

Research Article

Qian Liu[#], Ling-bing Meng[#], Tian-qi Qi, Ya-Qing Ma, Guo-wei Liang^{*}

Exploring cerebrospinal fluid metabolites, cognitive function, and brain atrophy: Insights from Mendelian randomization

<https://doi.org/10.1515/med-2025-1237>

received April 9, 2025; accepted June 13, 2025

Abstract: Disruption of cerebrospinal fluid (CSF) metabolites affects brain function and cognition, potentially altering the brain structure. To elucidate the causal relationships between CSF metabolites and the neurological outcomes, we conducted a two-sample Mendelian randomization analysis. Genome-wide association data from 689 individuals of European descent provided exposure levels for metabolites, analyzed alongside gene associations for cognitive performance ($N = 257,841$), brain atrophy measures (cortical surface area and thickness; $N = 51,665$), and hippocampal volume ($N = 33,536$). Our analysis identified 30 metabolites exhibiting causal associations with brain atrophy and cognitive function: 20 linked to cognition and 10 to structural atrophy. Notably, butyrate correlated strongest with the cortical surface area, bilirubin with the cortical thickness, methionine sulfoxide with the hippocampal volume, threonate with cognitive performance, while oxidized Cys-gly, *N*6-succinyladenosine, and *N*-acetylglucosamine were linked to fluid intelligence, prospective memory, and reaction time,

respectively. Pathway analyses revealed that butanoate and niacinamide/niacin ester metabolism are significantly associated with brain atrophy and cognitive performance. These findings position CSF metabolites as promising therapeutic targets for neurodegenerative diseases, providing a causal framework to prioritize interventions. Experimental studies building on this genetic evidence hold potential to accelerate the development of mechanism-driven therapies targeting metabolic pathways in neurodegeneration.

Keywords: cerebrospinal fluid, metabolomics, cognition, brain atrophy, Mendelian randomization

1 Introduction

With increasing life expectancy, dementia has emerged as a significant global health challenge, with an estimated 150 million people projected to be affected by 2050 [1]. The brain starts changing years before clinical symptoms appear [2]. Early biomarkers help in studying the pathological mechanisms and potential prevention strategies.

In recent years, metabolomics research has revealed new insights into complex diseases such as diabetes, obesity, cancer, and Alzheimer's disease (AD) [3–6], elucidating their disease mechanisms through the identification of relevant metabolites and discovering potential biomarkers for diagnosis and prognosis. Although most studies focus on blood or urine samples, cerebrospinal fluid (CSF) is particularly crucial for psychiatric and neurological disorders [7]. CSF can directly reflect physiological changes in the central nervous system. For example, in AD, CSF is an important source of biomarkers like amyloid-beta and phosphorylated tau [8]. A wealth of observational cohort studies have investigated the relationship between metabolites and brain structure and cognition, suggesting that these metabolites may have neuroprotective, anti-inflammatory, or other effects [9–11].

Due to the tendency for observational studies to encounter reverse causation and confounding effects, we employed

[#] These authors contributed equally to this work and share first authorship.

*** Corresponding author: Guo-wei Liang**, Department of Clinical Laboratory, Aerospace Center Hospital, Beijing, 100049, China, e-mail: lgw721@126.com

Qian Liu, Tian-qi Qi: Department of Clinical Laboratory, Aerospace Center Hospital, Beijing, 100049, China

Ling-bing Meng: Cardiometabolic Medicine Center, National Clinical Research Center for Cardiovascular Diseases, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

Ya-Qing Ma: The Key Laboratory of Geriatrics, Beijing Institute of Geriatrics, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing Hospital/National Center of Gerontology of National Health Commission, Dong Dan, Beijing, 100730, China
ORCID: Guo-wei Liang 0000-0002-9772-7286

Mendelian randomization (MR) using genetic variants as instrumental variables (IVs) based on Mendel's law to explore causal relationships between exposure and outcome [12]. Cognitive function and brain atrophy markers are considered key features of AD and other forms of dementia [13,14], and a larger number of cases may be more suitable for dissecting potential causal relationships between CSF metabolites and dementia phenotypes.

This study aims to investigate 338 CSF metabolites closely associated with the onset and progression of AD. By identifying metabolites causally linked to AD, we seek to understand their biological significance and their feasibility as biomarkers or therapeutic agents. Therefore, in this study, we employ two-sample MR to assess causal relationships between CSF metabolite levels and measures of general cognitive function (with three additional domains of the fluid intelligence score, prospective memory result, and reaction time) and measures of brain atrophy (cerebral cortical surface area and thickness and hippocampal volume).

2 Method

2.1 Study design

We conducted a two-sample MR study, adhering to three hypotheses: (1) the IV is closely associated with the exposure of interest; (2) the IV is independent of confounding factors; and (3) the IV does not affect the outcome except through the exposure [15]. The data used in this study are summary statistics from genome-wide association studies (GWAS). The original studies obtained specific ethical approval and informed consent. The study design is illustrated in Figure 1.

2.2 Selection of IVs

Genetic instruments associated with 338 CSF metabolites were derived from recent GWAS studies. These studies

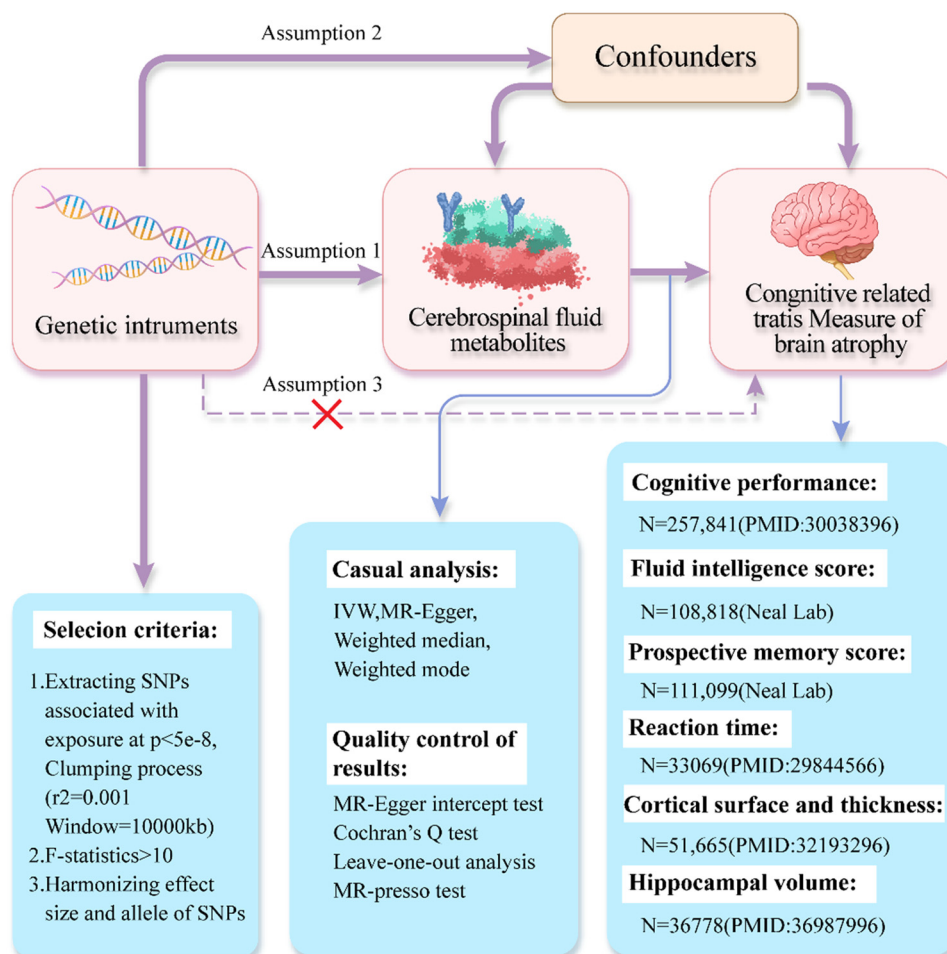


Figure 1: Overview of MR analysis. Assumption 1, genetic instruments are strongly associated with the exposures of interest; Assumption 2, genetic instruments are independent of confounding factors; Assumption 3, genetic instruments are not associated with outcome and affect outcome only via exposures. IVW, inverse variance weighted; MR-PRESSO, MR-Pleiotropy RESidual sum and outlier; SNPs, single-nucleotide polymorphisms.

performed genome-wide association analyses of 338 CSF metabolites in cognitively healthy participants with an average age of 60 years (data available for download from ftp://ftp.biostat.wisc.edu/pub/lu_group/Projects/MWAS/). Notably, this represents the most comprehensive report to date on genetic loci related to CSF metabolites, identifying nearly 1.8 million SNPs through genome-wide association scans and high-throughput metabolic profiling conducted by Panyard et al. [16]. The detailed names of the 338 metabolites are listed in Table S1, with metabolites named X having unknown chemical properties. Specifically, the study included 689 participants, of whom 532 were from the Wisconsin Alzheimer's Disease Research Center (WADRC) and 168 were from the Wisconsin Registry for Alzheimer's Prevention, each providing unique CSF samples for metabolite analysis. Selection criteria for WADRC participants included age ≥ 45 years, decision-making capacity, and ability to fast for 12 h, with exclusion criteria such as specific medical histories including renal dysfunction, congestive heart failure, and major neurological disorders (excluding dementia).

To meet hypothesis (1), we identified IVs related to blood metabolites through rigorous screening criteria from multiple perspectives. Given the moderate number of SNPs associated with metabolites, we selected a stringent significance threshold of $p < 5 \times 10^{-8}$ to choose SNPs associated with the metabolites. We used a p -value threshold of less than 5×10^{-8} for metabolite GWAS IVs, as suggested by Burgess and Thompson [26], to strengthen genetic effects, though this could exclude SNPs with minor associations. Subsequently, we aggregated SNPs by eliminating linkage disequilibrium (LD, $R^2 > 0.001$ and within 10,000 kb). To minimize weak instrument bias, we computed the R^2 and F statistics for each SNP. The R^2 and F statistics were calculated as follows:

$$R^2 = \frac{2 \times \beta^2 \times \text{EAF} \times (1 - \text{EAF})}{[2 \times \beta^2 \times (1 - \text{EAF}) + 2 \times (\text{se}(\beta))^2 \times N \times \text{EAF} \times (1 - \text{EAF})]},$$

$$F = \frac{N - k - 1}{k} \times \frac{R^2}{1 - R^2}.$$

Here, β represents the effect size of the target genetic variation, EAF denotes the allele frequency of the target genetic variation, $\text{se}(\beta)$ indicates the standard error of the effect size of the target genetic variation, R^2 is the coefficient of determination for the IVs explaining the exposure level in the regression equation, N is the sample size of the exposure, and k is the number of SNPs (IVs). SNPs with $F < 10$ are defined as poor genetic variations and are excluded [17]. Next, we filtered SNPs associated with metabolites from the results ($p < 5 \times 10^{-8}$). We further matched SNPs for exposure and outcome, removing those with reverse effects or allele inconsistency (e.g., A/G versus A/C). Then, to

satisfy hypothesis (3), we excluded SNPs from IVs associated with the outcome ($p < 5 \times 10^{-5}$). Finally, causal inference analysis was performed on metabolites with two or more SNPs [18].

2.3 Cognitive-related traits

Cognitive performance was evaluated using a meta-analysis weighted by sample size, utilizing data from the Cognitive Genomics Consortium (COGENT, $n = 35,295$) and the UK Biobank ($n = 222,543$), and extracting genetic association statistics pertaining to general cognitive performance from the most recent GWAS conducted on individuals of European ancestry ($N = 257,841$) [19]. Participants ranged in age from 16 to 102 years and had no history of stroke or dementia. In COGENT, cognitive performance was defined as the first unrotated principal component score derived from at least three different neuropsychological tests included in the study. In the UK Biobank, verbal-numerical reasoning (VNR) was assessed using 13 multiple-choice questions, with VNR scores reflecting the number of correctly answered questions within 2 min, designed to measure fluid intelligence. This assessment has been demonstrated to possess adequate reliability and validity [19,20].

GWAS data on fluid intelligence score ($N = 149,051$) and prospective memory score ($N = 152,605$) were obtained from the UK Biobank, evaluated using fluid intelligence questions and image tests, respectively [21]. Data for reaction time ($N = 330,069$) came from a study on genetic loci related to general cognitive function conducted by the Centre for Cognitive Ageing and Epidemiology at the University of Edinburgh [22]. This measure is based on the total time taken by the participants to correctly complete card-matching tasks.

2.4 Measurements of brain atrophy

Genetic associations with the cortical surface area (mm^2) and thickness (mm) were identified using a genome-wide meta-analysis of brain $T1$ -weighted magnetic resonance imaging (MRI) data from 51,665 healthy individuals of European ancestry aged 3.3–91.4 years across 60 cohorts, conducted by the Enhancing Neuro Imaging Genetics through Meta Analysis (ENIGMA) consortium [23]. Among these 60 cohorts, less than 1% of the participants had psychiatric disorders in the included case-control studies, with nearly all others being healthy. Cortical surface area was measured at the

gray–white matter boundary, and thickness was defined as the average distance between the white matter and pial surface. The total surface area and mean thickness were calculated for each participant. Genetic associations were assessed using additive models within each cohort, adjusting for age, age squared, sex, interaction of sex with age and age squared, the first four multidimensional scaling components, diagnostic status (where applicable in case–control cohorts), and scanner-specific variables.

Genetic data for hippocampal volume were derived from a GWAS meta-analysis conducted by the ENIGMA and CHARGE consortia across 65 study centers, analyzing high-resolution brain MRI scans from 33,536 healthy individuals aged 11–98 years [24]. Bilateral average hippocampal volume (mm^3) was defined as the mean volume of the left and right hippocampus. Genetic associations were evaluated within each study center, adjusting for age, age squared, sex, intracranial volume, four multidimensional scaling components, and diagnostic status. For studies involving data from multiple centers or scanners, adjustments were made for center effects, and mixed-effects models were used to account for familial relationships pertaining to family data.

2.5 Statistical analysis

Causal relationships between CSF metabolites and brain atrophy and cognitive decline were primarily assessed based on results from inverse variance weighting (IVW). The IVW estimates are derived from meta-analyses of the Wald ratio for all genetic variants [24]. IVW assumes no horizontal pleiotropy across all SNPs, providing the most accurate assessment of causal relationships [25]. Initially, we used IVW-based estimates to screen CSF metabolites causally linked to brain atrophy and cognitive decline. To ensure robust findings, we applied two additional methods to further evaluate metabolites with significant IVW estimates ($p < 0.05$). MR-Egger, weighted median (WM), and weighted mode methods were employed as complementary analyses. These methods offer more robust estimates under relaxed assumptions; WM allows for less than 50% of the SNPs to be invalid, while MR-Egger provides tests for horizontal pleiotropy and heterogeneity in the presence of all SNPs having horizontal pleiotropy [26]. Weighted mode considers each genetic variant's contribution to causal inference more precisely, making it suitable for handling complex weight data [17]. The false discovery rate (FDR) correction was applied using a standard and well-established procedure to account for multiple comparisons and control the false discovery rate in statistical analyses.

Sensitivity analysis is crucial as it allows for examination of potential violations of horizontal pleiotropy and heterogeneity that could seriously bias MR estimates. Horizontal pleiotropy is observed when IV influences outcomes through pathways other than the exposure of interest. Therefore, several tests were conducted to ensure the credibility of estimates. In this study, we employed three methods to detect and correct for heterogeneity and pleiotropy, including Cochran's Q test and MR-Egger intercept test. Heterogeneity was deemed present if Cochran's Q test yielded a p -value < 0.05 [27]. The MR-Egger intercept was calculated to assess directional pleiotropy and bias due to invalid IVs [26]. For robustness of results, we conducted leave-one-out (LOO) analysis, where each SNP was sequentially removed and MR analysis was rerun to evaluate the impact of individual SNPs [26].

To summarize, we systematically screened CSF metabolites potentially associated with brain atrophy and cognitive decline using several criteria: (1) significant p -values in the main analysis ($p < 0.05$ by IVW). (2) Consistent direction of beta values across four MR methods. (3) No heterogeneity or horizontal pleiotropy observed in MR results. (4) Minimal disruption of MR estimates by individual SNPs.

2.6 Reverse analysis

In order to investigate potential reverse causation, we also examined the relationships between genetically influenced cognitive performance and brain atrophy markers with 338 CSF metabolites. To achieve this, we extracted independent genetic variants significantly associated with cognitive performance, cortical surface area, and thickness, as well as hippocampal volume from the same GWAS dataset, utilizing these variants as IVs in our primary analyses. SNPs associated with the cortical surface area and thickness were excluded to mitigate potential pleiotropic effects. All analyses were performed using R (v3.6.3) statistical software (R Foundation for Statistical Computing, Vienna, Austria). MR analyses were conducted using the R package “TwoSample MR.” A p -value of 0.05 was considered statistically significant.

Ethics and consent declarations: This study utilized publicly available GWAS summary statistics and did not involve new human subject experimentation. Therefore, institutional ethics approval and individual participant consent were not required. The original studies obtained ethical approval in accordance with the Declaration of Helsinki.

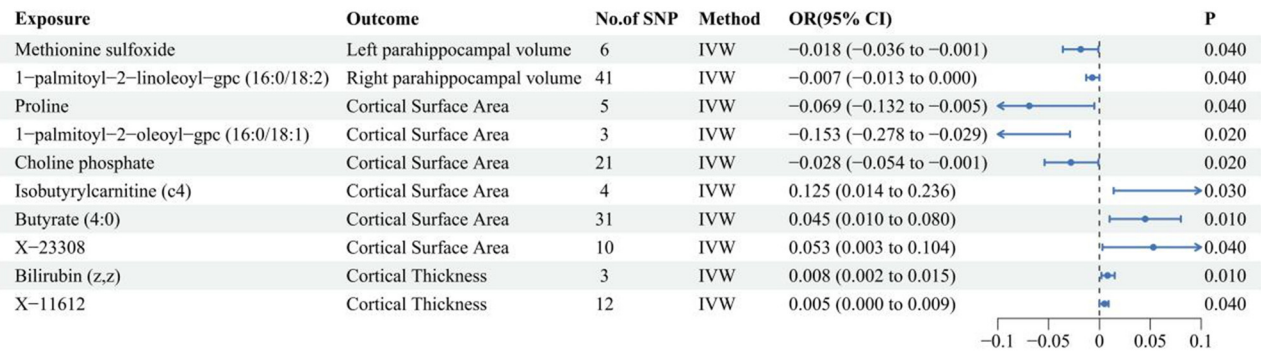


Figure 2: Association of CSF metabolites with measures of brain atrophy.

3 Results

3.1 MR results

Following strict control of the IV P1 threshold ($5e-8$) and removal of linkage disequilibrium, we proceeded with further analysis using 228 out of the 338 metabolites, as outlined in Table S1. From the screened exposure-related SNP data, SNPs corresponding to the outcome were extracted to ensure analysis consistency and comparability based on the same SNP set. A threshold P2 ($5e-5$) was applied to exclude SNPs highly correlated with the outcome, aiming to mitigate confounding effects or spurious associations. SNP datasets demonstrating no reverse causation, identified through Steiger test results, were used to

finalize the IVs for analysis (refer to Table S2). All F-statistics for SNPs associated with metabolites exceeded 10, confirming the strength of the IVs. Subsequent to IVW analysis, combined with complementary and sensitivity analyses, we pinpointed 30 metabolites potentially causally associated with brain atrophy and cognitive function.

Ten metabolites meeting strict screening criteria are associated with brain atrophy, as shown in Figure 2. Methionine sulfoxide and 1-palmitoyl-2-linoleoyl-gpc (16:0/18:2) show increased levels correlated with smaller hippocampal volume. Proline, 1-palmitoyl-2-oleoyl-gpc (16:0/18:1), choline phosphate levels, isobutyrylcarnitine (c4), butyrate (4:0), and X-23308 are correlated with the cortical surface area. Increased bilirubin (z,z) and X-11612 levels are associated with thicker cortical thickness.

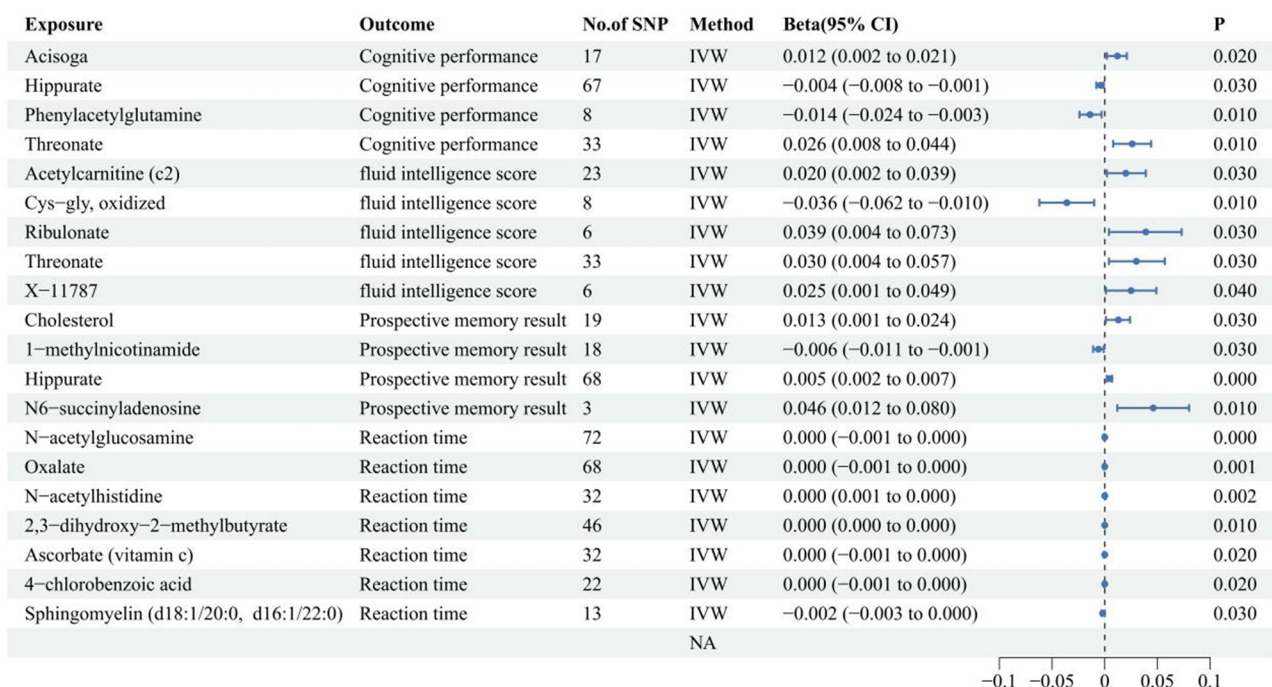


Figure 3: Association of CSF metabolites with measures of cognitive function.

Table 1: Heterogeneity and pleiotropy of CSF metabolites on brain atrophy

Exposure	Outcome	Heterogeneity test		Pleiotropy test	
		MR Egger Q	P value	MR-Egger intercept	P value
Methionine sulfoxide	Left parahippocampal volume	1.751	0.781	0.000	0.999
1-Palmitoyl-2-linoleoyl-gpc (16:0/18:2)	Right parahippocampal volume	50.357	0.105	−0.002	0.740
Proline	Cortical surface area	0.586	0.900	−0.001	0.940
1-Palmitoyl-2-oleoyl-gpc (16:0/18:1)	Cortical surface area	0.007	0.931	−0.006	0.892
Choline phosphate	Cortical surface area	14.510	0.753	0.016	0.369
Isobutyrylcarnitine (c4)	Cortical surface area	2.483	0.289	−0.002	0.929
Butyrate (4:0)	Cortical surface area	19.452	0.909	0.004	0.849
X-23308	Cortical surface area	7.579	0.476	−0.023	0.265
Bilirubin (z,z)	Cortical thickness	1.515	0.218	−0.001	0.811
X-11612	Cortical thickness	12.429	0.257	0.000	0.830

Detailed results of all MR analyses for CSF metabolites and brain atrophy are presented in Table S3.

Twenty metabolites meeting stringent screening criteria associated with cognitive function are detailed in Figure 3. Acisoga, hippurate, phenylacetylglutamine, and threonate are correlated with cognitive performance. Acetylcarnitine (c2), Cys-gly, oxidized, ribulonate, threonate, and X-11787 are associated with fluid intelligence scores. Cholesterol, 1-methylnicotinamide, hippurate, and N6-succinyladenosine are associated with prospective memory. N-Acetylglucosamine, oxalate, N-acetylhistidine, 2,3-dihydroxy-2-methylbutyrate, ascorbate, 4-chlorobenzoic acid, and sphingomyelin (d18:1/20:0, d16:1/22:0) are associated with the reaction time.

Comprehensive results of all MR analyses on CSF metabolites and cognitive function are provided in Table S4. The substances causing these changes include N-acetylglucosamine and oxalate, both of which, after adjustment using FDR multiple comparisons, have adjusted *p*-values less than 0.05.

Thirty metabolites exhibited statistically significant estimated values from IVW ($p < 0.05$), with consistent direction and magnitude of estimates across IVW, MR-Egger, and WM approaches (Tables S3 and S4). Cochran's *Q* test ($p > 0.05$) and MR-Egger intercept test ($p > 0.05$) provided robust evidence against heterogeneity and pleiotropy (Tables 1 and 2). LOO analysis supported the absence of MR estimation bias due to individual SNPs (Figure S1).

Table 2: Heterogeneity and pleiotropy of CSF metabolites on cognitive function

Exposure	Outcome	Heterogeneity test		Pleiotropy test	
		MR Egger Q	P value	MR-Egger intercept	P value
Acisoga	Cognitive performance	10.838	0.764	−0.014	0.146
Hippurate	Cognitive performance	76.364	0.158	−0.002	0.606
Phenylacetylglutamine	Cognitive performance	2.976	0.812	−0.007	0.110
Threonate	Cognitive performance	39.203	0.148	0.001	0.811
Acetylcarnitine (c2)	Fluid intelligence score	25.906	0.210	0.008	0.564
Cys-gly, oxidized	Fluid intelligence score	3.986	0.679	−0.001	0.938
Ribulonate	Fluid intelligence score	1.831	0.767	0.011	0.515
Threonate	Fluid intelligence score	40.427	0.120	−0.001	0.860
X-11787	Fluid intelligence score	3.518	0.475	0.003	0.814
Cholesterol	Prospective memory result	15.087	0.589	0.001	0.743
1-Methylnicotinamide	Prospective memory result	12.953	0.676	−0.002	0.756
Hippurate	Prospective memory result	58.792	0.723	0.000	0.942
N6-succinyladenosine	Prospective memory result	0.092	0.762	0.003	0.886
N-acetylglucosamine	Reaction time	11.947	1.000	0.000	0.964
Oxalate	Reaction time	18.912	1.000	0.000	0.352
N-acetylhistidine	Reaction time	3.346	1.000	−0.001	0.290
2,3-dihydroxy-2-methylbutyrate	Reaction time	13.532	1.000	0.000	0.552
Ascorbate (vitamin c)	Reaction time	8.217	1.000	0.000	0.958
4-Chlorobenzoic acid	Reaction time	6.364	0.998	0.000	0.525
Sphingomyelin (d18:1/20:0, d16:1/22:0)	Reaction time	0.725	1.000	0.000	0.729

3.2 Reverse analysis

In total, 147, 35, 42, 3, 12, 6, and 13 SNPs were, respectively, identified as associated with cognitive performance, reaction time, fluid intelligence score, prospective memory result, cortical surface area, cortical thickness, and hippocampal volume, which also appeared in CSF metabolite GWAS. Using the IVW method, changes in choline phosphate, hippurate, dimethylmalonic acid, homoarginine, citramalate, inosine, 3-hydroxy-3-methylglutarate, pyridoxate, myo-inositol, *N*-acetylglutamate, glycerol, argininate, isovalerate (i5:0), and X-12104 metabolites associated with cognitive performance were detected (Table S7). Comprehensive results of all MR analyses for cognitive performance and metabolites are presented in Table S8.

3.3 Metabolic pathway analysis

Metabolic pathway analysis based on metabolites identified two important metabolic pathways associated with brain atrophy and cognitive performance (Table S9). Our results indicate that “butanoate metabolism” ($p = 0.0376$) is associated with improvement in brain atrophy (Figure 4a), and “niacinamide and niacin ester metabolism” is associated with cognitive function ($p = 0.0283$) (Figure 4b).

4 Discussion

We identified 30 metabolites potentially and causally associated with brain atrophy and cognitive function, including 27 with known chemical compositions and 3 with unknown identities. The 27 known metabolites span categories such as amino acids, peptides, lipids, nucleic acids, vitamins, carbohydrates, and xenobiotics. Metabolic pathway analysis revealed two key pathways associated with brain atrophy and cognitive function: “butanoate metabolism” is linked to improvement in brain atrophy, and “niacinamide and niacin ester metabolism” is associated with cognitive performance. Additionally, there is a reverse causality between hippurate and cognitive performance.

The hippocampus is a crucial neural structure in the brain, particularly essential for learning and memory functions. Changes in its volume are typically associated with cognitive decline and neurodegenerative diseases such as AD. Generally, a reduction in the hippocampal volume is considered an early sign of cognitive decline. Our study found that increases in methionine sulfoxide and 1-palmitoyl-2-linoleoyl-gpc (16:0/18:2) are correlated with decreased hippocampal volume, consistent with the previous research findings. Methionine sulfoxide is an oxidized form of sulfur-containing amino acid, commonly associated with oxidative stress and antioxidant processes [28]. Methionine sulfoxide has also been reported to be closely related to the occurrence, development, and effective treatment of AD [29]. A specific phospholipid

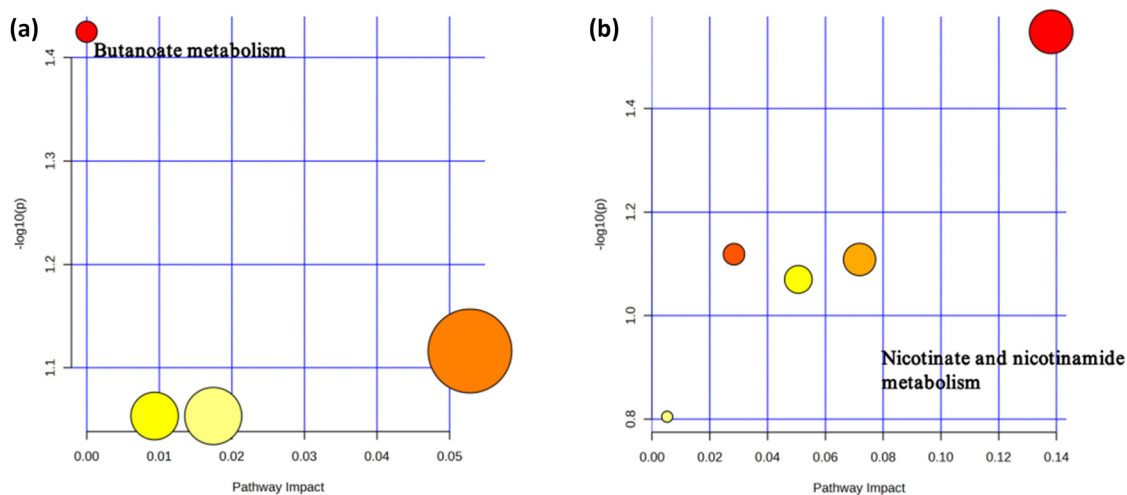


Figure 4: Potential metabolic pathways associated with brain atrophy and cognitive function. (a) Potential metabolic pathways involved in changes related to brain atrophy in MR positive analysis (based on metabolites). (b) Potential metabolic pathways involved in changes related to cognitive function in MR positive analysis (based on metabolites). Based on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The color and size of each circle represent the p -value from topological analysis (yellow: higher p -value, red: lower p -value) and pathway impact score (larger circles indicate higher impact scores).

called 16:0/18:2, which is one of the main components of the cell membrane, may play an important role in cell signal transduction, lipid metabolism, and neuroprotection.

The cortical surface area of the brain refers to the total surface area of the outer gray matter, serving as a critical indicator of the brain structure and functional development. Studies indicate that the cortical surface area may be affected in AD patients, particularly in the early stages of the disease. Cortical atrophy and thinning are observed in AD patients, possibly linked to pathological processes such as neuronal loss, synaptic damage, and brain inflammation [30]. Proline primarily functions as an amino acid involved in protein structural stability and antioxidative effects in organisms, and is a neurotoxin associated with schizophrenia [31]. Our research has found a correlation between proline and the reduced cortical surface area, consistent with its detrimental role in neurodegenerative diseases. Isobutyrylcarnitine (C4), a short-chain acylcarnitine, is the most abundant carnitine in the human body. Studies suggest that levels of isobutyryl-L-carnitine in blood or plasma decrease in traumatic brain injury patients [32,33], consistent with our finding that lower levels of isobutyrylcarnitine (C4) correlate with smaller cortical surface area. Bilirubin's antioxidant activity is particularly important for the brain, preventing excitotoxicity and neuronal death by clearing superoxide radicals during *N*-methyl-D-aspartate neurotransmission [34]. Similarly, our research correlates higher bilirubin levels with larger cortical surface area. Phospholipid 1-palmitoyl-2-oleoyl-gpc (16:0/18:1) also affects the cortical surface area by participating in the cell membrane structure and function. Choline phosphate is involved in phospholipid biosynthesis, impacting membrane stability and function. Butyric acid, a short-chain fatty acid, is metabolized in mitochondria through fatty acid metabolism as an energy source. Its role in the intestine and potential impact on overall health are supported by existing research, with lower butyric acid levels appearing correlated with multisystem atrophy [35]. Metabolomic pathway analysis reveals a significant association between butyric acid metabolism and brain atrophy, suggesting a potential beneficial role in neuroprotection and brain health via gut microbiota synthesis.

Cognitive performance, fluid intelligence scores, prospective memory outcomes, and reaction times are highly significant in studying cognitive functions, psychology, and neuroscience. Among these related CSF metabolites, acetylcarnitine (C2) participates in energy metabolism processes, potentially aiding in maintaining neuronal function and protecting the nervous system from damage [36]. Compounds such as Cys-gly, oxidized forms, and ascorbate (vitamin C) may benefit brain health through anti-

oxidative and anti-inflammatory processes [37,38]. 1-Methylnicotinamide, the main metabolite of nicotinamide, has been reported to be associated with neuroprotection and can improve cognitive deficits [39]. Consistent with these findings, our results indicate that 1-methylnicotinamide leads to changes in prospective memory results. In studies of neuroprotection and cognitive function, threonate has shown significant advantages in improving cognitive abilities and memory [40]. Similarly, our research findings indicate that an increase in threonate enhances cognitive performance and fluid intelligence scores. Hippurate levels decrease in cognitive impairment disorders [41], which aligns with our finding that it alters cognitive performance and fluid intelligence scores. Phenylacetylglutamine and *N*-acetylhistidine are metabolites of amino acids. Ribulonate participates in the metabolic pathway of ribose. N6-Succinyladenosine is an adenosine analogue that may be involved in protein synthesis or energy metabolism pathways. Cholesterol and sphingomyelin (d18:1/20:0, d16:1/22:0) are lipids involved in the cell membrane structure and signal transduction. 4-Chlorobenzoic acid is a chlorinated benzoic acid used in chemical research or pharmaceutical synthesis. Oxalate participates in calcium metabolism and stone formation. These compounds have shown certain correlations with cognitive functions related to the brain in our research. Notably, two compounds remained significantly associated with the reaction time after FDR correction, highlighting their potential biological relevance. *N*-Acetylglucosamine, a biologically significant monosaccharide, plays a critical role in cellular processes, and its post-translational modifications may influence neuronal function and contribute to the pathogenesis of neurodegenerative diseases [42]. Oxalate, a naturally occurring organic compound primarily derived from dietary sources such as vegetables, legumes, and nuts, exhibits strong chelating properties and potential neurotoxicity. Consequently, tight regulation of its concentration in CSF may be essential to prevent neuronal damage [43]. In line with these findings, our results demonstrate that both *N*-acetylglucosamine and oxalate are negatively associated with reaction time, suggesting a potential link between these metabolites and cognitive processing speed.

In our reverse analysis, we found that only cognitive performance leads to changes in CSF metabolites, whereas other factors such as cortical area, thickness, hippocampal volume, fluid intelligence scores, brain cognition, prospective memory outcomes, and reaction times did not affect CSF metabolites. These two metabolites are worth noting. Elevated levels of hippurate are associated with the risk of age-related cognitive impairment, possibly due to an increase in hippurate following a high polyphenol diet,

which enhances anti-inflammatory and antioxidative effects [41]. A high polyphenol diet appears to be the best dietary approach for preventing or delaying cognitive decline [44]. Additionally, from the perspective of gut microbiota, cognitive impairment significantly disrupts the composition and function of intestinal microbiota, where intestinal microbiota are involved in the synthesis and metabolism of hippurate [45]. Choline phosphate plays a crucial role in the synthesis of essential phospholipids for the human brain [46]. Higher levels of choline in brain metabolites are associated with the onset of psychiatric disorders [47]. Choline phosphate in the CSF is correlated with cognitive performance, though specific mechanisms require further research.

5 Advantages and limitations

To the best of our knowledge, this study represents the first comprehensive investigation of causal relationships between CSF metabolites and both cognitive function and structural brain atrophy. By employing a two-sample MR design (including bidirectional MR studies), we maximally avoided the pitfalls of reverse causation and residual confounding in observational studies. Second, this MR analysis extensively incorporates genetic variables, enabling us to make effective causal inferences. Finally, bidirectional MR design largely mitigates reverse causation and residual confounding factors [48]. However, several limitations need consideration in interpreting the results. First, the biological roles of certain identified metabolites and their associated pathways in cortical development and cognitive processes remain incompletely characterized, potentially constraining mechanistic interpretations of the MR findings. Second, as the analysis was restricted to European-ancestry populations, generalizability to other ethnic groups may be limited. Another limitation of the study is that the chemical properties of metabolites such as X-23308 and X-11612 remain unknown, which may affect the comprehensive interpretation of their roles and mechanisms.

6 Conclusion

In summary, this MR study establishes robust causal links between CSF metabolites and genetic determinants, while demonstrating their associations with cognitive function and structural brain metrics, including cortical and

hippocampal volumes. These findings collectively underscore the potential role of CSF metabolites in modulating neurocognitive health. Future studies should prioritize multi-ethnic cohorts and longitudinal designs to strengthen causal inference, thereby advancing biomarker discovery for early diagnosis and disease monitoring in neurodegenerative disorders.

Acknowledgments: None.

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

Author contributions: G.W.L. conceived the study and supervised research; Q.L. and Y.Q.M. designed the methodology; Q.L. and L.B.M. performed data analysis. All authors contributed to data interpretation. Q.L. and L.B.M. drafted the manuscript; G.W.L., T.Q.Q., and L.B.M. provided critical revisions. All authors reviewed the manuscript.

Conflict of interest: The authors declare no competing interests.

Data availability statement: The datasets generated or analyzed during this study are included in the supplementary material or the data repositories listed in the methods.

References

- [1] Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet*. 2020;396:413–46. doi: 10.1016/s0140-6736(20)30367-6.
- [2] Beason-Held LL, Goh JO, An Y, Kraut MA, O'Brien RJ, Ferrucci L, et al. Changes in brain function occur years before the onset of cognitive impairment. *J Neurosci*. 2013;33:18008–14. doi: 10.1523/jneurosci.1402-13.2013.
- [3] Jin Q, Ma RCW. Metabolomics in diabetes and diabetic complications: insights from epidemiological studies. *Cells*. 2023;10:2832. doi: 10.3390/cells10112832.
- [4] Oberbach A, Blüher M, Wirth H, Till H, Kovacs P, Kullnick Y, et al. Combined proteomic and metabolomic profiling of serum reveals association of the complement system with obesity and identifies novel markers of body fat mass changes. *J Proteome Res*. 2011;10:4769–88. doi: 10.1021/pr2005555.
- [5] Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature*. 2009;457:910–4. doi: 10.1038/nature07762.

- [6] Trushina E, Dutta T, Persson XM, Mielke MM, Petersen RC. Identification of altered metabolic pathways in plasma and CSF in mild cognitive impairment and Alzheimer's disease using metabolomics. *PLoS One*. 2013;8:e63644. doi: 10.1371/journal.pone.0063644.
- [7] Schwarz E, Bahn S. Biomarker discovery in psychiatric disorders. *Electrophoresis*. 2008;29:2884–90. doi: 10.1002/elps.200700710.
- [8] Olsson B, Lautner R, Andreasson U, Öhrfelt A, Portelius E, Björke M, et al. CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol*. 2016;15:673–84. doi: 10.1016/s1474-4422(16)00070-3.
- [9] Jiang YM, Long LL, Zhu XY, Zheng H, Fu X, Ou SY, et al. Evidence for altered hippocampal volume and brain metabolites in workers occupationally exposed to lead: a study by magnetic resonance imaging and (1)H magnetic resonance spectroscopy. *Toxicol Lett*. 2008;181:118–25. doi: 10.1016/j.toxlet.2008.07.009.
- [10] Meier TB, Drevets WC, Wurfel BE, Ford BN, Morris HM, Victor TA, et al. Relationship between neurotoxic kynurenine metabolites and reductions in right medial prefrontal cortical thickness in major depressive disorder. *Brain Behav Immun*. 2016;53:39–48. doi: 10.1016/j.bbi.2015.11.003.
- [11] Scott AS, Nashid HC, Usman AK, Beatriz P, Martin AS, Iris A, et al. Imaging patients with psychosis and a mouse model establishes a spreading pattern of hippocampal dysfunction and implicates glutamate as a driver. *Neuron*. 2013;78:81–93. doi: 10.1016/j.neuron.2013.02.011.
- [12] Peggy S, Fabiola DGM, Cristian P, Anna K. Mendelian randomization as an approach to assess causality using observational data. *J Am Soc Nephrol*. 2016;27:2257–69. doi: 10.1681/ASN.2016010098.
- [13] Jessen F, Amariglio RE, van Boxtel M, Breteler M, Ceccaldi M, Chételat G, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers Dement*. 2014;10:844–52. doi: 10.1016/j.jalz.2014.01.001.
- [14] Pini L, Pievani M, Bocchetta M, Altomare D, Bosco P, Cavedo E, et al. Brain atrophy in Alzheimer's disease and aging. *Ageing Res Rev*. 2016;30:25–48. doi: 10.1016/j.arr.2016.01.002.
- [15] Boef AG, Dekkers OM, le Cessie S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. *Int J Epidemiol*. 2015;44:496–511. doi: 10.1093/ije/dyv071.
- [16] 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. doi: 10.1038/s42003-020-01583-z.
- [17] Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37:658–65. doi: 10.1002/gepi.21758.
- [18] Gill D, Brewer CF, Monori G, Trégouët DA, Franceschini N, Giambartolomei C, et al. Effects of genetically determined iron status on risk of venous thromboembolism and carotid atherosclerotic disease: a Mendelian randomization study. *J Am Heart Assoc*. 2019;8:e012994. doi: 10.1161/jaha.119.012994.
- [19] Lee JJ, Wedow R, Okbay A, Kong E, Maghziyan O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*. 2018;50:1112–21. doi: 10.1038/s41588-018-0147-3.
- [20] Fawns-Ritchie C, Deary IJ. Reliability and validity of the UK Biobank cognitive tests. *PLoS One*. 2020;15:e0231627. doi: 10.1371/journal.pone.0231627.
- [21] Haworth S, Mitchell R, Corbin L, Wade KH, Dudding T, Budu-Aggrey A, et al. Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nat Commun*. 2019;10:333. doi: 10.1038/s41467-018-08219-1.
- [22] Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun*. 2018;9:2098. doi: 10.1038/s41467-018-04362-x.
- [23] Katrina LG, Neda J, Jodie NP, Lucía C-C, Janita B, Derrek PH, et al. The genetic architecture of the human cerebral cortex. *Science*. 2020;367:eaay6690. doi: 10.1126/science.aay6690.
- [24] Hibar DP, Adams HHH, Jahanshad N, Chauhan G, Stein JL, Hofer E, et al. Novel genetic loci associated with hippocampal volume. *Nat Commun*. 2017;8:13624. doi: 10.1038/ncomms13624.
- [25] Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol*. 2013;178:1177–84. doi: 10.1093/aje/kwt084.
- [26] Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32:377–89. doi: 10.1007/s10654-017-0255-x.
- [27] Cohen JF, Chalumeau M, Cohen R, Korevaar DA, Khoshnood B, Bossuyt PM. Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. *J Clin Epidemiol*. 2015;68:299–306. doi: 10.1016/j.jclinepi.2014.09.005.
- [28] Suzuki S, Koda Y, Saito T, Fujimoto K, Momozono A, Hayashi A, et al. Methionine sulfoxides in serum proteins as potential clinical biomarkers of oxidative stress. *Sci Rep*. 2016;6:38299. doi: 10.1038/srep38299.
- [29] Schöneich C. Methionine oxidation by reactive oxygen species: reaction mechanisms and relevance to Alzheimer's disease. *Biochim Biophys Acta*. 2005;1703:111–9. doi: 10.1016/j.bbapap.2004.09.009.
- [30] Anders MF, Linda M, Dominic H, Anders MD, Kristine, BW. What is normal in normal aging? Effects of aging, amyloid and Alzheimer's disease on the cerebral cortex and the hippocampus. *Prog Neurobiol*. 2014;118:1–20. doi: 10.1016/j.pneurobio.2014.02.0064.
- [31] Patriarca EJ, Cermola F, D'Aniello C, Fico A, Guardiola O, De Cesare D, et al. The Multifaceted roles of proline in cell behavior. *Front Cell Dev Biol*. 2021;9:728576. doi: 10.3389/fcell.2021.728576.
- [32] Makarova E, Makrecka-Kuka M, Vilks K, Volska K, Sevostjanovs E, Grinberga S, et al. Decreases in circulating concentrations of long-chain acylcarnitines and free fatty acids during the glucose tolerance test represent tissue-specific insulin sensitivity. *Front Endocrinol (Lausanne)*. 2019;10:870. doi: 10.3389/fendo.2019.00870.
- [33] Jeter CB, Hergenroeder GW, Ward 3rd NH, Moore AN, Dash PK. Human mild traumatic brain injury decreases circulating branched-chain amino acids and their metabolite levels. *J Neurotrauma*. 2013;30:671–9. doi: 10.1089/neu.2012.2491.
- [34] Vasavda C, Kothari R, Malla AP, Tokhunts R, Lin A, Ji M, et al. Bilirubin links heme metabolism to neuroprotection by scavenging superoxide. *Cell Chem Biol*. 2019;26:1450–60.e1457. doi: 10.1016/j.chembiol.2019.07.006.
- [35] Voigt RM, Wang Z, Brown JM, Engen PA, Naqib A, Goetz CG, et al. Gut microbial metabolites in Parkinson's disease: Association with lifestyle, disease characteristics, and treatment status. *Neurobiol Dis*. 2022;170:105780. doi: 10.1016/j.nbd.2022.105780.
- [36] Zanelli SA, Solenski NJ, Rosenthal RE, Fiskum G. Mechanisms of ischemic neuroprotection by acetyl-L-carnitine. *Ann N Y Acad Sci*. 2005;1053:153–61. doi: 10.1196/annals.1344.013.

- [37] Iantomasi T, Favilli F, Marraccini P, Stio M, Treves C, Quattrone A, et al. Age and GSH metabolism in rat cerebral cortex, as related to oxidative and energy parameters. *Mech Ageing Dev.* 1993;70:65–82. doi: 10.1016/0047-6374(93)90060-5.
- [38] Olajide OJ, Yawson EO, Gbadamosi IT, Arogundade TT, Lambe E, Obasi K, et al. Ascorbic acid ameliorates behavioural deficits and neuropathological alterations in rat model of Alzheimer's disease. *Environ Toxicol Pharmacol.* 2017;50:200–11. doi: 10.1016/j.etap.2017.02.010.
- [39] Mu RH, Tan YZ, Fu LL, Nazmul Islam M, Hu M, Hong H, et al. 1-Methylnicotinamide attenuates lipopolysaccharide-induced cognitive deficits via targeting neuroinflammation and neuronal apoptosis. *Int Immunopharmacol.* 2019;77:105918. doi: 10.1016/j.intimp.2019.105918.
- [40] Zhang C, Hu Q, Li S, Dai F, Qian W, Hewlings S, et al. A Magtein[®], magnesium L-threonate, -based formula improves brain cognitive functions in healthy chinese adults. *Nutrients.* 2022;14:5235. doi: 10.3390/nu14245235.
- [41] Angela G, Andrea T, Franca A, Antonio N, Beatrice P, Silvana P, et al. Insights about urinary hippuric and citric acid as biomarkers of fruit and vegetable intake in patients with kidney stones: The role of age and sex. *Nutrition.* 2018;50:1–7. doi: 10.1016/j.nut.2017.12.006.
- [42] Shafi R, Iyer SP, Ellies LG, O'Donnell N, Marek KW, Chui D, et al. The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc Natl Acad Sci U S A.* 2000;97:5735–9. doi: 10.1073/pnas.100471497.
- [43] Joshi S, Saylor BT, Wang W, Peck AB, Khan SR. Apocynin-treatment reverses hyperoxaluria induced changes in NADPH oxidase system expression in rat kidneys: a transcriptional study. *PLoS One.* 2012;7:e47738. doi: 10.1371/journal.pone.0047738.
- [44] Zhou Y, Wang J, Cao L, Shi M, Liu H, Zhao Y, et al. Fruit and vegetable consumption and cognitive disorders in older adults: a meta-analysis of observational studies. *Front Nutr.* 2022;9:871061. doi: 10.3389/fnut.2022.871061.
- [45] Xi J, Ding D, Zhu H, Wang R, Su F, Wu W, et al. Disturbed microbial ecology in Alzheimer's disease: evidence from the gut microbiota and fecal metabolome. *BMC Microbiol.* 2021;21:226. doi: 10.1186/s12866-021-02286-z.
- [46] Zeisel SH. Choline: an essential nutrient for humans. *Nutrition.* 2000;16:669–71. doi: 10.1016/s0899-9007(00)00349-x.
- [47] Byun MS, Choi JS, Yoo SY, Kang DH, Choi CH, Jang DP, et al. Depressive Symptoms and brain metabolite alterations in subjects at ultra-high risk for psychosis: a preliminary study. *Psychiatry Investig.* 2009;6:264–71. doi: 10.4306/pi.2009.6.4.264.
- [48] Bowden J, Holmes MV. Meta-analysis and mendelian randomization: a review. *Res Synth Methods.* 2019;10:486–96. doi: 10.1002/jrsm.1346.