Jie *et al. Journal of Translational Medicine* (2024) 22:1077
https://doi.org/10.1186/s12967-024-05796-2

Journal of Translational Medicine

RESEARCH Open Access

The role of cerebrospinal fluid metabolites in mediating the impact of lipids on Late-Onset Alzheimer's Disease: a two-step mendelian randomization analysis

Jie Jie[1], Yonglu Gong[1], Hongbo Hu[1] and Su Liu[1*]

Abstract

Background Although research has indicated correlations between lipids, cerebrospinal fluid (CSF) metabolites, and Late-Onset Alzheimer's Disease (LOAD), the specific causal relationships among these elements, as well as the roles and mechanisms of the cerebrospinal fluid metabolites, remain unclear.

Methods Statistical datasets derived from Genome-Wide Association Studies (GWAS) were utilized to assess the bidirectional causal relationships between lipids and LOAD. Subsequently, genetic variants associated with CSF metabolites and established lipids underwent a two-step Mendelian randomization (MR) analysis to explore potential mediators and analyze mediation effects. Sensitivity analyses were employed to assess the robustness of the detection systems.

Results Genetically predicted cholesterol (IVW OR = 0.989; 95% CI 0.982–0.996) was found to reduce the risk of LOAD, whereas Phosphatidylcholine (PC) (18:1_0:0) (IVW OR = 1.015; 95% CI 1.005–1.025) posed a risk factor. The potential mediator, CSF metabolite N-acetylneuraminate (NeuAC), was identified with a mediation proportion of 21.02% (3.25%, 45.50%). No pleiotropy or heterogeneity was detected across MR analyses.

Conclusions The findings underscore the pivotal role of CSF metabolomics in elucidating the lipid-mediated pathogenesis of LOAD, highlighting potential diagnostic and preventative biomarkers.

Keywords Lipids, Cerebrospinal fluid metabolites, Late-Onset Alzheimer's Disease, Mediator, Mendelian randomization

*Correspondence:
Su Liu
jary618@outlook.com
[1]Changde Hospital, Xiangya School of Medicine, Central South University (The First People's Hospital of Changde City), 818 Renmin Road, Changde City, Hunan Province 415000, China

Background

Alzheimer's Disease (AD) is a neurodegenerative condition characterized predominantly by progressive memory loss and constitutes the most prevalent form of dementia, affecting over 50 million people globally. This increasing prevalence negatively impacts families, communities, and healthcare systems worldwide [1–3]. Among individuals aged over 65, the incidence of AD ranges from 1–3% [4], with over 90% of these cases classified as Late-Onset Alzheimer's Disease (LOAD) [5]. Previous studies have demonstrated a close association between lipid profiles and LOAD. The APOE ε4 allele, a major genetic risk factor for AD, encodes a lipid transport protein involved in cholesterol metabolism [6]. Genome-Wide Association Studies (GWAS) have identified Single nucleotide polymorphisms (SNPs) related to lipid processes in LOAD, including genes like CLU, ABCA1, and ABCA7 [7–9]. However, despite extensive studies, confounding factors have obscured definitive evidence of a causal relationship between lipids and LOAD, and the underlying mechanisms remain unclear [9].

Intriguingly, accumulating evidence suggests that AD's pathophysiology is closely linked to disturbances in brain energy metabolism and homeostasis [10–13]. In neuropsychiatric disorders, CSF metabolites are considered superior analytes since they are in direct contact with brain and spinal cord cells, thus reflecting physiological changes in the central nervous system [14, 15]. Previous research has indicated that lipids participate in the metabolism of substances within CSF [16]. Disorders in lipid metabolism, such as those involving cholesterol and APOE, may constitute a principal mechanism underlying AD [17–19].

Mendelian randomization (MR) offers a robust approach to overcoming the limitations inherent in traditional epidemiological and observational studies [20]. By employing genetic variants (SNPs) from non-experimental data as instrumental variables (IVs), MR infers the effects of exposures on outcomes [21]. Alleles are randomly assigned to offspring during meiosis, akin to a randomized control trial [22]. This process, generally unaffected by external factors, maximizes the minimization of confounding and reverse causation [23]. Two-sample MR utilizes separate cohorts to measure exposures and outcomes, enabling more extensive and substantiated studies.

Given the established correlations between lipids, CSF metabolites, and AD, this study leverages the latest aggregated data from GWAS. It aims to infer potential causal relationships between lipids, CSF metabolites, and LOAD using a two-sample MR approach. Furthermore, a two-step MR analysis explores the mediating role of CSF metabolites in the association between lipids and LOAD, identifying potential biomarkers for early diagnosis and clinical intervention.

Methods

Study design

Figure 1 illustrates the research design. MR is an analysis technique using genetic IVs, specifically SNPs, to assess the impact of exposures on various outcomes and to analyze causal relationships between them. The genetic variants selected as IVs in MR analysis must satisfy three key assumptions: relevance, independence, and exclusion restriction. The two-step MR analysis must meet the following criteria: (1) a causal relationship between the exposure and the outcome; (2) a causal relationship between the mediator and the outcome, independent of the exposure; (3) a causal relationship between the exposure and the mediator. Reverse MR analysis is conducted at each step to confirm the absence of reverse causality in the final positive results [24]. Our study adheres to the STROBE-MR guidelines [25] as detailed in Supplementary Table S1.

Data sources

The datasets utilized in this study are derived from publicly available GWAS summary data (https://www.ebi.ac.uk/gwas/). therefore, no additional ethical approval was required. Table 1 summarizes the GWAS data employed in this study. Lipids data were obtained from a study by Linda Ottensmann et al., which conducted shotgun lipidomic analysis via mass spectrometry on 7,174 individuals from the GeneRISK cohort, identifying 179 subspecies-level lipid molecules spanning 13 lipid classes and four major lipid categories: triglycerides, glycerophospholipids, sphingolipids, and sterols [26]. Data on 338 CSF metabolites were sourced from a metabolome-wide association study by Daniel J. Panyard et al., involving 291 individuals of European descent [15]. LOAD data were derived from a genome-wide association study by Douglas P. Wightman et al., involving 1,126,563 individuals (90,338 cases and 1,036,225 controls) of European ancestry [27]. Replication data for blood metabolites were obtained from a study by Yiheng Chen et al., which included 8,299 individuals from a Canadian aging longitudinal cohort [28]. All GWAS data originated from distinct consortia, minimizing the risk of bias due to sample overlap.

Instrumental variable selection

Instrumental variables were selected following rigorous data cleaning principles to ensure the validity of our MR analysis. First, SNPs were required to be significantly associated with the exposure, so we set the significance threshold at $P<5\times10^{-8}$. However, because the GWAS data on lipids and CSF metabolites did not yield

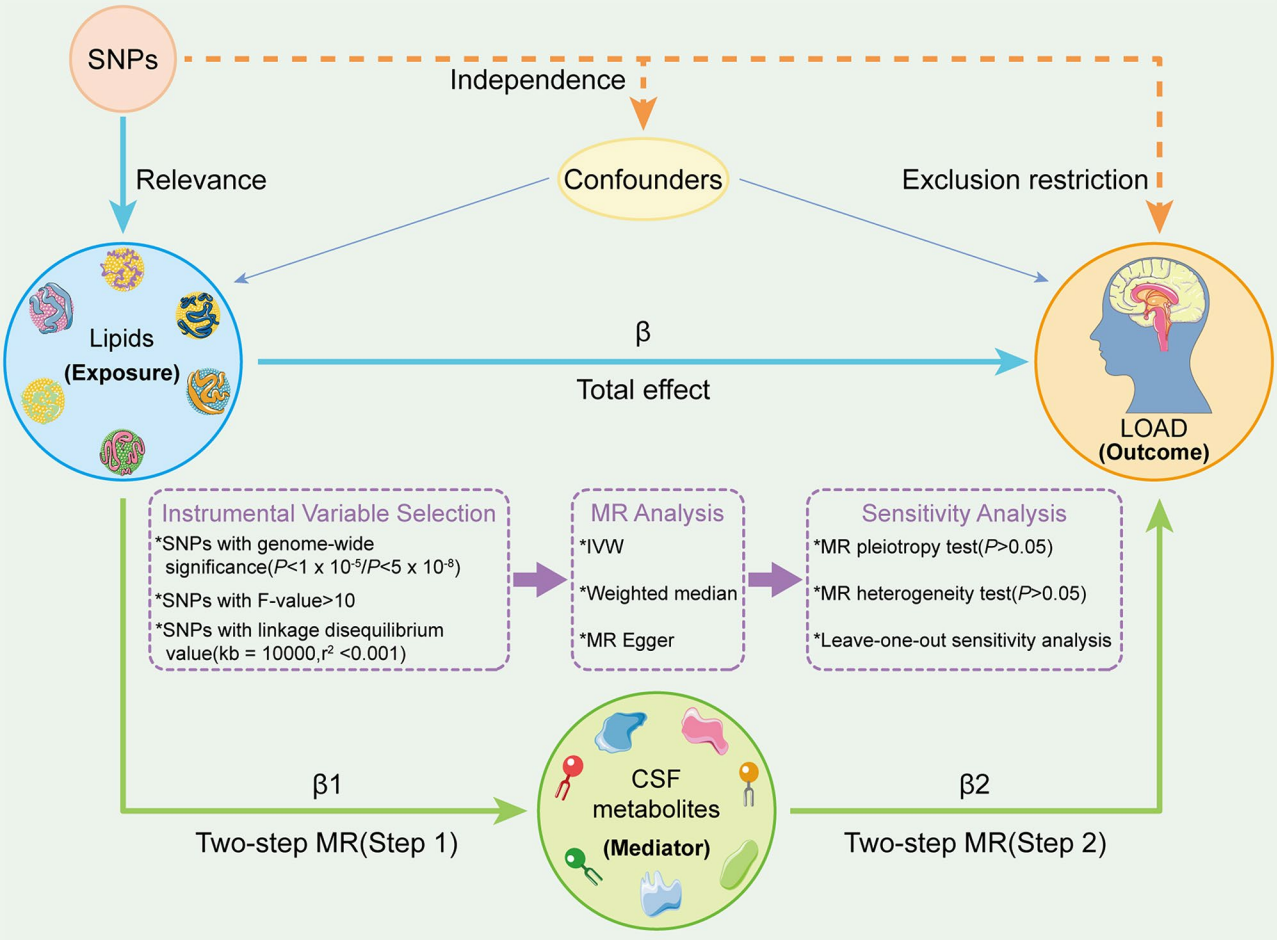

Fig. 1 Overview and research content of Mendelian Randomization. MR three assumptions: (1) Instrumental variables(IVs) are strongly associated with the exposure (Relevance), (2) IVs are not affected by confounding factors (Independence), (3) IVs influence the outcome solely through their effect on the exposure (Exclusion Restriction) SNP. Single glycerine polymorphism; LOAD, Late onset Alzheimer's disease; CSF, Cerebrospinal fluid; IVW, Inverse variance weighted

Table 1 Overview of GWAS data included in Mendelian randomization analyses

Phenotype	Sample Size	Ancestry	Author	Year of publication	PubMed ID
Lipids	7174	Finland	Ottensmann et al.	2023	37,907,536
CSF metabolites	291	European	Panyard et al.	2021	33,437,055
LOAD	1,126,563(90338 cases and 1036225 controls)	Finland, Iceland, Norway, Spain, Sweden, U.K., NR, U.S	Wightman et al.	2021	34,493,870
Blood metabolites(Copy)	8299	European	Chen et al.	2023	36,635,386

CSF, Cerebrospinal fluid; LOAD, Late-Onset Alzheimer's Disease

enough instrumental variables for subsequent analysis, we adjusted the threshold to $P<1\times10^{-5}$ to obtain sufficient instruments. Second, since linkage disequilibrium (LD) is more likely to occur between closely located genetic variants, potentially introducing bias [29], we performed a clumping process (window size=10,000 kb, r^2 = 0.001) to remove SNPs in LD. An important step in MR is to ensure that the effect of SNPs on the exposure corresponds to the same allele as their effect on the outcome. To avoid distortion due to strand orientation or allele coding errors, we removed palindromic SNPs (e.g., those with A/T or G/C alleles). During harmonization, we aligned the alleles with the human genome reference sequence (GRCh37) and eliminated ambiguous and duplicate SNPs [30]. Finally, we calculated the F-statistic for each SNP using the formula $F=\frac{R^2}{(1-R^2)}\times\frac{(n-k-1)}{k}$, where the proportion of phenotypic variance explained (R^2) is calculated as $R^2=2\times MAF\times(1\text{-}MAF)\times\beta^2$. Here, n

represents the sample size, k is the number of instrumental variables, β is the effect size, and MAF is the minor allele frequency. Weak instrumental variables can produce misleading results because the estimated effect of the exposure-outcome association may be biased [23]; therefore, SNPs with an F-statistic less than 10 were excluded.

Causal total effect analysis for genetic prediction

We initiated a two-sample Mendelian Randomization analysis to explore the causal relationship between lipids profiles and LOAD, with β representing the total effect. The primary method employed was the inverse-variance weighted (IVW) fixed-effect model for estimation, shifting to a random-effects model in the presence of heterogeneity among the genetic instruments [31]. Notably, the IVW method operates under the assumption that all genetic variants are valid instrumental variables, providing estimates through the slope of a weighted linear regression, known for its robust causal inference capabilities [32]. To enhance the credibility of the IVW results, both the weighted median approach and MR-Egger regression were utilized to assess the reliability and stability of the outcomes. The weighted median method provides consistent estimates even when 50% of the instruments are deemed invalid [33]. MR-Egger can detect violations of the IV assumptions and offers effect estimates that are not influenced by these violations [34].

Mediation effect analysis of the association "lipids-CSF metabolites-LOAD"

The two-step MR test was employed to delineate the direct and indirect impacts of lipids and CSF metabolites on LOAD. In the first step, we assessed the effect (β1) of lipids on the potential mediators, and in the second step, we evaluated the effect (β2) of these mediators on LOAD [35, 36]. SNPs used in the second step of the analysis were distinct and did not overlap with those used in the first step.

Reverse MR analysis

In this study, reverse MR analysis will be conducted using candidate lipids that have a causal relationship with LOAD and CSF metabolites. The purpose of this approach is to eliminate potential confounding from reverse causality, thereby ensuring the reliability of the research findings.

Replicate analysis

Building on the potential CSF metabolites identified earlier, we will conduct a mediation effect analysis of "lipid-blood metabolites-LOAD" using the corresponding metabolites. This analysis aims to explore whether the potential mediators in the CSF can exert similar mediating roles in the blood, thereby elucidating the underlying mechanisms of these mediators within the pathway.

Sensitivity analysis

To ensure the robustness and reliability of our findings, we conducted multiple sensitivity analyses aimed at detecting and correcting potential pleiotropy, heterogeneity, and the influence of individual SNPs. First, when IVs affect other exposures related to the outcome of interest, pleiotropy can lead to confounding and bias. Therefore, we employed the MR-Egger regression method to detect horizontal pleiotropy. MR-Egger regression assesses the average pleiotropic effect across all IVs, and the significance of its intercept term (i.e., the MR-Egger intercept) is used to determine the presence of horizontal pleiotropy [37]. If the P-value of the MR-Egger intercept is less than 0.05, it indicates evidence of horizontal pleiotropy. In such cases, we further applied the MR-PRESSO method to detect and correct bias caused by pleiotropy [38]. MR-PRESSO can identify and remove outlier SNPs, after which we re-conducted the MR analysis and sensitivity analyses to obtain more accurate and reliable estimates. Second, to assess whether there was heterogeneity among the IVs, we used Cochran's Q test [39]. Heterogeneity refers to significant differences in the effects of different IVs on the outcome, which may affect the accuracy of MR analysis. Cochran's Q test detects inconsistency in the effects of the IVs, and if the P-value corresponding to the Q statistic is less than 0.05, it indicates significant heterogeneity. Additionally, to ensure that our results were not driven by extreme effects of individual SNPs, we performed a leave-one-out analysis [23]. In this analysis, we sequentially excluded each IV and recalculated the overall effect estimate. If the results changed significantly after excluding any SNP, it suggests that the SNP may have an undue influence on the findings and warrants further careful consideration.

Statistical analysis

MR analysis and causal effect estimation were conducted using R (version 4.3.3) and two principal packages: "TwoSampleMR" (version 0.5.11) and "MendelianRandomization" (version 0.9.0). Continuous outcomes were reported with beta (β) values and 95% confidence intervals (CIs), while binary outcomes were estimated using odds ratios (ORs) and 95% CIs, reflecting the risk alteration in outcomes with each standard deviation (SD) increase in genetically predicted exposure. Stepwise regression was employed to select exposures and mediators with genuine effects [40]. The product of coefficients method (β1×β2) was utilized for estimating indirect effects. The mediation proportion was calculated using

the formula (β1×β2)/β, with standard errors and 95% CIs computed using the delta method.

Metabolic pathways involved in CSF metabolites

Extended analysis of CSF metabolites causally linked with LOAD (P_{IVW}<0.05) was conducted using the online network tool, MetaboAnalyst 6.0 (https://www.metaboanalyst.ca/MetaboAnalyst/). MetaboAnalyst 6.0 is a publicly available online tool that performs pathway analysis of metabolomics data based on the Kyoto Encyclopedia of Genes and Genomes (KEGG).

Results

Genetic instruments for exposures

For the 179 lipids, 4,512 instrumental variables were selected with a median F-statistic of 22.46 (ranging from 18.80 to 1969.07). From the GWAS data of 338 CSF metabolites, 21,147 SNPs were selected with a median F-statistic of 22.27 (ranging from 19.29 to 904.75). For LOAD, used as the exposure, 38 SNPs were extracted with a median F-statistic of 50.23 (ranging from 30.23 to 1569.32). For the replicated blood metabolites, 33 instrumental variables were selected with a median F-statistic of 19.33 (ranging from 17.13 to 69.86). All SNPs had an F-statistic greater than 10, which excludes the interference from weak instrumental variables (Supplementary Tables S2-5).

Genetic causal relationship between lipids and LOAD

An initial two-sample MR analysis using the IVW method identified ten lipid species associated with LOAD, involving six categories: cholesterol, triglycerides (TG), phosphatidylcholine (PC), phosphatidylcholine ether (PCO), phosphatidylethanolamine (PE), and phosphatidylinositol (PI) (Supplementary Table S6). After Bonferroni correction for multiple testing ($P<0.05/10=5.00\times10^{-3}$), as shown in Fig. 2, only Cholesterol (IVW OR=0.989; 95% CI 0.982–0.996; $P=3.14\times10^{-3}$) and PC (18:1_0:0) (IVW OR=1.015; 95% CI 1.005–1.025; $P=2.82\times10^{-3}$) remained significant (Supplementary Table S6). Elevated levels of Cholesterol are associated with reduced risk of LOAD, whereas PC (18:1_0:0) demonstrates the opposite effect. Additionally, sensitivity analyses indicated no evidence of pleiotropy or heterogeneity in these positive findings (Supplementary Table S7). In assessing the causal impact of LOAD on these lipids, we observed inverse causality for PC(O-16:0_22:5) (IVW OR=0.602; 95% CI 0.439–0.824; $P=1.55\times10^{-3}$) and PC(O-16:1_20:3) (IVW OR=0.544; 95% CI 0.406–0.731; $P=5.18\times10^{-3}$) (Supplementary Table S8).

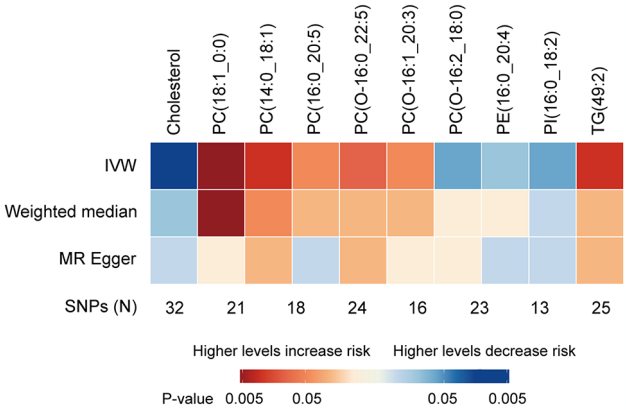

Fig. 2 Mendelian randomization analyses show causal effects between lipids on LOAD. PC, Phosphatidylcholine; PE, Phosphatidylethanolamine; PI, Phosphatidylinositol; TG, Triacylglycerol. The lipid species are named in the following notation: class name < sum of carbon atoms>:< sum of double bonds>;< sum of hydroxyl groups>. Acyl chains, "_"

Mediated analysis of potential CSF metabolites

Following the analysis, we identified two positive findings regarding the causal relationship between lipids and LOAD, which led us to further explore potential mediating effects using two-step MR analysis. Initially, MR analysis was performed between positive lipids and CSF metabolites (Fig. 3). For Cholesterol as the exposure, 27 causally linked CSF metabolites were identified, and for PC (18:1_0:0) as the exposure, 30 causally linked CSF metabolites were found (Supplementary Table S9). Subsequently, these findings were subjected to two-sample MR analysis with LOAD to confirm potential mediators (Fig. 4). Ultimately, three CSF metabolites were found to causally relate to LOAD. Elevated N-acetylthreonine (IVW OR=1.024; 95% CI 1.007–1.041; $P=5.71\times10^{-3}$) was associated with increased LOAD risk, while increases in Citrulline (IVW OR=0.978; 95% CI 0.964–0.991; $P=1.48\times10^{-3}$) and N-acetylneuraminate (NeuAC, IVW OR=0.974; 95% CI 0.954–0.995; $P=0.014$) were associated with reduced risk (Supplementary Table S10). Subsequently, the indirect effects and mediation proportions mediated by CSF metabolites were calculated. It was found that for N-acetylthreonine and Citrulline, the directions of direct and indirect effects in the mediation pathway were inconsistent, making it impossible to calculate mediation proportions. Overall, after a series of analyses, the mediation proportion of NeuAC was found to be 21.02% (3.25%, 45.50%) as shown in Table 2. All tests passed checks for pleiotropy and heterogeneity, as detailed in Supplementary Tables S11-12. Furthermore, the positive results showed no evidence of reverse causation (Supplementary Tables S13-14).

Replication-mediated effects in blood metabolites

Following the aforementioned methodology, we conducted a mediation test for NeuAC in blood using a two-step MR analysis. We found no causal relationship between NeuAC in blood, lipid bodies, and LOAD with

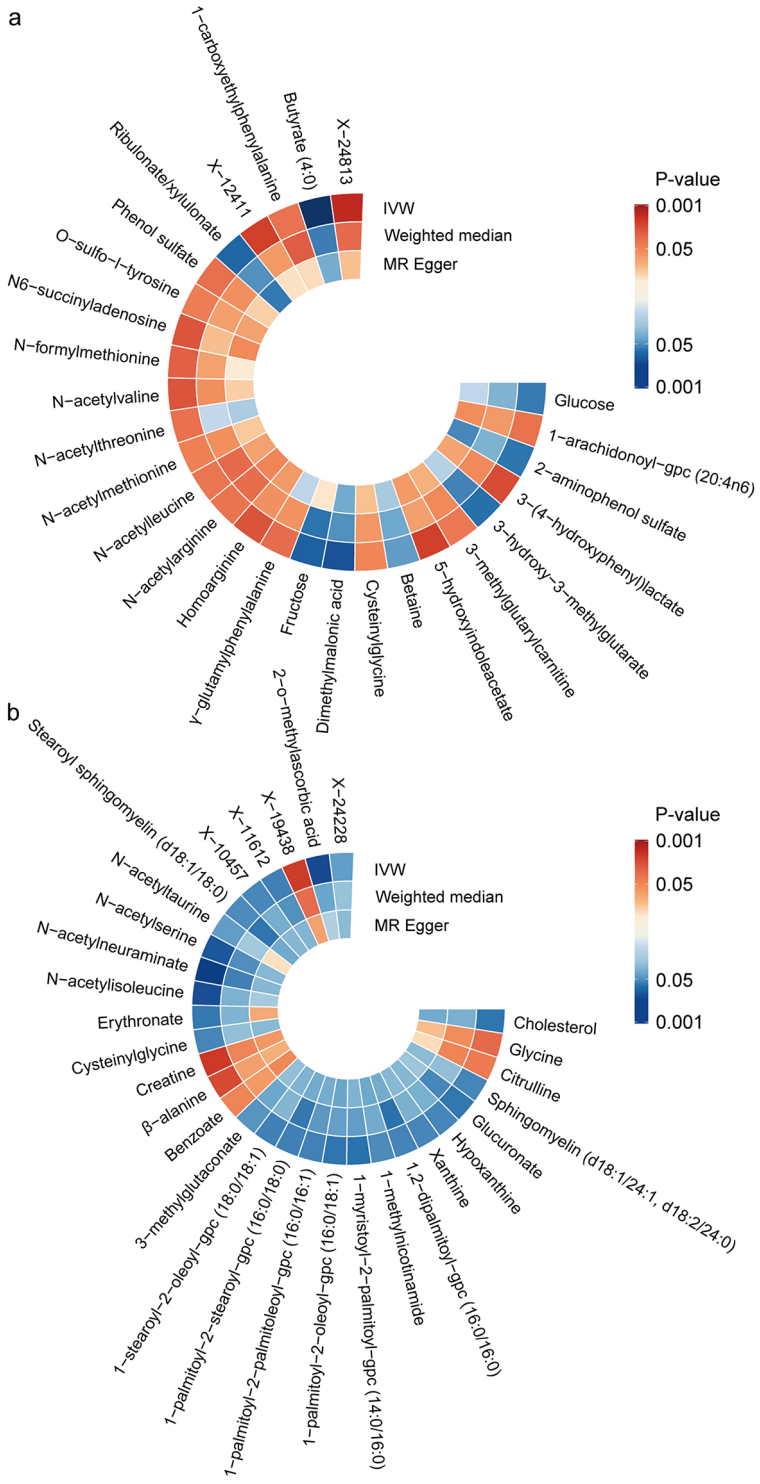

Fig. 3 Mendelian randomization analyses show causal effects between lipids on CSF metabolites. (**a**). Heatmap of cholesterol as exposure. (**b**). Heatmap of Phosphatidylcholine (18:1_0:0) as exposure

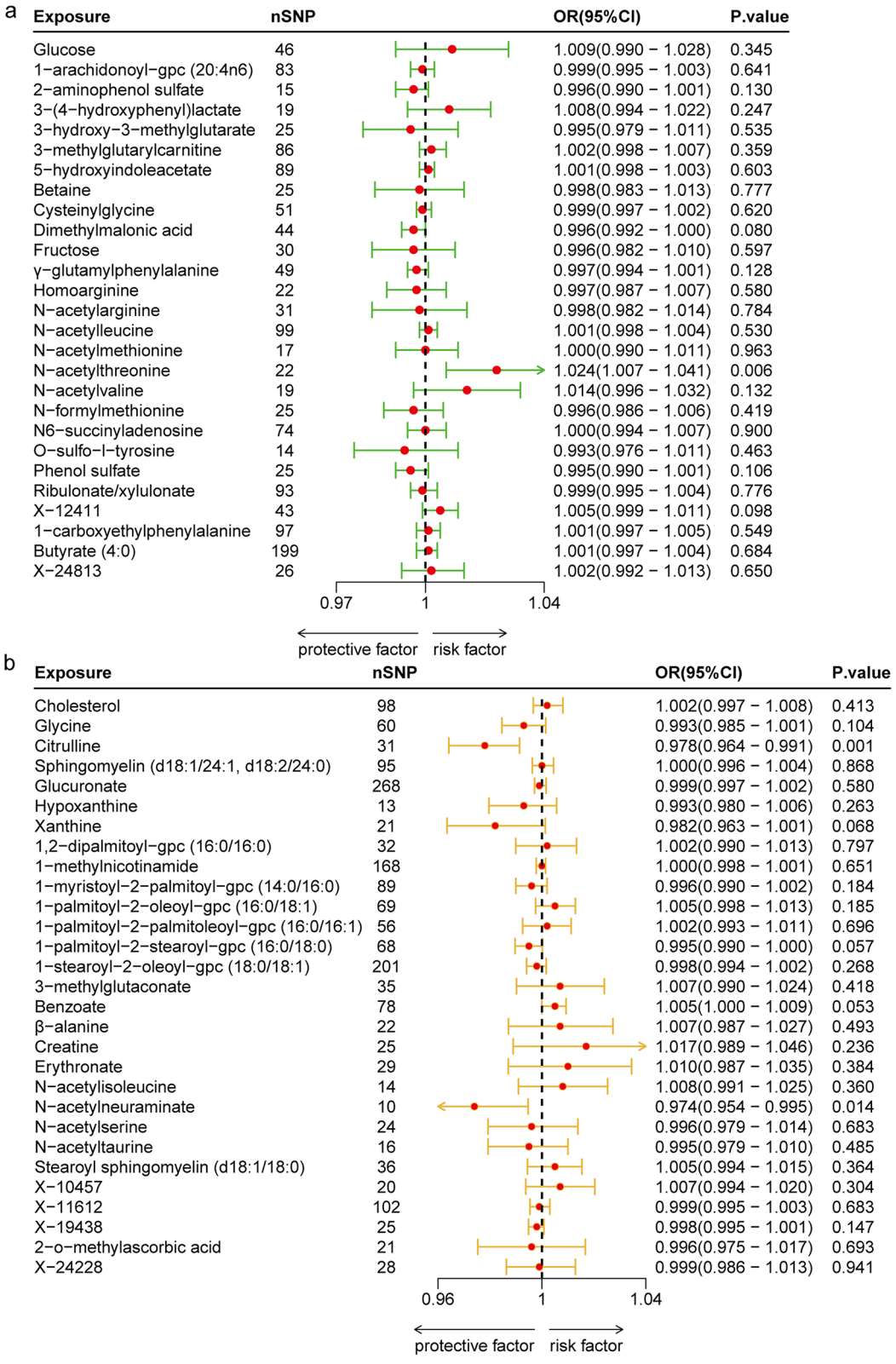

Fig. 4 Mendelian randomization analyses show causal effects between CSF metabolites on LOAD. (**a**). The exposure comprises positive CSF metabolites identified using cholesterol as the exposure variable. (**b**). The exposure comprises positive CSF metabolites identified using PC (18:1_0:0) as the exposure variable

Table 2 The mediation effect of lipids on LOAD via CSF metabolites

Mediator	Total effect	Step 1 effect	Step 2 effect	Mediated proportion (%)
	β (95% CI)	β1 (95% CI)	β2 (95% CI)	(95% CI)
N-acetylthreonine	-0.011(-0.015, -0.007)	0.055(0.031, 0.079)	0.023(0.015, 0.032)	-
Citrulline	0.015(0.010, 0.020)	0.060(0.032, 0.088)	-0.023(-0.030, -0.015)	-
N-acetylneuraminate	0.015(0.010, 0.020)	-0.118(-0.153, -0.082)	-0.026(-0.037, -0.016)	21.02(3.25, 45.50)

Total effect: The effect of lipids on LOAD, Step 1 effect: The effect of lipids on CSF metabolites, Step 1 effect: The effect of CSF metabolites on LOAD

the IVW method yielding P-values of 0.195 and 0.577, respectively (Table 3).

Metabolic pathway analysis

Further analysis using the IVW method on CSF metabolites with a positive causal relationship with LOAD (Supplementary Table S15) revealed 15 metabolic pathways potentially involved in the development and progression of LOAD. However, only the pathways "Valine, leucine, and isoleucine biosynthesis" ($P=1.86\times10^{-5}$), "Lysine degradation" ($P=1.22\times10^{-3}$), and "Valine, leucine and isoleucine degradation" ($P=2.85\times10^{-3}$) were considered significant (Fig. 5).

Discussion

We conducted a systematic two-sample and two-step MR study to explore the causal relationship between lipids and LOAD, and to analyze the potential mediating role of CSF metabolites. Results indicated that increased levels of Cholesterol are associated with a reduced risk of LOAD, while PC (18:1_0:0) acts as a risk factor. The two-step MR analysis identified 27 and 30 CSF metabolites causally linked with Cholesterol and PC (18:1_0:0) respectively. Subsequent validation suggested potential mediation by N-acetylthreonine, Citrulline, and N-acetylneuraminate. Only N-acetylthreonine was identified as a risk factor, while the other two metabolites were protective. However, only PC (18:1_0:0) was shown to increase LOAD risk, potentially through a reduction in CSF NeuAC levels, as indicated by directionality discrepancies in IVW estimates. All mediation analyses excluded the possibility of reverse causation and showed no signs of horizontal pleiotropy or heterogeneity. Replication experiments in blood metabolites did not yield corresponding positive results.

Cholesterol metabolism plays a significant role in the pathogenesis of LOAD, but the relationship between cholesterol levels and LOAD risk is complex and varies by cholesterol subtype. A study suggest that higher total cholesterol levels in midlife increase Alzheimer's risk by contributing to amyloid-beta deposition and neurofibrillary tangle formation—the hallmark features of Alzheimer's pathology [41]. Conversely, other studies indicate that higher cholesterol levels in later life might protect against cognitive decline [42, 43]. This inconsistency highlights the importance of considering cholesterol subtypes separately. High-density lipoprotein cholesterol (HDL-C), known as "good cholesterol" [44], facilitates reverse cholesterol transport and possesses anti-inflammatory and antioxidant properties that may confer neuroprotective effects. Higher HDL-C levels are associated with a reduced risk of cognitive decline and AD [45]. In contrast, low-density lipoprotein cholesterol (LDL-C), or "bad cholesterol," when elevated, can lead to atherosclerotic plaque formation, reducing cerebral blood flow and contributing to cognitive impairment. Elevated LDL-C levels are linked to an increased risk of AD; oxidized LDL induces inflammatory responses and oxidative stress, leading to neuronal damage and promoting amyloid-beta production [46]. Therefore, differentiating between cholesterol subtypes is crucial for understanding their distinct roles in LOAD pathogenesis, and future research should focus on HDL-C and LDL-C separately to refine our understanding of lipid metabolism in LOAD.

Presently, research on PC in LOAD is sparse. Earlier studies reported reductions in three specific PCs in AD patients: 16:0/20:5, 16:0/22:6, and 18:0/22:6 [47]. PC, a type of 1,2-diacylglycerol phospholipid, is an essential component of cellular membranes and constitutes

Table 3 Detailed results of MR analysis for causal effects of lipids, blood metabolite and LOAD

Exposure	Outcome	Method	Beta	SE	OR	95%CI	*P*-value
PC(18:1_0:0)	NeuAC (copy)	IVW	-0.050	0.038	0.951	0.882–1.026	0.195
		Weighted median	-0.044	0.055	0.957	0.859–1.066	0.427
		MR Egger	0.038	0.104	1.038	0.847–1.273	0.721
NeuAC (copy)	LOAD	IVW	-0.002	0.004	0.998	0.989–1.006	0.577
		Weighted median	-0.002	0.006	0.998	0.985–1.011	0.737
		MR Egger	-0.004	0.013	0.996	0.971–1.021	0.738

SE, standard error; OR, odds ratio; CI, confidence interval; IVW, Inverse variance weighted. The lipid species are named in the following notation: class name < sum of carbon atoms>:< sum of double bonds>;< sum of hydroxyl groups>. Acyl chains, "_"

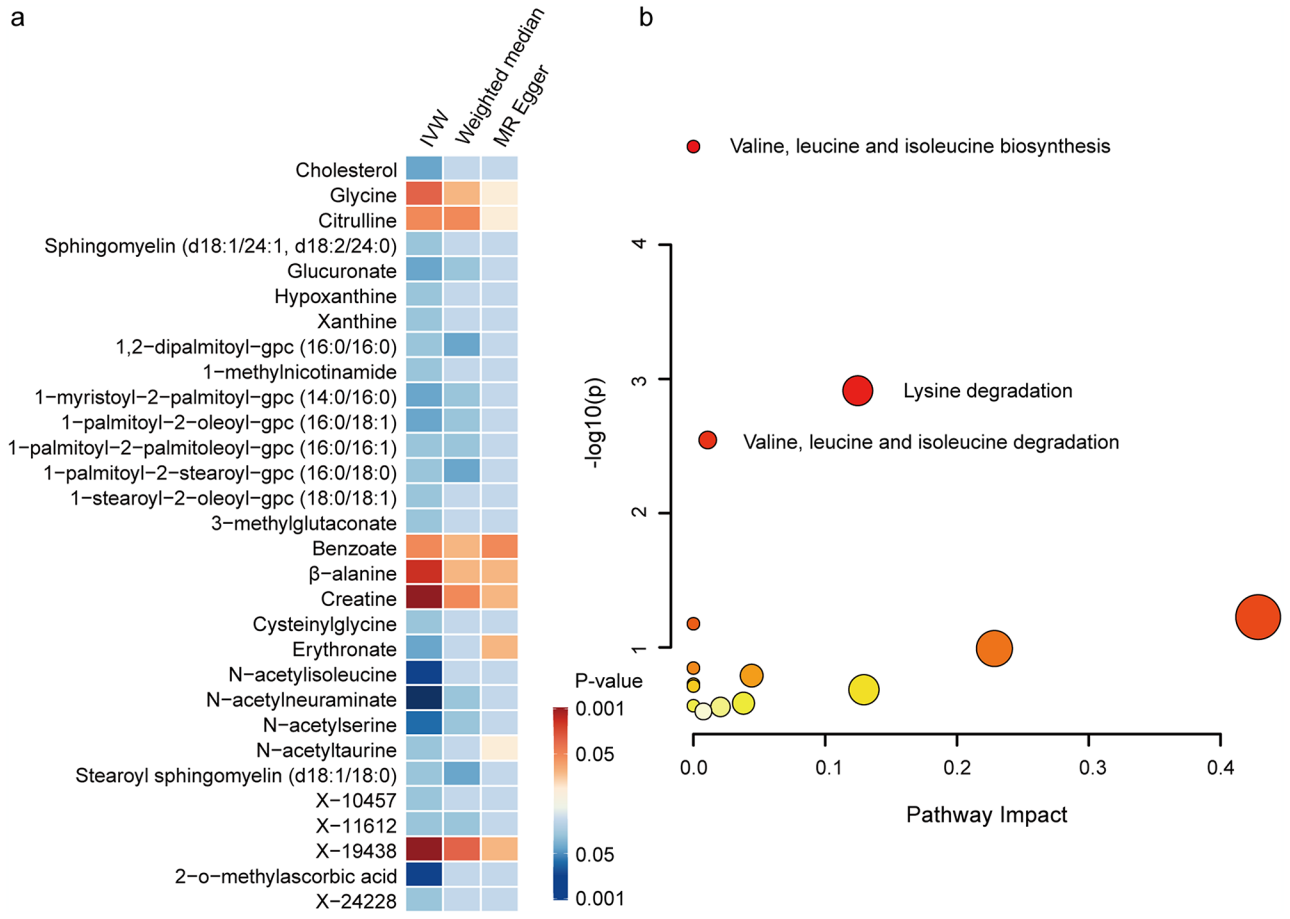

Fig. 5 (**a**). Mendelian randomization analyses show causal effects between all positive CSF metabolites on LOAD. (**b**). Potential metabolic pathways associated with LOAD. Pathway analysis utilizing the "Kyoto Encyclopedia of Genes and Genomes" (KEGG) was conducted. The circles in the visualization are colored and sized based on their respective *P*-values (yellow indicating higher *P*-value and red indicating lower *P*-value) and pathway impact values (larger circles correspond to higher impact scores) derived from topological analysis

approximately 95% of the total choline compounds in most tissues [48, 49]. The findings of Luke Whiley and others show some deviations from our conclusions. This variance may be attributed to the diverse molecular forms of PC, which have different functions across pathways. Moreover, the small sample size of Whiley's cohort may have amplified certain confounding factors, potentially obscuring the true effects. Despite these challenges, we maintain that the outcomes predicted by genetic studies are reliable. However, further research is required to fully understand these mechanisms.

N-acetylneuraminate (also known as sialic acid), an acidic monosaccharide with a nine-carbon backbone, is a structural and functional component of brain gangliosides [50]. It correlates with the content of DHA and total long-chain polyunsaturated fatty acids in the brain's sphingolipid sites [51]. Genetic studies in mice and human diseases have demonstrated that ganglioside metabolism plays a crucial role in regulating neuronal excitability, axon-myelin interactions, axonal stability, and regeneration [51, 52]. Notably, the concentration of NeuAC in gangliosides varies significantly with age. While sialic acid levels rise rapidly early after birth, they plateau before the age of 50 and subsequently decline [53], a pattern that corresponds with the onset age and causes of LOAD. Additionally, NeuAC monomers can polymerize into polysialic acid, a major component of the neural cell adhesion molecule (NCAM) expressed on the surfaces of central nervous system (CNS) cells [54]. NCAM is critical in synaptogenesis, neuronal plasticity, and memory formation [55, 56]. This is consistent with our identification of NeuAC as a mediator in the protective effects within regulatory pathways. Despite its significance, NeuAC has been scarcely studied in the context of LOAD. Future research in this field may lead to the clinical application of NeuAC as a novel biomarker.

The presence of the blood-brain barrier and blood-spinal cord barrier results in significant differences between the metabolite profiles of blood and cerebrospinal fluid [57, 58], explaining the inability to replicate the

NeuAC-mediated effects in blood metabolites. Furthermore, pathway analysis of cerebrospinal fluid metabolites via KEGG identified three critical pathways related to LOAD, involving the metabolism of valine, leucine, isoleucine, and lysine. Leucine, isoleucine, and valine, as essential branched-chain amino acids (BCAAs), cannot be synthesized in the body. However, nitrogen is transferred to α-ketoglutarate (αKG) via branched-chain amino acid transaminases 1 and 2 (BCAT 1/2), resulting in the production of glutamine and specific branched-chain keto acids (BCKA). These keto acids are then metabolized by the branched-chain α-keto acid dehydrogenase complex to produce branched-chain acyl-CoA (R-CoA), which are ultimately metabolized into the tricarboxylic acid cycle intermediates acetyl-CoA and succinyl-CoA [59]. Many enzymes acting on these amino acids utilize all BCAAs as substrates, thereby similarly influencing the levels of all BCAAs. This reflects the similar chemical properties and metabolism of BCAAs [60]. Glutamate, an important excitatory neurotransmitter in the brain, is synthesized in astrocytes surrounding neurons, with BCAAs, particularly leucine, playing a crucial role. Leucine enters the brain from the bloodstream more rapidly than other amino acids and contributes approximately 25% of all α-amino groups to glutamate synthesis [60]. Overall, the cerebrospinal fluid metabolites obtained via the two-step MR approach play a crucial role in the pathophysiological pathways of LOAD.

The strength of this study lies in the utilization of the most recent and comprehensive GWAS data for exposure and outcome variables, with no overlapping samples. Rigorous selection criteria were applied to exclude confounding and weak genetic instruments, enhancing the robustness and reliability of our results through systematic MR analysis and sensitivity testing. The diversity of the study cohort, encompassing multiple ethnicities, lends generalizability to our findings. Additionally, the direct contact between cerebrospinal fluid and the brain's extracellular space allows its metabolites to reflect the brain's pathophysiological conditions [61], which is crucial for early diagnosis and treatment of diseases. Nonetheless, this study has limitations. The primary limitation of our study is the relatively small sample size of the CSF metabolite GWAS cohort ($n=291$). A smaller sample size may reduce the statistical power to detect true causal associations and increases the risk of both type I and type II errors, potentially affecting the reliability and generalizability of our results. Despite this, we believe that our findings provide valuable preliminary insights into the potential mediating role of CSF metabolites in the relationship between lipids and LOAD. To mitigate the impact of the small sample size, we employed rigorous two-sample MR methods and conducted extensive sensitivity analyses to enhance the robustness of our results. We also ensured that the instrumental variables used were strong and valid by applying stringent selection criteria and quality control measures. However, we acknowledge that future studies with larger and more diverse cohorts are necessary to validate our findings, improve their generalizability, and fully elucidate the underlying biological mechanisms. Replication of our results in larger samples will strengthen the evidence for the causal relationships identified and may facilitate the development of potential biomarkers for early diagnosis and intervention in LOAD, thereby enhancing the validity and applicability of the results to broader populations.

Finally, MR analysis has only revealed genetic causality. The role of PC (18:1_0:0) as a lipid mediator in the association with LOAD still requires further validation through experimental and clinical data.

Conclusions

Our study is the first to reveal causal relationships between certain lipids and CSF metabolites and LOAD. Building on this, we also established that CSF metabolites mediate the relationship between lipids and LOAD. These findings suggest potential biomarkers for diagnosing and preventing the disease. Going forward, we advocate for further in vivo and in vitro experimental studies in this field to explore the causality, mechanisms, and therapeutic potential of these findings.

Abbreviations

AD	Alzheimer's Disease
BCAA	Branched-chain amino acid
BCAT	Branched-chain amino acid transaminase
CI	Confidence interval
CNS	Central nervous system
CSF	Cerebrospinal fluid
GWAS	Genome-Wide Association Study
IV	Instrumental variable
IVW	Inverse-variance weighted
LD	Linkage disequilibrium
LOAD	Late-Onset Alzheimer's Disease
KEGG	Kyoto Encyclopedia of Genes and Genomes
MR	Mendelian randomization
NCAM	Neural cell adhesion molecule
NeuAC	N-acetylneuraminate
OR	Odds ratio
PC	Phosphatidylcholine
PCO	Phosphatidylcholine ether
PE	Phosphatidylethanolamine
PI	Phosphatidylinositol
SD	Standard deviation
SNP	Single nucleotide polymorphism
TG	Triglyceride

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-024-05796-2.

Supplementary Material 1

Supplementary Material 2: Figure S1. Leave-one-out sensitivity analysis for lipids on LOAD.

Supplementary Material 3: Figure S2. Leave-one-out sensitivity analysis for lipids on CSF metabolites.

Supplementary Material 4: Figure S3. Leave-one-out sensitivity analysis for CSF metabolites on LOAD.

Acknowledgements
We sincerely thank the researchers and participants of the GWAS studies, as well as the data platforms, for their dedication to gathering and organizing extensive data resources. This collaborative work has greatly enhanced our knowledge of Late-Onset Alzheimer's Disease and the possible interventions.

Author contributions
JJ: Conceptualization, Software, Writing – original draft, Writing – review & editing. YLG: Writing – review. HBH: Writing – original draft. SL: Visualization, Conceptualization.

Funding
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability
All primary data from this study are detailed within the article and its supplementary materials. The GWAS summary data is available at https://www.ebi.ac.uk/gwas/. For additional information, direct inquiries to the corresponding author.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors have disclosed no conflicts of interest.

Received: 17 April 2024 / Accepted: 23 October 2024
Published online: 28 November 2024

References

1. Alzheimer's Disease facts and figures. Alzheimers Dement. 2022;18:700–89.
2. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the global burden of Disease Study 2016. Lancet Neurol. 2019;18:88–106.
3. Hou X-H, Suckling J, Shen X-N, Liu Y, Zuo C-T, Huang Y-Y, et al. Multipredictor risk models for predicting individual risk of Alzheimer's disease. J Translational Med. 2023;21:768.
4. Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL. Alzheimer's disease. Nat Rev Dis Primers. 2015;1:15056.
5. Lau V, Ramer L, Tremblay M-È. An aging, pathology burden, and glial senescence build-up hypothesis for late onset Alzheimer's disease. Nat Commun. 2023;14:1670.
6. Mahley RW. Central Nervous System Lipoproteins: ApoE and Regulation of Cholesterol Metabolism. Arterioscler Thromb Vasc Biol. 2016;36:1305–15.
7. Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45:1452–8.
8. He J, Huang M, Li N, Zha L, Yuan J. Genetic Association and Potential Mediators between Sarcopenia and Coronary Heart Disease: A Bidirectional Two-Sample, Two-Step Mendelian Randomization Study. Nutrients. 2023;15:3013.
9. Reitz C. Dyslipidemia and the Risk of Alzheimer's Disease. Curr Atheroscler Rep. 2013;15:307.
10. Mills JD, Nalpathamkalam T, Jacobs HIL, Janitz C, Merico D, Hu P, et al. RNA-Seq analysis of the parietal cortex in Alzheimer's disease reveals alternatively spliced isoforms related to lipid metabolism. Neurosci Lett. 2013;536:90–5.
11. Arvanitakis Z, Wilson RS, Bienias JL, Evans DA, Bennett DA. Diabetes mellitus and risk of Alzheimer disease and decline in cognitive function. Arch Neurol. 2004;61:661–6.
12. Pasinetti GM, Eberstein JA, METABOLIC SYNDROME AND THE ROLE OF DIETARY, LIFESTYLES IN ALZHEIMER'S DISEASE. J Neurochem. 2008;106:1503–14.
13. Czech C, Berndt P, Busch K, Schmitz O, Wiemer J, Most V, et al. Metabolite Profiling of Alzheimer's Disease Cerebrospinal Fluid. PLoS ONE. 2012;7:e31501.
14. Nagata Y, Hirayama A, Ikeda S, Shirahata A, Shoji F, Maruyama M, et al. Comparative analysis of cerebrospinal fluid metabolites in Alzheimer's disease and idiopathic normal pressure hydrocephalus in a Japanese cohort. Biomark Res. 2018;6:5.
15. Panyard DJ, Kim KM, Darst BF, Deming YK, Zhong X, Wu Y, et al. Cerebrospinal fluid metabolomics identifies 19 brain-related phenotype associations. Commun Biol. 2021;4:63.
16. Refolo LM, Malester B, LaFrancois J, Bryant-Thomas T, Wang R, Tint GS, et al. Hypercholesterolemia accelerates the Alzheimer's amyloid pathology in a transgenic mouse model. Neurobiol Dis. 2000;7:321–31.
17. Wolozin B, Kellman W, Ruosseau P, Celesia GG, Siegel G. Decreased prevalence of Alzheimer disease associated with 3-hydroxy-3-methyglutaryl coenzyme A reductase inhibitors. Arch Neurol. 2000;57:1439–43.
18. Puglielli L, Tanzi RE, Kovacs DM. Alzheimer's disease: the cholesterol connection. Nat Neurosci. 2003;6:345–51.
19. Wang Q, Zang F, He C, Zhang Z, Xie C. Alzheimer's Disease Neuroimaging Initiative. Dyslipidemia induced large-scale network connectivity abnormality facilitates cognitive decline in the Alzheimer's disease. J Translational Med. 2022;20:567.
20. Davey Smith G, Ebrahim S. Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32:1–22.
21. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601.
22. Smith GD, Ebrahim S. Mendelian randomization: Prospects, potentials, and limitations. Int J Epidemiol. 2004;33:30–42.
23. Burgess S, Thompson SG, CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol. 2011;40:755–64.
24. Qiu Y-F, Ye J, Xie J-J, Mao X-T, Liu Y-L, Fang Q, et al. Pancreatitis affects gut microbiota via metabolites and inflammatory cytokines: an exploratory two-step Mendelian randomisation study. Mol Genet Genomics. 2024;299:36.
25. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. JAMA. 2021;326:1614–21.
26. Ottensmann L, Tabassum R, Ruotsalainen SE, Gerl MJ, Klose C, Widén E, et al. Genome-wide association analysis of plasma lipidome identifies 495 genetic associations. Nat Commun. 2023;14:6934.
27. Wightman DP, Jansen IE, Savage JE, Shadrin AA, Bahrami S, Holland D, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. Nat Genet. 2021;53:1276–82.
28. Chen Y, Lu T, Pettersson-Kymmer U, Stewart ID, Butler-Laporte G, Nakanishi T, et al. Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases. Nat Genet. 2023;55:44–53.
29. Sekula P, Del Greco MF, Pattaro C, Köttgen A. Mendelian randomization as an approach to assess causality using observational data. J Am Soc Nephrology: JASN. 2016;27:3253–65.
30. Long Y, Tang L, Zhou Y, Zhao S, Zhu H. Causal relationship between gut microbiota and cancers: a two-sample Mendelian randomisation study. BMC Med. 2023;21:66.
31. Lin Z, Deng Y, Pan W. Combining the strengths of inverse-variance weighting and Egger regression in Mendelian randomization using a mixture of regressions model. PLoS Genet. 2021;17:e1009922.
32. Burgess S, Butterworth A, Thompson SG. Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. Genet Epidemiol. 2013;37:658–65.
33. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40:304–14.
34. 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.

35. Burgess S, Daniel RM, Butterworth AS, Thompson SG, the EPIC-InterAct Consortium. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. Int J Epidemiol. 2015;44:484–95.
36. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. Int J Epidemiol. 2012;41:161–76.
37. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017;32:377–89.
38. Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018;50:693–8.
39. Cohen JF, Chalumeau M, Cohen R, Korevaar DA, Khoshnood B, Bossuyt PMM. Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. J Clin Epidemiol. 2015;68:299–306.
40. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51:1173–82.
41. Anstey KJ, Lipnicki DM, Low L-F. Cholesterol as a risk factor for dementia and cognitive decline: A systematic review of prospective studies with meta-analysis. Am J Geriatric Psychiatry: Official J Am Association Geriatric Psychiatry. 2008;16:343–54.
42. Notkola IL, Sulkava R, Pekkanen J, Erkinjuntti T, Ehnholm C, Kivinen P, et al. Serum total cholesterol, apolipoprotein E epsilon 4 allele, and Alzheimer's disease. Neuroepidemiology. 1998;17:14–20.
43. Mielke MM, Zandi PP, Sjögren M, Gustafson D, Östling S, Steen B, et al. High total cholesterol levels in late life associated with a reduced risk of dementia. Neurology. 2005;64:1689–95.
44. Kjeldsen EW, Nordestgaard LT, Frikke-Schmidt R. HDL cholesterol and non-cardiovascular disease: A narrative review. Int J Mol Sci. 2021;22:4547.
45. Matsuzaki T, Sasaki K, Hata J, Hirakawa Y, Fujimi K, Ninomiya T, et al. Association of alzheimer disease pathology with abnormal lipid metabolism. Neurology. 2011;77:1068–75.
46. Anstey KJ, Ashby-Mitchell K, Peters R. Updating the evidence on the association between serum cholesterol and risk of late-life dementia: Review and meta-analysis. J Alzheimers Dis. 2017;56:215.
47. Whiley L, Sen A, Heaton J, Proitsi P, García-Gómez D, Leung R, et al. Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. Neurobiol Aging. 2014;35:271–8.
48. Frisardi V, Panza F, Seripa D, Farooqui T, Farooqui AA. Glycerophospholipids and glycerophospholipid-derived lipid mediators: a complex meshwork in Alzheimer's disease pathology. Prog Lipid Res. 2011;50:313–30.
49. Schaeffer EL, da Silva ER, Novaes B, de Skaf A, Gattaz HD. Differential roles of phospholipases A2 in neuronal death and neurogenesis: implications for Alzheimer disease. Prog Neuropsychopharmacol Biol Psychiatry. 2010;34:1381–9.
50. Kolter T. Ganglioside Biochemistry. ISRN Biochem. 2012;2012:506160.
51. Wang B, Brand-Miller J. The role and potential of sialic acid in human nutrition. Eur J Clin Nutr. 2003;57:1351–69.
52. Schnaar RL, Gerardy-Schahn R, Hildebrandt H. Sialic Acids in the Brain: Gangliosides and Polysialic Acid in Nervous System Development, Stability, Disease, and Regeneration. Physiol Rev. 2014;94:461–518.
53. Schnaar RL. Glycolipid-mediated cell-cell recognition in inflammation and nerve regeneration. Arch Biochem Biophys. 2004;426:163–72.
54. Wang B. Molecular mechanism underlying sialic acid as an essential nutrient for brain development and cognition. Adv Nutr. 2012;3:S465–72.
55. Wang B. Sialic Acid Is an Essential Nutrient for Brain Development and Cognition. Annu Rev Nutr. 2009;29:177–222.
56. Colley KJ, Kitajima K, Sato C. Polysialic acid: biosynthesis, novel functions and applications. Crit Rev Biochem Mol Biol. 2014;49:498–532.
57. Banks WA, Erickson MA. The blood–brain barrier and immune function and dysfunction. Neurobiol Dis. 2010;37:26–32.
58. Zhang SL, Yue Z, Arnold DM, Artiushin G, Sehgal A. A Circadian Clock in the Blood-Brain Barrier Regulates Xenobiotic Efflux. Cell. 2018;173:130–e13910.
59. Sivanand S, Vander Heiden MG. Emerging roles for branched chain amino acid metabolism in cancer. Cancer Cell. 2020;37:147–56.
60. Shimomura Y, Kitaura Y. Physiological and pathological roles of branched-chain amino acids in the regulation of protein and energy metabolism and neurological functions. Pharmacol Res. 2018;133:215–7.
61. Ibáñez C, Simó C, Barupal DK, Fiehn O, Kivipelto M, Cedazo-Mínguez A, et al. A new metabolomic workflow for early detection of Alzheimer's disease. J Chromatogr A. 2013;1302:65–71.

Publisher's note