</td>
<td>0.056</td>
<td>-2.6 x 10^{-4}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>AFT_score - Age &gt; 60</td>
<td>10.5</td>
<td>0.24</td>
<td>7.9 x 10^{-4}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>AFT_score - All age</td>
<td>22.3</td>
<td>0.087</td>
<td>-1.0 x 10^{-3}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>COLTIME_score - Age &gt; 60</td>
<td>10.5</td>
<td>0.25</td>
<td>3.9 x 10^{-3}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>COLTIME_score- All age</td>
<td>21.1</td>
<td>0.097</td>
<td>3.3 x 10^{-3}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>MAT_score - Age &gt; 60</td>
<td>0</td>
<td>0.69</td>
<td>-2.2 x 10^{-4}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>MAT_score- All age</td>
<td>0</td>
<td>0.71</td>
<td>-2.5 x 10^{-4}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>REYI_score - Age &gt; 60</td>
<td>47.4</td>
<td>0.0049</td>
<td>-7.5 x 10^{-4}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>REYI_score - All age</td>
<td>26.3</td>
<td>0.057</td>
<td>-1.2 x 10^{-4}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>REYII_score - Age &gt; 60</td>
<td>22.4</td>
<td>0.095</td>
<td>-2.6 x 10^{-3}</td>
</tr>
<tr>
<td>25(OH)D</td>
<td>REYII_score- All age</td>
<td>15.8</td>
<td>0.16</td>
<td>-2.6 x 10^{-3}</td>
</tr>
</tbody>
</table>
Figure 1 Forest plot of MR studies investigating the effect of the natural log transformed 25(OH)D on PACC, AFT, AFT2, REYI, REYII, COLTIME, and MAT scores.

Forest plots of the main study and the sensitivity analyses. Significance threshold as a Bonferroni $P < 7.1 \times 10^{-3}$

Abbreviations: MR: mendelian randomization; IVW(re) random effect inverse weighted variance; Wmedian, weighted median. $25\text{OHD}$, 25-hydroxyvitamin. PACC, AFT, AFT2, REYI, REYII, COLTIME, and MAT are defined in Table 1.
Figure 2. Scatter plots of the natural log 25 (OH)D MR analyses for (A) PACC score for all individuals, (B) PACC score for aged 60 years. Each dot represents a genetic instrumental variable. Three lines represent causal estimate (βIV) by the inverse variance weighted (IVW) method (pink), MR–Egger method (blue), and Weighted median method (orange). Error bars represent 95%CIs. MR, Mendelian randomization. For visualization purposes, (B) the IVW line was dashed since it overlaps with the MR-Egger line. A genetically predicted change in the natural log of 25(OH) D does not have an effect on a 1 SD of PACC score (for all individuals: beta = -0.0050, 95% CI: -0.056–0.048, P = 0.85; For individuals 60 years and more: beta = 0.00093, 95% CI: -0.070–0.071, P = 0.98)
Figure 3. PACC score GWAS. A) Manhattan plot for PACC score ($P_{\text{significance}} = 5 \times 10^{-8}; \ N=17,166$). B) QQplot of PACC score stratified by allele frequency. C) Manhattan plot for PACC score for individuals age $\geq 60$ ($P_{\text{significance}} = 5 \times 10^{-8}; \ N=9,489$). D) QQplot of PACC score stratified by allele frequency for individuals age $\geq 60$. Black points are for all SNP regardless of allelic frequency, blue points are for SNPs with MAF $> 1\%$, and red points are for SNPs with MAF $< 1\%$. All associations are adjusted the PACC cognition score for age, sex, and the first 20 principal components.