

Investigating Causal Relations Between Circulating Metabolites and Alzheimer's Diseases: a Mendelian Randomization Study

Shu-Yi Huang

Fudan University Huashan Hospital Department of Neurology

Yu-Xiang Yang

Fudan University Huashan Hospital Department of Neurology

Ya-Ru Zhang

Fudan University Huashan Hospital Department of Neurology

Kevin Kuo

Fudan University Huashan Hospital Department of Neurology

Hong-Qi Li

Fudan University Huashan Hospital Department of Neurology

Xue-Ning Shen

Fudan University Huashan Hospital Department of Neurology

Shi-Dong Chen

Fudan University Huashan Hospital Department of Neurology

Ke-Liang Chen

Fudan University Huashan Hospital Department of Neurology

Qiang Dong

Fudan University Huashan Hospital Department of Neurology

Lan Tan

Fudan University Huashan Hospital Department of Neurology

Jin-Tai Yu (✉ jintai_yu@fudan.edu.cn)

Huashan Hospital, Fudan University <https://orcid.org/0000-0002-7686-0547>

Research Article

Keywords: Metabolite, Mendelian randomization, Alzheimer's disease

Posted Date: April 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-467119/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Metabolomics is a promising approach that can be used to understand pathophysiological pathways of Alzheimer disease (AD). However, the relationships between metabolism and AD are poorly understood. The aim of this study is to investigate the causal association between circulating metabolites and risk of AD by combining metabolomics with genomics through two-sample Mendelian randomization (MR) approach.

Method

Genetic associations with 123 circulating metabolic traits were utilized as exposures. A large summary statistics data from International Genomics of Alzheimer's Project was used in primary analysis, including 21,982 AD cases and 41,944 controls. Validation was further performed using family history of AD data from UK Biobank (27,696 cases of maternal AD, 14,338 cases of paternal AD and 272,244 controls). We utilized the inverse-variance weighted method as primary analysis and four additional MR methods (MR-Egger, weighted median, weighted mode, and MR pleiotropy residual sum and outlier) for sensitivity analyses.

Results

We found one-fold increased risk of developing AD per standard deviation increase in the levels of circulating ApoB (odds ratio (OR)=3.18; 95% confidence interval (CI): 1.52–6.66, P=0.0022) and glycoprotein acetyls (OR=1.21; 95% CI: 1.05–1.39, P=0.0093), serum total cholesterol (OR=2.73; 95% CI: 1.41-5.30, P=0.0030), and low-density lipoprotein (LDL) cholesterol (OR=2.34; 95% CI: 1.53-3.57, P=0.0001). Whereas glutamine (OR=0.81; 95% CI: 0.71-0.92, P=0.0011) were significantly associated with lower risk of AD. We also detected causal effects of several different composition of LDL fractions on increased AD risk, which has been verified in validation. However, we found no association between circulating high-density lipoprotein cholesterol and AD.

Conclusions

Our findings provided robust evidence supporting causal effects of circulating glycoprotein acetyls, ApoB, LDL cholesterol, and serum total cholesterol on higher risk of AD, whereas glutamine showed the protective effect. Further research is required to decipher the biological pathways underpinning associations.

Background

Alzheimer's disease (AD) is the leading cause of dementia, affecting dramatically increased number of aging populations and placing an enormous burden on patients, families, and health-care systems [1]. The hallmarks of AD include amyloid plaques, tau neurofibrillary tangles, neurodegeneration, and

synaptic loss [1, 2]. Unfortunately, there is currently no effective prevention and treatment for AD [1]. Metabolomics is the newest systematic biology approach which measures the biochemical products of cell processes downstream of genomic, transcriptomic, proteomic systems, and influences from the environment. It offers great potential for the diagnosis and prognosis of neurodegenerative diseases by capturing snapshots of the complex and multifactorial biochemical pathways that may be altered in AD [3]. Recent studies have showed that metabolomics can be used to measure alterations in biochemical pathways related to AD [4–7]. However, one of the key challenges in these metabolomics studies is the inability to ensure whether the relationships between circulating metabolites and AD are causal. It is of great importance to understand the causality between metabolites and AD, as well as the potential pathophysiological pathways of AD, to inspire drug discovery and to detect biomarkers that aid in early detection of high-risk individuals to initiate prevention, monitoring, and treatment.

Mendelian randomization (MR) is an analytic approach that uses genetic variants as instrumental variables (IVs) to assess causal inferences between an exposure and the outcome of interest [8]. The MR approach is largely independent of unmeasured confounding biases and reverse causality inherent in observational studies, given that allocation of genotypes from parents to offspring is random and genetic variation is unlikely to be affected by environmental factors [8, 9]. This is particularly relevant in a metabolomic study where the inter-individual variability of circulating metabolites can be affected by a wide range of potential confounders such as age, sex, physical diseases, medication, exercise, weight, and time of sampling. In addition, two-sample MR analysis is an extension in which the effects of the genetic instrument on the exposure and on the outcome can be obtained from publicly available genome-wide association studies (GWAS) summary data [10]. The two-sample MR approach enables us to link circulating metabolites with risk of AD using GWAS estimates on both metabolic phenotypes and AD.

The aim of this study was applying a two-sample MR approach to investigate the causal relationship between circulating metabolites and risk of AD, combining genome-wide and metabolome-wide datasets generated from large scale cohorts.

Method

Genetically determined metabolites

We obtained the summary statistics from a GWAS for 123 circulating metabolites [11]. Briefly, *Kettunen et al.* conducted a comprehensive GWAS estimated with quantitative human serum/plasma metabolites as phenotypes [11]. Metabolomics data were acquired based on human fasting blood samples, otherwise the fasting time effect were adjusted in original study. The 123 metabolites represent a broad molecular signature of systemic metabolism and were assigned to 12 classes (carboxylic acids and derivatives, fatty acyls, glycerolipids, glycerophospholipids, hydroxy acids and derivatives, keto acids and derivatives, lipoprotein, organooxygen compounds, protein, ratio, sphingolipids, steroids and steroid derivatives) based on human metabolome database classification and expert opinion [12, 13]. A total of up to 24,925 individuals from 14 Europe cohorts were meta-analyzed. The mean age was 44.6 years, and females

accounted for 54.6%. Written informed consent was obtained from all participants, and the study was approved by ethical committees.

Selection of instrumental variables

For each of the 123 metabolites, single nucleotide polymorphisms (SNPs) associated at genome-wide significance P-value ($P < 5 \times 10^{-8}$) with a minor allele frequency greater than 0.01 were considered as potential instruments. Independent SNPs were selected at a threshold of linkage disequilibrium $r^2 > 0.05$ and a distance of 1000kb. For palindromic SNPs, we aligned strands using allele frequency and discarded palindromic SNP(s) that had minor allele frequency above 0.42. Then exposure–outcome datasets were harmonized. We have considered the palindromic SNPs and checked original datasets to avoid reverse effects.

The proportion of variance explained by IVs were computed. And F-statistic of each metabolite was calculated to judge the strength of IVs. Typically, a strong instrument was defined as an F-statistic > 10 [14]. Additionally, power calculations were conducted using the R code provided by Burgess S with a two-sided type-I error rate $\alpha = 0.05$ [15]. The proportion of variance explained by IVs, F-statistics and power were presented in Additional file 1.

IGAP AD dataset

In primary analysis, genetic variants associated with late-onset AD were obtained from a meta-analysis GWAS performed by International Genomics of Alzheimer's Project (IGAP) [16]. There is no sample overlap of IGAP with cohorts of circulating metabolites. IGAP is a large two-stage study based upon GWAS on individuals of European ancestry. Data from stage 1 was used in the present study, including 63,926 individuals (21,982 AD cases and 41,944 cognitively normal controls) of European descent from four consortia: Alzheimer Disease Genetics Consortium (ADGC), Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE), The European Alzheimer's Disease Initiative (EADI), and the Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES) [16]. All AD diagnoses were autopsy-confirmed or satisfied the NINCDS-ADRDA criteria or DSM-IV guidelines [16]. The average age at onset of AD ranged from 74.4 to 81.9 years, and the average age at examination for 83% controls is ≥ 76 years.

UK Biobank proxy-AD dataset

We additionally set out to validate our results using a proxy-AD dataset, based on individuals in the UK Biobank (<http://www.ukbiobank.ac.uk>) [17, 18]. The high heritability of AD implies that individuals with family history of AD are likely to have a higher genetic AD risk load. Thus, individuals with one or two parents with AD were defined as proxy cases, that is, having family history of AD [19]. In this dataset, the proxy-AD case-control status was ascertained via self-report. Over 500,000 community-dwelling individuals aged between 37 and 73 years were recruited in the United Kingdom between 2006 and 2010 [17]. An array of 314,278 participants with available AD information on at least one parent were meta-

analyzed in this analysis, including 27,696 cases of maternal AD, 14,338 cases of paternal AD, and 272,244 controls [18]. All data sources used in this MR study received approval from an ethics standards committee on human experimentation and obtained informed consent from all participants.

Statistical analysis for Mendelian randomization

We used inverse-variance weighted (IVW) method as the primary analysis to determine the causal relationships between genetically determined circulating metabolites and AD. The IVW method will return an unbiased estimate in the absence of horizontal pleiotropy or when horizontal pleiotropy is balanced [20]. Results are presented as odds ratio (OR) per standard deviation (SD) increase in genetically determined metabolites on AD for the outcome was dichotomous.

We conducted sensitivity analysis using weighted median [21], weighted mode-based estimate (MBE) [22], MR Egger regression [23], and Mendelian randomization pleiotropy residual sum and outlier (MR-PRESSO) [24], and displayed in scatter plots. These methods hold different assumptions at the costs of reduced statistical power. The weighted median allows for 50% of the IVs to be invalid or present pleiotropy [21]. The weighted MBE method used the mode of the IVW empirical density function as the weighted MBEs and obtained a causal effect estimate robust to horizontal pleiotropy [22]. MR-Egger regression allows >50% of the variants to be invalid [23]. MR-Egger is based on the “NO Measurement Error” (NOME) assumption (no measurement error in the SNP exposure effects), which is evaluated by the regression dilution I^2_{GX} statistic (i.e., less than 0.9 indicates a violation of NOME) [25]. Thus, an I^2_{GX} statistic was calculated to test the presence of bias with MR-Egger.

We checked each SNP prioritized in the present study using the GWAS Catalog (<http://www.ebi.ac.uk/gwas>) to ensure no genetic instruments were associated directly with AD. SNPs significantly associated with AD were excluded from the IVs. Where SNPs significantly ($P < 5 \times 10^{-8}$) associated with AD are detected, the analysis was re-run to determine whether removing those potential pleiotropic SNPs impacted the effects.

Furthermore, we conducted MR Egger intercept and MR-PRESSO global test to detect the presence of horizontal pleiotropy or heterogeneity. In the case of horizontal pleiotropy, MR-PRESSO outlier test compares the observed and expected distributions of the tested variants to identify outlier variants. If significant outliers ($P < 0.05$) are detected, they were removed from the analysis to return an unbiased causal estimate [24]. Heterogeneity in the IVW estimates was tested using Cochran’s Q test, and displayed in forest plots. Moreover, visual inspection of funnel plot and leave-one-out plot were also used to assess the MR “no horizontal pleiotropy” assumption (see Additional file 6).

To correct for multiple comparisons, we applied false discovery rates (FDR) correction in IVW. An FDR corrected P-value < 0.05 was considered significant, and an unadjusted P-value < 0.05 was considered as the evidence of a suggestive association. Those metabolites that showed suggestive evidence of association ($P_{IVW} < 0.05$) with late-onset AD were assessed for validation using UK Biobank GWAS.

Results with consistent direction of point estimates across sensitivity analyses, validation, IVW, and estimates after correction of potential pleiotropy were considered as robust causal associations.

Analyses were conducted using R version 3.6.3, with the MR analysis performed using the “TwoSampleMR” package version 0.5.2 [26, 27].

Pathway enrichment analysis

Metabolic enrichment analysis was conducted using the web-based MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca/>). We analyzed the metabolite sets from Kyoto Encyclopedia of Genes and Genomes (KEGG) Database in our study [28]. All metabolites that were associated with AD risk with a P-value <0.05 were used to identify metabolic pathway.

Results

A flow diagram depicting the process of the MR analyses is shown in Fig 1. The IVW identified 32 circulating metabolites that were associated with AD risk, including 27 significant (FDR $P_{IVW} < 0.05$) traits and five suggestive ($P_{IVW} < 0.05$) causal traits (Fig 2). Of the 32 metabolites, the IVs were composed of 6–52 SNPs and could explain 3–20% of the variance of the corresponding metabolites (see Additional file 1). And the F-statistics were all highly above the threshold of weak instruments of F-statistic < 10 (see Additional file 1). Results of sensitivity analyses were showed in Additional file 2. The negative MR results were presented in Additional file 3.

Effects of Genetically Determined Metabolites on AD

Those associated metabolites include two proteins, two carboxylic acids and derivatives, 26 lipoproteins, and two steroids and steroid derivatives (Fig 2). We observed a 1-fold increased risk of developing AD per SD increase in the level of circulating ApoB (OR = 3.18; 95% confidence interval (CI): 1.52–6.66, $P_{IVW} = 0.0022$) and glycoprotein acetyls (Gp; OR = 1.21; 95% CI: 1.05–1.39, $P_{IVW} = 0.0093$), serum total cholesterol (serum-C; OR = 2.73; 95% CI: 1.41–5.30, $P_{IVW} = 0.0030$) and esterified cholesterol (OR = 2.42; 95% CI: 1.07–5.48, $P_{IVW} = 0.0345$). Citrate (OR = 0.83; 95% CI: 0.73–0.95, $P_{IVW} = 0.0055$) and glutamine (OR = 0.81; 95% CI: 0.71–0.92, $P_{IVW} = 0.0011$) were significantly associated with lower risk of AD. In case of lipoproteins, five intermediate-density lipoprotein (IDL) subfractions (i.e., IDL-C, IDL-FC, IDL-L, IDL-P, and IDL-PL) and 15 low-density lipoprotein (LDL) subfractions (i.e., LDL-C, L-LDL-C, L-LDL-CE, L-LDL-FC, L-LDL-L, L-LDL-P, L-LDL-PL, M-LDL-C, M-LDL-CE, M-LDL-L, M-LDL-P, M-LDL-PL, S-LDL-C, S-LDL-L, and S-LDL-P) were significantly associated with higher risk of AD. Four high-density lipoprotein (HDL) subfractions (i.e., XL-HDL-C, XL-HDL-CE, XL-HDL-FC, XL-HDL-P) were associated with lower risk of AD, while S-HDL-P showed association with higher risk of AD at a suggestive level. Besides, XS-VLDL-PL was significantly associated with higher risk of AD using IVW method (OR = 1.49; 95% CI: 1.11–2.02, $P_{IVW} = 0.0087$).

Pleiotropy and heterogeneity analysis

The Q-test detected no evidence of heterogeneity in the results of Gp, citrate, glutamine, S-HDL-P, XL-HDL-C, XL-HDL-CE, and XL-HDL-FC (Table 1). By checking each SNP in GWAS Catalog, nine AD-associated SNP were identified from IVs of 24 metabolites (see Additional file 4). IVW results of the 24 metabolites after removing those potential pleiotropy SNPs were showed in Additional file 5. Although this sensitivity results were all non-significant, the direction of point estimates was consistent with that in primary analysis. Only one metabolite, esterified cholesterol, showed opposite direction of point estimate (OR = 0.95) to the primary result (OR = 2.42; 95% CI: 1.07–5.48), indicating potential bias of the effect.

The MR Egger intercept test showed no evidence of pleiotropy in these IVW-identified metabolites (Table 1). MR-PRESSO global test also showed no pleiotropy in the results of Gp, citrate, glutamine, XL-HDL-C, XL-HDL-CE and serum-C. Others were analyzed by MR-PRESSO after removing the outlying SNPs identified by MR-PRESSO outlier test. The corrected results were significant except for XS-VLDL-PL ($P_{\text{cMR-PRESSO}} = 0.396$) and IDL-FC ($P_{\text{cMR-PRESSO}} = 0.074$). The corrected point estimates of MR-PRESSO were all consistent with IVW except for XL-HDL-FC ($OR_{\text{cMR-PRESSO}} = 1.64$ vs. $OR_{\text{IVW}} = 0.90$). Thus, effects of XS-VLDL-PL, IDL-FC, and XL-HDL-FC were considered unreliable.

Sensitivity analysis

The results for the sensitivity analyses of 32 IVW-identified metabolites are showed in Fig 4 and Additional file 2. The I^2_{GX} statistics of each MR-Egger result are all above 0.9. ApoB, Gp, glutamine, and serum-C present robust associations that all additional MR methods showed consistent direction of point estimates with IVW, though not all MR methods yielded significant results. However, as for citrate, the sensitivity results showed opposite effect estimates ($OR_{\text{weighted-median}} = 1.06$, $OR_{\text{weighted-mode}} = 1.05$) to primary result ($OR_{\text{IVW}} = 0.83$). In case of lipoproteins, an array of 16 IVW-identified significant metabolites (i.e., IDL-C, IDL-L, IDL-P, LDL-C, L-LDL-C, L-LDL-CE, L-LDL-FC, L-LDL-L, L-LDL-P, L-LDL-PL, M-LDL-CE, M-LDL-L, M-LDL-P, M-LDL-PL, S-LDL-C, and S-LDL-P) also showed concordant results across sensitivity methods and IVW. Sensitivity analysis indicates non-robust effects of S-LDL-L ($OR_{\text{weighted-median}} = 0.80$, $OR_{\text{weighted-mode}} = 0.80$ vs $OR_{\text{IVW}} = 2.69$), XL-HDL-CE ($OR_{\text{weighted-median}} = 1.20$, $OR_{\text{weighted-mode}} = 1.23$ vs $OR_{\text{IVW}} = 0.90$), XL-HDL-FC ($OR_{\text{MR-PRESSO}} = 0.0002$ vs $OR_{\text{IVW}} = 0.90$) that significant sensitivity results yielded discordant point estimates with IVW. Another six lipoproteins, IDL-FC, IDL-PL, M-LDL-C, S-HDL-P, XL-HDL-C, and XL-HDL-P also showed inconsistent results of weighted median and weighted mode with IVW.

Validation

In validation, we used UK Biobank dataset to verify the association between IVW-identified metabolites and proxy-AD (Fig 3). Of the 32 identified metabolites, 20 were still significant associated with proxy AD in validation, especially including ApoB, Gp, and LDL-C, and several different composition of LDL fractions. Among 32 selected metabolites, direction of point estimates in validation were all accordant with primary results in IVW except for the citrate ($OR_{\text{validation}} = 1.029$), which yields reversed effect with risk of AD ($OR_{\text{IVW}} = 0.83$).

Pathway enrichment analysis

Our study identified six significant metabolic pathways that were involved in the pathogenesis of AD (Table 2). The most significant metabolic pathway was D-glutamine and D-glutamate metabolism ($P = 1.30 \times 10^{-5}$) from the KEGG database. L-glutamine and L-glutamate were involved in this metabolic pathway. Another two metabolic pathways involving two circulating metabolites (i.e., glutamine and citrate) survived FDR correction, that is, the pathway of alanine, aspartate, and glutamate metabolism ($P = 0.0003$), and glyoxylate and dicarboxylate metabolism ($P = 0.0004$). We also identified three pathways at the nominal $P < 0.05$, including nitrogen metabolism ($P = 0.008$), arginine biosynthesis ($P = 0.018$), and citrate cycle (i.e., tricarboxylic acid (TCA) cycle; $P = 0.026$).

Discussion

By performing a two-sample MR analysis, the present study supports the hypothesis that circulating metabolites levels can be causally correlated to risk of AD. We suggest a significant association between higher levels of Gp and higher risk of AD, and genetically predicted glutamine levels are significantly associated with lower risk of AD. Our results also reinforce the idea that circulating lipid-related metabolites may play a role in the pathophysiological process of AD. Particularly, we observed robust evidences of causal effects with respect to ApoB, serum-C, three IDL subfractions (i.e., IDL-C, IDL-L, IDL-P), and 13 LDL subfractions (i.e., LDL-C, L-LDL-C, L-LDL-CE, L-LDL-FC, L-LDL-L, L-LDL-P, L-LDL-PL, M-LDL-CE, M-LDL-L, M-LDL-P, M-LDL-PL, S-LDL-C, and S-LDL-P) on higher risk of AD. To our knowledge, this is the most comprehensive MR analysis to examine the causal associations of circulating metabolites and risk of AD.

The measured Gp are mainly α -1-acid glycoprotein (AGP) [11], also called orosomucoid. Gp is an acute phase plasma α -globulin glycoprotein, involving in many activities including modulating immunity, binding and carrying drugs, maintaining the barrier function of capillary, and mediating the sphingolipid metabolism [29–31]. Gp is associated with AD due to its important role in modulating neuroinflammation [32]. Higher levels of plasma AGP were found in patients with cognitive impairment than in normal subjects [32]. Previous meta-analysis has reported that plasma levels of Gp were associated with increased risk of dementia and lower cognitive function [33]. These results support our findings suggesting a relationship between circulating Gp and AD.

ApoB is synthesized in the liver and circulates in the plasma as the major protein component of LDL, involving in the transport of cholesterol to peripheral tissues [34]. Previous studies have demonstrated that AD group has significantly higher levels ApoB in serum [35, 36] and plasma [37] than that of the control group, especially in AD subjects with APOE ϵ 4 allele [38]. In AD patients, higher serum levels of ApoB are significantly correlated with higher A β 42 levels in brain [36]. Additionally, genetic variants in the gene of *APOB* are strongly associated with early-onset AD [39], suggesting a link of ApoB to AD risk. However, previous studies of circulating ApoB levels in human are conflicting, with a large population study finding no association between circulating ApoB levels and incident AD [40]. Therefore, our MR

results is a significant evidence enhancing the association between circulating ApoB and AD and suggesting it as causality.

It is reasonable of our findings that many biological studies have reported coincident evidences. Plasma ApoB was found co-localized with cerebral amyloid plaque in a transgenic mouse AD model [41], and was positively correlated with Ab plaque abundance in brain [42]. Overexpressing *APOB* in a transgenic mouse model induces significant memory impairment and increases A β levels compared with wild-type mice, suggesting that increased ApoB levels can contribute to the development of AD-like pathology [43].

Whereas ApoB is involved in LDL-C metabolism and is regarded as a promising link between cholesterol and AD [44], many epidemiological evidences of association between LDL-C and AD are consistent with that of ApoB [36]. Observational studies have indicated that LDL-C levels were significantly increased in AD patients [45–47]. Likewise, Zhou et al. suggest that elevated concentration of LDL-C (> 121 mg/dl) may be a potential risk factor for AD [48]. Our MR analysis support these results and suggest a causal effect of high circulating LDL-C levels in increasing risk of AD. Consistent with our findings, another two published MR study also revealed similar effects of LDL-C using different datasets [49, 50], enhancing reliability of the results.

According to the molecular size, LDLs are further categorized as large (L), medium (M), and small (S) LDLs in initial study [11]. Variation in circulating levels and composition of these fractions may have different pathophysiologic significance. Particularly, plasma levels of L-LDL particles were significantly associated with greater cerebral amyloidosis and lower hippocampal volumes independent of LDL-C [51]. Except for LDL-C, our findings suggest that six L-LDL subfractions, four M-LDL subfractions, and two S-LDL subfractions can influence the AD risk, but further investigations are needed to fully understand the molecular mechanisms involved.

In results of observational studies, the effects of serum-C on risk of AD were highly heterogeneous [52]. With respect to serum-C, several meta-analyses revealed non-significant effect on AD [53]. While other epidemiological studies reported that serum-C levels were significantly increased in AD patients [45–47]. The significance of serum TC differs between mid-life and older adults [54]. Several studies state that high mid-life serum TC levels represent a risk factor for subsequent AD [55], but that there are no detectable differences in serum TC levels at older ages [56]. Additionally, except for long-term average serum TC levels, higher TC variability is significantly associated with increased risk of all-cause dementia and AD in the general population, independent of mean TC levels [57]. Thus, the in-coincident results between serum-C studies may be explained by the variations in total cholesterol levels and the disease progression. Taking the advantage of not being affected by unmeasured confounders inherent in observational studies [8], our MR results are more robust, suggesting the high serum-C levels may have a causal effect in increasing AD risk.

Many epidemiological evidences suggest a protective association of circulating HDL cholesterol (HDL-C) levels against AD risk [58]. While we found HDL-C levels are not associated with AD risk with enough power (see Additional file 3), concordant with a large population study [40]. In our study, four very large

HDL subfractions (i.e., XL-HDL-C, XL-HDL-CE, XL-HDL-FC, and XL-HDL-P) yield inverse effects in AD risk, however, these sensitivity results showed inconsistent effect estimates against IVW. Investigations are needed to further clarify whether the relationship between HDL and AD are causal.

Our study also reported several metabolic pathways that might be involved in the pathogenesis of AD, in which the D-glutamine and D-glutamate metabolism has been reported to be associated with AD [4, 59]. Observational study showed that glutamine concentrations in plasma is positively correlated with that in posterior cingulate cortex [60], which is associated with cognitive impairment in AD [61]. We found consistent results that genetically determined circulating glutamine show a protective effect against AD. Nevertheless, a cohort study found higher glutamine levels were associated with lower cognitive function and higher risk of dementia [33]. Whereas observational studies are prone to reverse causation and confounding bias, an MR analysis with balanced horizontal pleiotropy is more credible [9]. Consistent with our results, a published two-sample MR study came to a similar conclusion of glutamine using a different AD dataset [62]. Furthermore, by conducting a series of rigorous sensitivity, pleiotropy, and validation analyses, our results are more comprehensive and robust. Moreover, there also exist biological evidences of this result. Anderson et al. have observed that reduced glutamine metabolism, reduced TCA activity, and impaired oxidative glutamine metabolism precede amyloid plaque formation in AD mouse model compared to controls [63]. And glutamine is proved to protect against oxidative stress-induced injury that is intimately related to AD in AD mice model [64].

Citrate is key constituent of the TCA cycle, serves as a substrate in the cellular energy metabolism cycle involved in the fatty acid synthesis, glycolysis, and gluconeogenesis [65]. There are very few researches exploring the relationship between citrate and AD. However, our current analysis found a protective effect of citrate in AD risk at significant level. Although additional evidence is needed, it might provide valuable information to help understand the underlying biological mechanisms in the pathogenesis of AD.

Our study also has several limitations. First, a general challenge of MR is the persistent possibility of horizontal pleiotropic associations between exposure and outcome. In the present study, we conducted up-to-date analyses to detect and correct the potential pleiotropy. One limitation is that the Q-test and MR-PRESSO global test is significant in some metabolites. Nevertheless, MR-PRESSO outlier test was further performed to correct for horizontal pleiotropy and returned an unbiased causal estimate. Second, some metabolites yield opposite direction of effect estimates across sensitivity analysis and IVW. It is generally recommended that the emphasis of sensitivity analysis should be laid on the direction of point estimates among the IVW and sensitivity analyses, rather than just the P values. Although this standard ruled out several metabolites from robust results, a serious screening protocol ensure the reliability of our results. For instance, the Gp and glutamine showed the most robust casualty. However, we didn't have enough evidence that those "non-robust" metabolites are not associated with AD. Third, we used a proxy-AD GWAS dataset to verify our analysis. Hence, the phenotypes used for validation were different from that used in primary analysis, resulting in smaller effect sizes. However, the validation is an independent replication analysis for that there is no overlap between primary AD dataset and validation dataset.

Despite these limitations, strengths of the study are notable. Our study provided novel insight by combining metabolomics with genomics to help understand the pathogenesis of AD. The use of two sample MR approach also enabled us to use the very large AD case–control data, giving sufficient power to detect even small effects. And there is no overlapping among exposure and outcome datasets, as is unachievable in many MR studies that may bias effect estimates. Stronger evidence of causal relationships is of great importance because the AD underlying pathophysiological mechanisms are unclear. If these circulating metabolites levels truly reduce AD risk, it would be promising markers for early detection and potential avenues for effective therapeutic intervention in AD.

Conclusions

In conclusion, our study suggested increased levels of circulating Gp, ApoB, LDL-C, and serum-C were the most robust metabolites that were associated with higher risk of AD, whereas glutamine showed the contrary effect. We found strong evidence for causal effects of several different composition of LDL fractions on increased AD risk. The present study provides little evidence that recommending circulating HDL-C would help to prevent AD. Further research is required to decipher the biological pathways underpinning associations.

Abbreviations

AD, Alzheimer's disease; ADGC, Alzheimer Disease Genetics Consortium; AGP, α -1-acid glycoprotein; CI, confidence interval; CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium; EADI, The European Alzheimer's Disease Initiative; FDR, false discovery rates; GERAD/PERADES, Genetic and Environmental Risk in Alzheimer's Disease /Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium; GWAS, genome-wide association studies; Gp, glycoprotein acetyls; HDL, high-density lipoprotein; IDL, intermediate-density lipoprotein; KEGG: Kyoto Encyclopedia of Genes and Genomes; LDL, low-density lipoprotein; NOME, NO Measurement Error; OR, odds ratio; SD, standard deviation; SNP, single nucleotide polymorphism; IGAP, International Genomics of Alzheimer's Project; IV, instrumental variable; IVW, inverse variance weighting; MBE, mode-based estimate; MR, Mendelian randomization; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier; serum-C, serum total cholesterol; TCA, tricarboxylic acid.

Declarations

Ethics approval and consent to participate

All data sources used in this MR study received approval from an ethics standards committee on human experimentation and obtained informed consent from all participants.

Consent for publication

Not applicable.

Availability of data and materials

All the data used in this study can be acquired from the original genome-wide association studies that are mentioned in the text. Any other data generated in the analysis process can be requested from the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by grants from the National Natural Science Foundation of China (91849126), the National Key R&D Program of China (2018YFC1314702), Shanghai Municipal Science and Technology Major Project (No.2018SHZDZX03) and ZHANGJIANG LAB, Tianqiao and Chrissy Chen Institute, and the State Key Laboratory of Neurobiology and Frontiers Center for Brain Science of Ministry of Education, Fudan University.

Authors' contributions

Jin-Tai Yu had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Jin-Tai Yu.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Shu-Yi Huang

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Shu-Yi Huang

Obtained funding: Jin-Tai Yu.

Administrative, technical, or material support: Jin-Tai Yu

Supervision: Jin-Tai Yu.

Acknowledgments

This work was made possible by the generous sharing of GWAS summary statistics. We thank the participants, researchers, and staff associated with the many other studies from which we used data for this report. We thank the UK Biobank for providing summary statistics for these analyses. We also thank the IGAP for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in

analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i-Select chips were funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX DISTALZ grant, Inserm, Institut Pasteur de Lille, Université de Lille 2 and the Lille University Hospital. GERAD was supported by the Medical Research Council (Grant n° 503480), Alzheimer's Research UK (Grant n° 503176), the Wellcome Trust (Grant n° 082604/2/07/Z) and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant n° 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01-AG-12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC-10-196728.

References

1. Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, Cedazo-Minguez A, Dubois B, Edvardsson D, Feldman H, et al. Defeating Alzheimer's disease and other dementias: a priority for European science and society. *The Lancet Neurology*. 2016;15:455-532.
2. Scheltens P, Blennow K, Breteler MMB, de Strooper B, Frisoni GB, Salloway S, Van der Flier WM. Alzheimer's disease. *The Lancet*. 2016;388:505-17.
3. Wilkins JM, Trushina E. Application of Metabolomics in Alzheimer's Disease. *Front Neurol*. 2017;8:719.
4. Cui M, Jiang Y, Zhao Q, Zhu Z, Liang X, Zhang K, Wu W, Dong Q, An Y, Tang H, et al. Metabolomics and incident dementia in older Chinese adults: The Shanghai Aging Study. *Alzheimers Dement*. 2020;16:779-88.
5. Varma VR, Oommen AM, Varma S, Casanova R, An Y, Andrews RM, O'Brien R, Pletnikova O, Troncoso JC, Toledo J, et al. Brain and blood metabolite signatures of pathology and progression in Alzheimer disease: A targeted metabolomics study. *PLoS medicine*. 2018;15:e1002482.
6. Huo Z, Yu L, Yang J, Zhu Y, Bennett DA, Zhao J. Brain and blood metabolome for Alzheimer's dementia: findings from a targeted metabolomics analysis. *Neurobiology of aging*. 2020;86:123-33.
7. Ibáñez C, Cifuentes A, Simó C. Recent advances and applications of metabolomics to investigate neurodegenerative diseases. *Int Rev Neurobiol*. 2015;122:95-132.
8. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*. 2018;362:k601.
9. Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. *Jama*. 2017;318:1925-6.
10. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol*. 2016;45:908-15.
11. Kettunen J, Demirkan A, Würtz P, Draisma HH, Haller T, Rawal R, Vaarhorst A, Kangas AJ, Lyytikäinen LP, Pirinen M, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel

- systemic effects of LPA. *Nat Commun.* 2016;7:11122.
12. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 2018;46:D608-d17.
 13. Hagenbeek FA, Pool R, van Dongen J, Draisma HHM, Jan Hottenga J, Willemsen G, Abdellaoui A, Fedko IO, den Braber A, Visser PJ, et al. Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. *Nat Commun.* 2020;11:39.
 14. Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int J Epidemiol.* 2011;40:740-52.
 15. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol.* 2014;43:922-9.
 16. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, van der Lee SJ, Amlie-Wolf A, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat Genet.* 2019;51:414-30.
 17. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, 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 medicine.* 2015;12:e1001779.
 18. Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, Davies G, Ritchie CW, Gale CR, Starr JM, et al. GWAS on family history of Alzheimer's disease. *Translational psychiatry.* 2018;8:99.
 19. Liu JZ, Erlich Y, Pickrell JK. Case-control association mapping by proxy using family history of disease. *Nat Genet.* 2017;49:325-31.
 20. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Statistics in medicine.* 2017;36:1783-802.
 21. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic epidemiology.* 2016;40:304-14.
 22. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol.* 2017;46:1985-98.
 23. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44:512-25.
 24. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet.* 2018;50:693-8.
 25. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I² statistic. *Int J Epidemiol.* 2016;45:1961-74.

26. Hemani G, Zheng J, Elsworth B, Wade KH, 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.
27. Rasooly D, Patel CJ. Conducting a Reproducible Mendelian Randomization Analysis Using the R Analytic Statistical Environment. *Curr Protoc Hum Genet*. 2019;101:e82.
28. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res*. 2012;40:D109-14.
29. Duprez DA, Jacobs DR, Jr. GlycA, a composite low-grade inflammatory marker, predicts mortality: prime time for utilization? *J Intern Med*. 2019;286:610-2.
30. Gruppen EG, Kunutsor SK, Kieneker LM, van der Vegt B, Connelly MA, de Bock GH, Gansevoort RT, Bakker SJL, Dullaart RPF. GlycA, a novel pro-inflammatory glycoprotein biomarker is associated with mortality: results from the PREVEND study and meta-analysis. *J Intern Med*. 2019;286:596-609.
31. Luo Z, Lei H, Sun Y, Liu X, Su DF. Orosomucoid, an acute response protein with multiple modulating activities. *J Physiol Biochem*. 2015;71:329-40.
32. Jo M, Kim JH, Song GJ, Seo M, Hwang EM, Suk K. Astrocytic Orosomucoid-2 Modulates Microglial Activation and Neuroinflammation. *J Neurosci*. 2017;37:2878-94.
33. van der Lee SJ, Teunissen CE, Pool R, Shipley MJ, Teumer A, Chouraki V, Melo van Lent D, Tynkkynen J, Fischer K, Hernesniemi J, et al. Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies. *Alzheimers Dement*. 2018;14:707-22.
34. Chan L. Apolipoprotein B, the major protein component of triglyceride-rich and low density lipoproteins. *J Biol Chem*. 1992;267:25621-4.
35. Caramelli P, Nitrini R, Maranhão R, Lourenço AC, Damasceno MC, Vinagre C, Caramelli B. Increased apolipoprotein B serum concentration in Alzheimer's disease. *Acta Neurol Scand*. 1999;100:61-3.
36. Kuo YM, Emmerling MR, Bisgaier CL, Essenburg AD, Lampert HC, Drumm D, Roher AE. Elevated low-density lipoprotein in Alzheimer's disease correlates with brain abeta 1-42 levels. *Biochem Biophys Res Commun*. 1998;252:711-5.
37. Raygani AV, Rahimi Z, Kharazi H, Tavilani H, Pourmotabbed T. Association between apolipoprotein E polymorphism and serum lipid and apolipoprotein levels with Alzheimer's disease. *Neurosci Lett*. 2006;408:68-72.
38. Song F, Poljak A, Crawford J, Kochan NA, Wen W, Cameron B, Lux O, Brodaty H, Mather K, Smythe GA, et al. Plasma apolipoprotein levels are associated with cognitive status and decline in a community cohort of older individuals. *PLoS One*. 2012;7:e34078.
39. Wingo TS, Cutler DJ, Wingo AP, Le NA, Rabinovici GD, Miller BL, Lah JJ, Levey AI. Association of Early-Onset Alzheimer Disease With Elevated Low-Density Lipoprotein Cholesterol Levels and Rare Genetic Coding Variants of APOB. *JAMA Neurol*. 2019;76:809-17.
40. Tynkkynen J, Hernesniemi JA, Laatikainen T, Havulinna AS, Sundvall J, Leiviskä J, Salo P, Salomaa V. Apolipoproteins and HDL cholesterol do not associate with the risk of future dementia and Alzheimer's disease: the National Finnish population study (FINRISK). *Age (Dordr)*. 2016;38:465-73.

41. Takechi R, Galloway S, Pallegage-Gamarallage M, Wellington C, Johnsen R, Mamo JC. Three-dimensional colocalization analysis of plasma-derived apolipoprotein B with amyloid plaques in APP/PS1 transgenic mice. *Histochem Cell Biol.* 2009;131:661-6.
42. Takechi R, Galloway S, Pallegage-Gamarallage MM, Wellington CL, Johnsen RD, Dhaliwal SS, Mamo JC. Differential effects of dietary fatty acids on the cerebral distribution of plasma-derived apo B lipoproteins with amyloid-beta. *Br J Nutr.* 2010;103:652-62.
43. Löffler T, Flunkert S, Havas D, Sántha M, Hutter-Paier B, Steyrer E, Windisch M. Impact of ApoB-100 expression on cognition and brain pathology in wild-type and hAPPsl mice. *Neurobiology of aging.* 2013;34:2379-88.
44. Ishii M. Apolipoprotein B as a New Link Between Cholesterol and Alzheimer Disease. *JAMA Neurol.* 2019;76:751-3.
45. Liu Y, Zhong X, Shen J, Jiao L, Tong J, Zhao W, Du K, Gong S, Liu M, Wei M. Elevated serum TC and LDL-C levels in Alzheimer's disease and mild cognitive impairment: A meta-analysis study. *Brain Res.* 2020;1727:146554.
46. Wu Y, Wang Z, Jia X, Zhang H, Zhang H, Li J, Zhang K. Prediction of Alzheimer's disease with serum lipid levels in Asian individuals: a meta-analysis. *Biomarkers.* 2019;24:341-51.
47. Wang P, Zhang H, Wang Y, Zhang M, Zhou Y. Plasma cholesterol in Alzheimer's disease and frontotemporal dementia. *Transl Neurosci.* 2020;11:116-23.
48. Zhou Z, Liang Y, Zhang X, Xu J, Lin J, Zhang R, Kang K, Liu C, Zhao C, Zhao M. Low-Density Lipoprotein Cholesterol and Alzheimer's Disease: A Systematic Review and Meta-Analysis. *Frontiers in aging neuroscience.* 2020;12:5.
49. Østergaard SD, Mukherjee S, Sharp SJ, Proitsi P, Lotta LA, Day F, Perry JR, Boehme KL, Walter S, Kauwe JS, et al. Associations between Potentially Modifiable Risk Factors and Alzheimer Disease: A Mendelian Randomization Study. *PLoS medicine.* 2015;12:e1001841; discussion e.
50. Benn M, Nordestgaard BG, Frikke-Schmidt R, Tybjaerg-Hansen A. Low LDL cholesterol, PCSK9 and HMGCR genetic variation, and risk of Alzheimer's disease and Parkinson's disease: Mendelian randomisation study. *Bmj.* 2017;357:j1648.
51. Lee S, Parekh T, King SM, Reed B, Chui HC, Krauss RM, Yassine HN. Low-Density Lipoprotein Particle Size Subfractions and Cerebral Amyloidosis. *Journal of Alzheimer's disease : JAD.* 2019;68:983-90.
52. Sáiz-Vazquez O, Puente-Martínez A, Ubillos-Landa S, Pacheco-Bonrostro J, Santabárbara J. Cholesterol and Alzheimer's Disease Risk: A Meta-Meta-Analysis. *Brain sciences.* 2020;10.
53. Xu W, Tan L, Wang HF, Jiang T, Tan MS, Tan L, Zhao QF, Li JQ, Wang J, Yu JT. Meta-analysis of modifiable risk factors for Alzheimer's disease. *Journal of neurology, neurosurgery, and psychiatry.* 2015;86:1299-306.
54. Solomon A, Kåreholt I, Ngandu T, Winblad B, Nissinen A, Tuomilehto J, Soininen H, Kivipelto M. Serum cholesterol changes after midlife and late-life cognition: twenty-one-year follow-up study. *Neurology.* 2007;68:751-6.

55. Kivipelto M, Helkala EL, Laakso MP, Hänninen T, Hallikainen M, Alhainen K, Iivonen S, Mannermaa A, Tuomilehto J, Nissinen A, et al. Apolipoprotein E epsilon4 allele, elevated midlife total cholesterol level, and high midlife systolic blood pressure are independent risk factors for late-life Alzheimer disease. *Annals of internal medicine*. 2002;137:149-55.
56. Tan ZS, Seshadri S, Beiser A, Wilson PW, Kiel DP, Tocco M, D'Agostino RB, Wolf PA. Plasma total cholesterol level as a risk factor for Alzheimer disease: the Framingham Study. *Arch Intern Med*. 2003;163:1053-7.
57. Chung HS, Lee JS, Kim JA, Roh E, Lee YB, Hong SH, Kim NH, Yoo HJ, Seo JA, Kim SG, et al. Variability in Total Cholesterol Concentration Is Associated With the Risk of Dementia: A Nationwide Population-Based Cohort Study. *Front Neurol*. 2019;10:441.
58. Pedrini S, Chatterjee P, Hone E, Martins RN. High-density lipoprotein-related cholesterol metabolism in Alzheimer's disease. *J Neurochem*. 2020.
59. Huang D, Liu D, Yin J, Qian T, Shrestha S, Ni H. Glutamate-glutamine and GABA in brain of normal aged and patients with cognitive impairment. *Eur Radiol*. 2017;27:2698-705.
60. Takado Y, Sato N, Kanbe Y, Tomiyasu M, Xin L, Near J, Yoshikawa K, Sahara N, Higashi T, Suhara T, et al. Association between Brain and Plasma Glutamine Levels in Healthy Young Subjects Investigated by MRS and LC/MS. *Nutrients*. 2019;11.
61. Uddin LQ, Kelly AM, Biswal BB, Castellanos FX, Milham MP. Functional connectivity of default mode network components: correlation, anticorrelation, and causality. *Hum Brain Mapp*. 2009;30:625-37.
62. Adams CD. Circulating Glutamine and Alzheimer's Disease: A Mendelian Randomization Study. *Clin Interv Aging*. 2020;15:185-93.
63. Andersen JV, Christensen SK, Aldana BI, Nissen JD, Tanila H, Waagepetersen HS. Alterations in Cerebral Cortical Glucose and Glutamine Metabolism Precedes Amyloid Plaques in the APP^{swe}/PSEN1^{dE9} Mouse Model of Alzheimer's Disease. *Neurochem Res*. 2017;42:1589-98.
64. Wang Y, Wang Q, Li J, Lu G, Liu Z. Glutamine Improves Oxidative Stress through the Wnt3a/ β -Catenin Signaling Pathway in Alzheimer's Disease In Vitro and In Vivo. *Biomed Res Int*. 2019;2019:4690280.
65. Westergaard N, Waagepetersen HS, Belhage B, Schousboe A. Citrate, a Ubiquitous Key Metabolite with Regulatory Function in the CNS. *Neurochem Res*. 2017;42:1583-8.

Tables

Table 1.

Pleiotropy and heterogeneity analyses for the association between metabolites and Alzheimer's disease in primary analysis

Metabolites	MR-Egger intercept P-value	Q test P-value	MR-PRESSO global test P-value	Corrected MR-PRESSO	
				OR (95%CI)	P-value
Protein					
ApoB	0.80	<0.001	<0.001	2.20 (1.66, 2.92)	2.74E-04
Glycoprotein acetyls	0.83	0.12	0.20	-	-
Carboxylic acids and derivatives					
Citrate	0.48	0.93	0.96	-	-
Glutamine	0.92	1.00	0.88	-	-
Lipoprotein					
IDL-C	0.53	<0.001	<0.001	1.72 (1.48, 1.99)	1.65E-07
IDL-FC	0.27	<0.001	<0.001	1.08 (0.99, 1.17)	0.074
IDL-L	0.26	<0.001	<0.001	1.24 (1.09, 1.40)	0.002
IDL-P	0.33	<0.001	<0.001	1.35 (1.16, 1.57)	7.22E-04
IDL-PL	0.83	<0.001	<0.001	1.69 (1.43, 1.98)	1.19E-06
LDL-C	0.44	<0.001	<0.001	1.59 (1.40, 1.81)	2.83E-08
L-LDL-C	0.41	<0.001	<0.001	1.67 (1.46, 1.91)	6.58E-08
L-LDL-CE	0.73	<0.001	<0.001	1.69 (1.49, 1.91)	8.22E-09
L-LDL-FC	0.41	<0.001	<0.001	1.69 (1.47, 1.94)	9.34E-08
L-LDL-L	0.51	<0.001	<0.001	1.61 (1.40, 1.85)	1.77E-07
L-LDL-P	0.34	<0.001	<0.001	1.68 (1.47, 1.92)	2.38E-08

L-LDL-PL	0.34	<0.001	<0.001	1.64 (1.45, 1.84)	3.45E-09
M-LDL-C	0.42	<0.001	<0.001	1.81 (1.57, 2.09)	4.10E-08
M-LDL-CE	0.40	<0.001	<0.001	1.74 (1.50, 2.01)	8.98E-08
M-LDL-L	0.42	<0.001	<0.001	1.89 (1.62, 2.21)	1.45E-07
M-LDL-P	0.17	<0.001	<0.001	1.78 (1.52, 2.09)	3.32E-07
M-LDL-PL	0.33	<0.001	<0.001	1.91 (1.59, 2.30)	3.53E-06
S-LDL-C	0.82	<0.001	<0.001	1.96 (1.61, 2.39)	4.07E-06
S-LDL-L	0.44	<0.001	<0.001	2.09 (1.62, 2.69)	5.46E-05
S-LDL-P	0.14	<0.001	<0.001	2.09 (1.62, 2.70)	4.38E-05
S-HDL-P	0.19	0.45	<0.001	1.93 (1.60, 2.33)	2.62E-06
XL-HDL-C	0.81	0.11	0.11	-	-
XL-HDL-CE	0.47	0.12	0.13	-	-
XL-HDL-FC	0.72	0.09	<0.001	1.64 (1.45, 1.84)	3.45E-09
XL-HDL-P	0.96	<0.001	<0.001	0.91 (0.86, 0.98)	0.012
XS-VLDL-PL	0.13	<0.001	<0.001	1.03 (0.97, 1.09)	0.396
Steroids and steroid derivatives					
Serum-C	0.78	<0.001	0.60	-	-
Esterified cholesterol	0.48	<0.001	<0.001	2.18 (1.54, 3.08)	0.004

CI, confidence interval; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier; OR, odds ratio.

Table 2.

Significant metabolic pathways involved in the pathogenesis of Alzheimer's disease.

Metabolic pathway	Metabolites involved	P value	FDR	Database
D-Glutamine and D-glutamate metabolism	L-Glutamine; L-Glutamine	1.30E-05	0.001	KEGG
Alanine, aspartate and glutamate metabolism	L-Glutamine; Citric acid	0.0003	0.012	KEGG
Glyoxylate and dicarboxylate metabolism	L-Glutamine; Citric acid	0.0004	0.012	KEGG
Nitrogen metabolism	L-Glutamine	0.0079	0.165	KEGG
Arginine biosynthesis	L-Glutamine	0.0183	0.308	KEGG
Citrate cycle (TCA cycle)	Citric acid	0.0261	0.366	KEGG

FDR, false discovery rates; TCA, tricarboxylic acid.

Figures

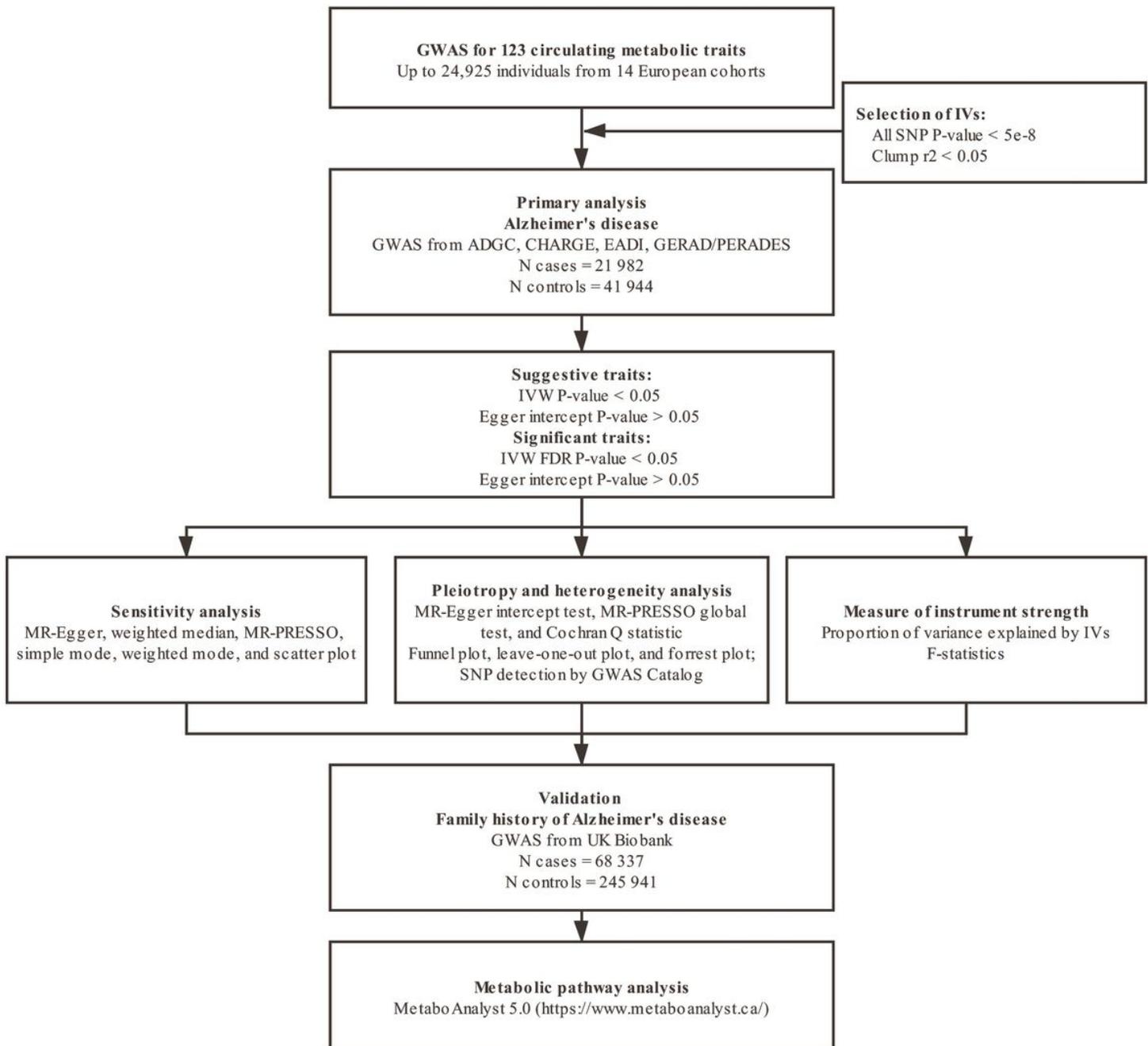


Figure 1

A flow diagram for the process for identifying genetic variants to be included in the Mendelian randomization analysis ADGC, Alzheimer Disease Genetics Consortium; CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium; EADI, The European Alzheimer's Disease Initiative; GERAD/PERADES, Genetic and Environmental Risk in Alzheimer's Disease /Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium; GWAS, genome-wide association studies; N, number; SNP, single nucleotide polymorphism; IV, instrumental variable; IVW, inverse variance weighting; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier.

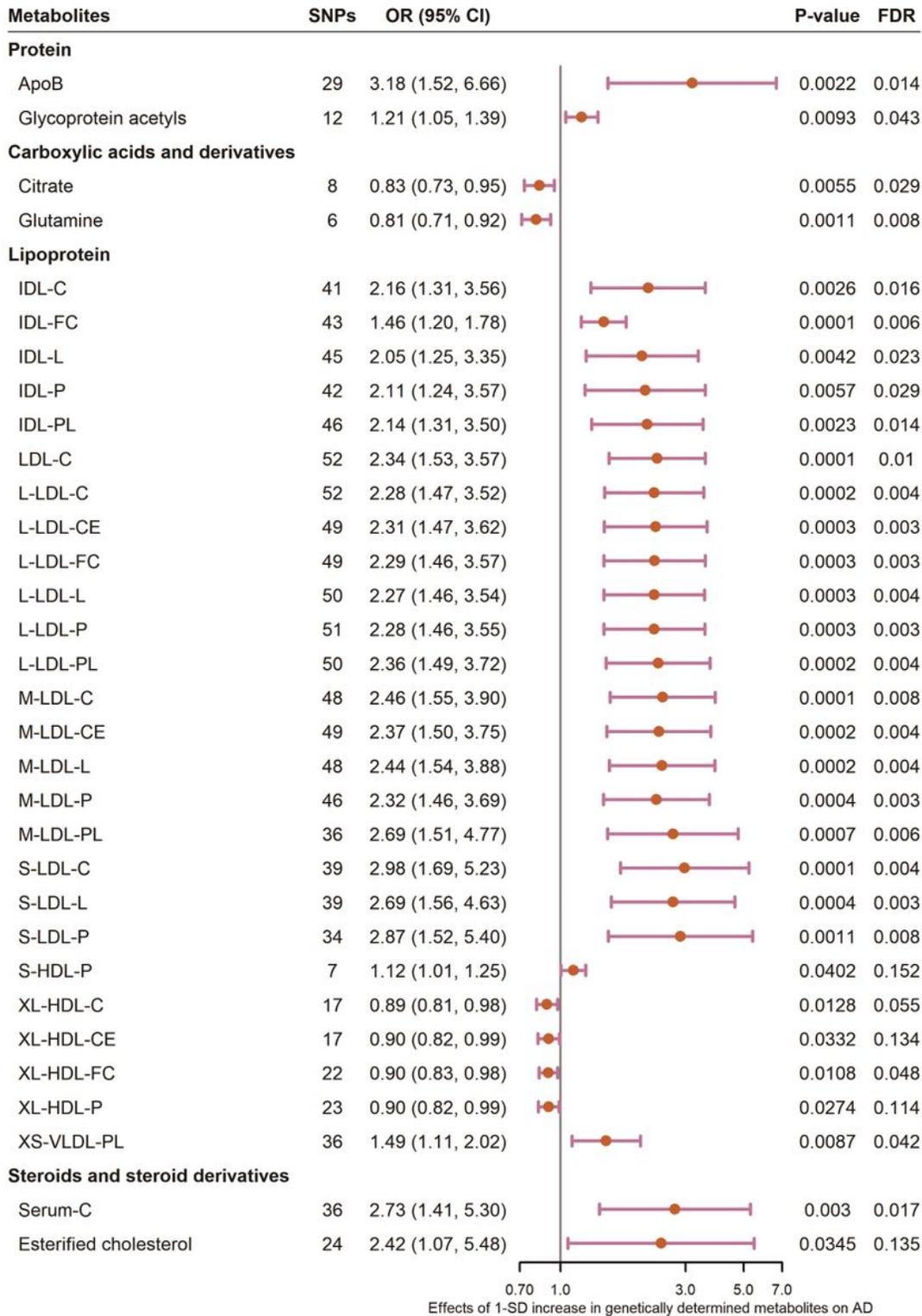


Figure 2

Mendelian randomization results for the circulating metabolites and the risks they pose in risk of Alzheimer's disease in primary analysis AD, Alzheimer's disease; CI, confidence interval; FDR, false discovery rates; SD, standard deviation; SNPs, single nucleotide polymorphisms; OR, odds ratio.

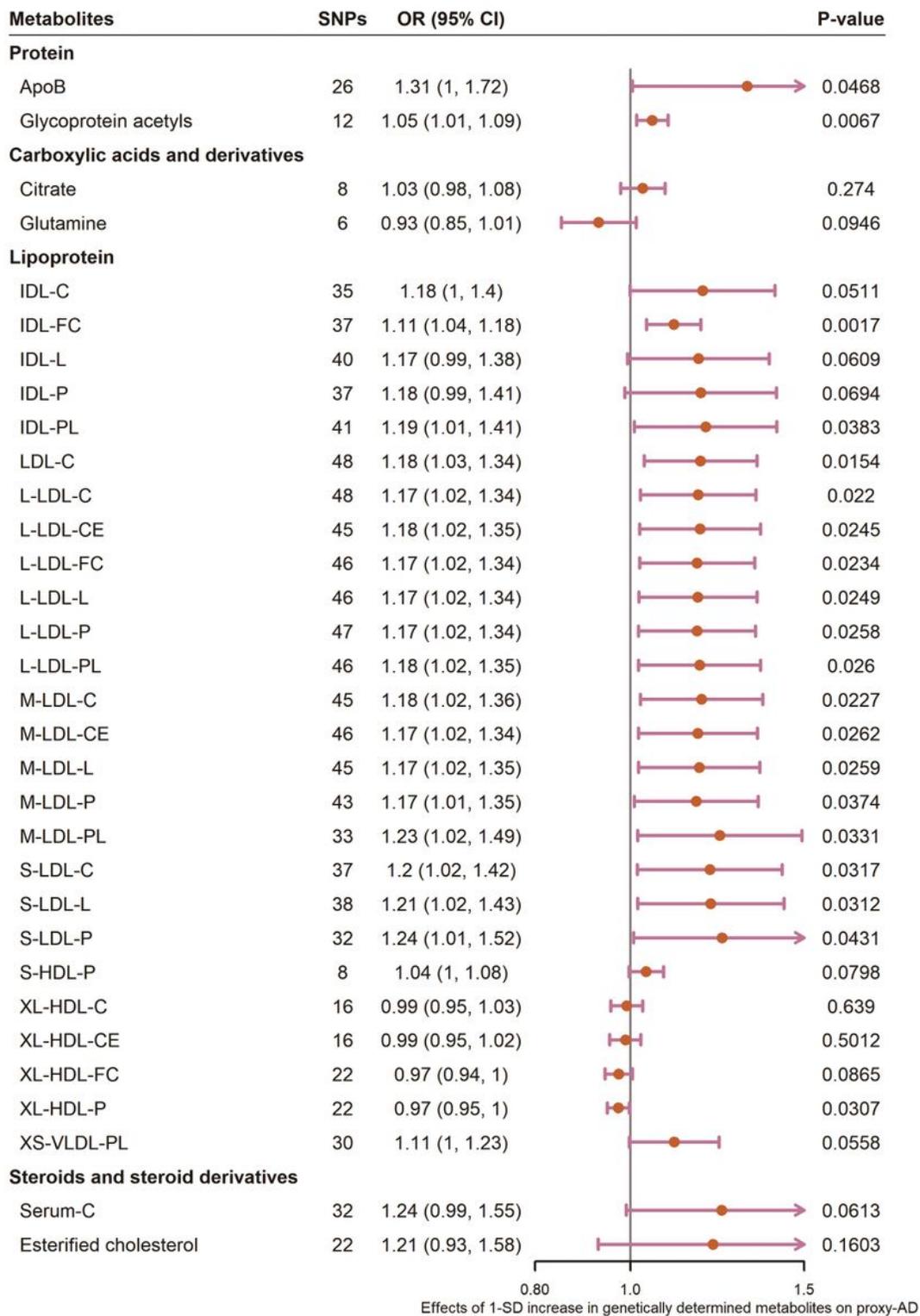
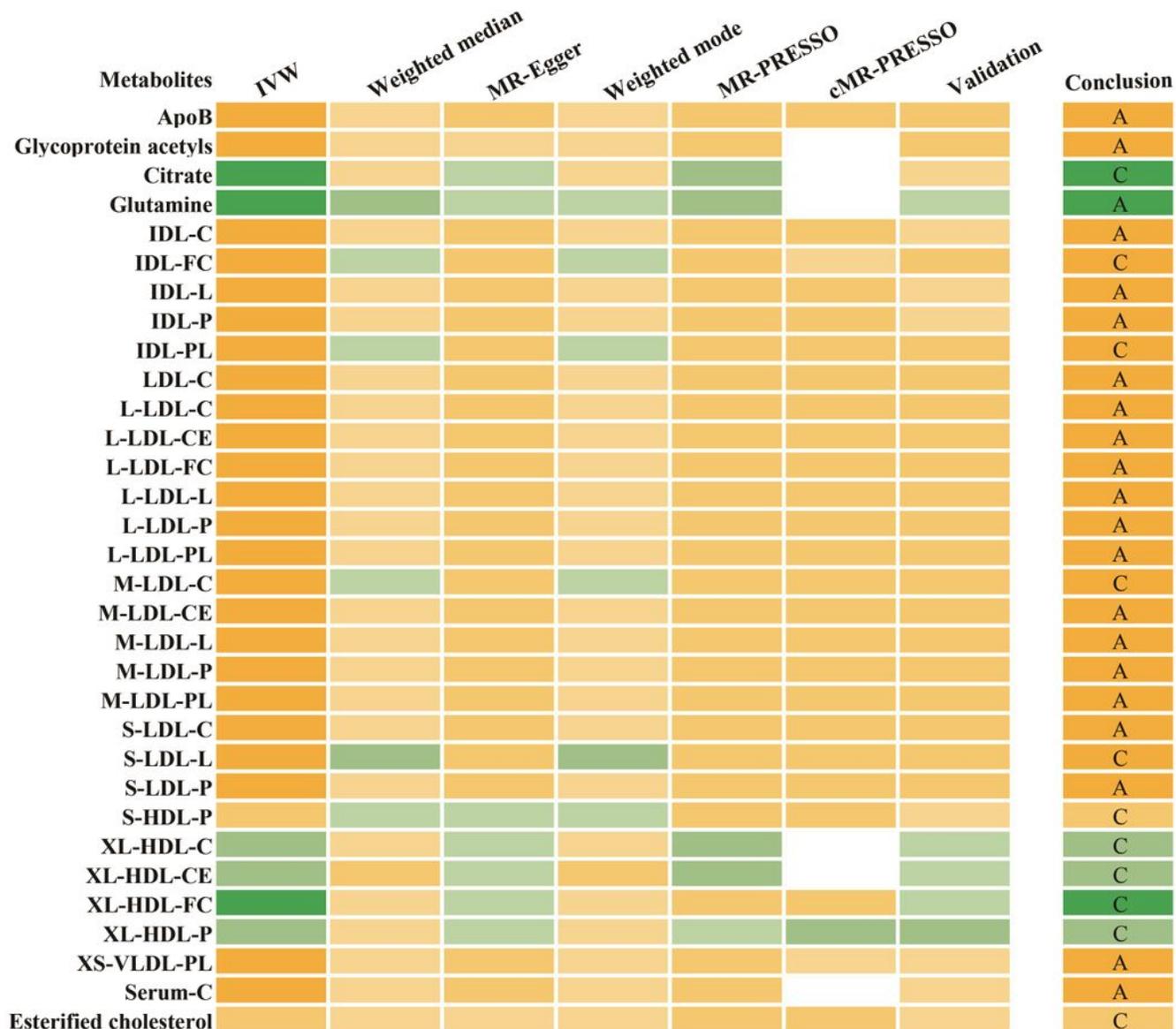


Figure 3

Mendelian randomization results for associations between circulating metabolites and family history of Alzheimer's disease in validation AD, Alzheimer's disease; CI, confidence interval; SD, standard deviation; SNPs, single nucleotide polymorphisms; OR, odds ratio.



Light Green	Protective effect on AD risk
Medium Green	Protective effect on AD risk with P<0.05
Dark Green	Protective effect on AD risk with FDR P<0.05 in IVW
Light Orange	Adverse effect on AD risk
Medium Orange	Adverse effect on AD risk with P<0.05
Dark Orange	Adverse effect on AD risk with FDR P<0.05 in IVW
A	Consistent evidence across main and sensitivity methods
B	Suggestive evidence, consistent across methods
C	Conflicting evidence across methods

Figure 4

Overview of the design and results of this MR study on circulating metabolites and Alzheimer's disease. All results can be found in Fig. 2, Fig. 3, Table 1, and Additional file 2. AD, Alzheimer's disease; cMR-PRESSO, corrected Mendelian randomization pleiotropy residual sum and outlier after removing of significant outliers; IVW, inverse variance-weighted; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier; SD, standard deviation; SNPs, single nucleotide polymorphisms.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.docx](#)
- [Additionalfile2.docx](#)
- [Additionalfile3.docx](#)
- [Additionalfile4.docx](#)
- [Additionalfile5.docx](#)
- [Additionalfile6.docx](#)