ARTICLE OPEN

Check for updates

Circulating metabolome in relation to cognitive impairment: a community-based cohort of older adults

Yuhui Huang[1,8], Xuehui Sun[2,3,8], Qingxia Huang[2,4,8], Qiumin Huang[2,8], Xiao Chen[5], Xiaofeng Zhou[2], Hui Chen[1], Jie Shen[1], Mengyan Gao[1], Yiying Gong[1], Hui Zhang[2,3], Huiru Tang[2,4], Xiaofeng Wang[2,3], Xiaoyan Jiang[6,9✉], Yan Zheng[2,9✉] and Changzheng Yuan[1,7,9✉]

The role of circulating metabolome in cognitive impairment is inconclusive, and whether the associations are in the severity-dependent manner remains unclear. We aimed to identify plasma metabolites associated with cognitive impairment and evaluate the added predictive capacity of metabolite biomarkers on incident cognitive impairment beyond traditional risk factors. In the Rugao Longevity and Ageing Study (RuLAS), plasma metabolome was profiled by nuclear magnetic resonance spectroscopy. Participants were classified into the cognitively normal, moderately impaired, and severely impaired groups according to their performance in two objective cognitive tests. A two-step strategy of cross-sectional discovery followed by prospective validation was applied. In the discovery stage, we included 1643 participants (age: 78.9 ± 4.5 years) and conducted multinomial logistic regression. In the validation stage, we matched 68 incident cases of cognitive impairment (moderately-to-severely impaired) during the 2-year follow-up with 204 cognitively normal controls by age and sex at a 1:3 ratio, and conducted conditional logistic regression. We identified 28 out of 78 metabolites cross-sectionally related to severely impaired cognition, among which IDL particle number, ApoB in IDL, leucine, and valine were prospectively associated with 28%, 28%, 29%, and 33% lower risk of developing cognitive impairment, respectively. Incorporating 13 metabolite biomarkers selected through Lasso regression into the traditional risk factors-based prediction model substantially improved the ability to predict incident cognitive impairment (AUROC: 0.839 vs. 0.703, $P < 0.001$; AUPRC: 0.705 vs. 0.405, $P < 0.001$). This study identified specific plasma metabolites related to cognitive impairment. Incorporation of specific metabolites substantially improved the prediction performance for cognitive impairment.

Translational Psychiatry (2024)14:469 ; https://doi.org/10.1038/s41398-024-03147-9

INTRODUCTION

Cognitive impairment affects memory, learning, concentration, and decision-making abilities of individuals. With the global population ageing, the number of people with dementia and mild cognitive impairment (MCI) is steadily increasing [1, 2]. Due to the limited disease-modifying therapies for dementia, it is crucial to further understand the etiology of cognitive impairment and detect biomarkers contributing to the early identification of high-risk individuals [3, 4].

As the end-products of various metabolic processes, circulating metabolites reflect how biological systems respond to changes in intrinsic and extrinsic factors [5]. Therefore, circulating metabolome has the potential to help understand the etiology and discover biomarkers for cognitive impairment [6]. Prospective epidemiological studies have found that some metabolites were associated with onset of cognitive impairment. For example, several amino acids (e.g., glutamine, glutamic acid), organic acids (e.g., anthranilic acid, isocitrate), fatty acids (e.g., saturated fatty acid to total fatty acid ratio), and lipoprotein fractions ratios (e.g., cholesterol to total lipids ratio in very large VLDL) were related to higher risk of dementia, Alzheimer's disease (AD), or MCI [7–11]. Additionally, several metabolites were related to lower risk of dementia or AD, including amino acids (e.g., isoleucine, leucine, valine), organic acids (e.g., taurine), fatty acids (e.g., docosahexaenoic acid), and lipoprotein fractions ratios (e.g., triglycerides to total lipids ratio in very large VLDL) [7–9, 11, 12]. However, most of the aforementioned relationships were observed in only one study, suggesting that the evidence remains inconclusive at this time. Moreover, the majority of current studies were conducted in the Western population, which has distinct genetic and cultural backgrounds that could influence the circulating metabolome compared to the Asian population [13]. Furthermore, few studies have investigated whether these relationships were in the severity-dependent manner to cognitive impairment. Overall, further studies are required to elucidate the

[1]School of Public Health, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [2]State Key Laboratory of Genetic Engineering, School of Life Sciences and Human Phenome Institute, Fudan University, Shanghai, China. [3]Fudan University-the People's Hospital of Rugao Joint Research Institute of Longevity and Aging, Rugao, Jiangsu, China. [4]Zhangjiang Fudan International Innovation Center, Zhongshan Hospital, Fudan University, Shanghai, China. [5]Department of Nutrition and Food Hygiene, School of Public Health, Institute of Nutrition, Fudan University, Shanghai, China. [6]State Key Laboratory of Cardiology, Department of Pathology and Pathophysiology, School of Medicine, Tongji University, Shanghai, China. [7]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [8]These authors contributed equally: Yuhui Huang, Xuehui Sun, Qingxia Huang, Qiumin Huang. [9]These authors jointly supervised this work: Xiaoyan Jiang, Yan Zheng, Changzheng Yuan.
✉email: xiaoyanjiang001@163.com; yan_zheng@fudan.edu.cn; chy478@zju.edu.cn

Received: 28 April 2024 Revised: 30 September 2024 Accepted: 1 October 2024
Published online: 11 November 2024

associations of circulating metabolites with cognitive impairment, especially among Asian individuals.

In this study, we utilized data from a community-based cohort of older adults in China, with two objective cognitive tests used to classify participants into three groups in a severity-dependent manner. We aimed to examine the associations between plasma metabolites and risk of cognitive impairment, and evaluate the added predictive capacity of metabolite biomarkers on incident cognitive impairment beyond traditional risk factors in this high-risk population.

SUBJECTS AND METHODS

Study design and population

This study was based on the ageing arm of Rugao Longevity and Ageing Study (RuLAS), a community-based open cohort study [14–16]. Briefly, 1788 individuals aged 70–84 years were recruited at the baseline survey in 2014 (Wave 1) and new eligible participants were additionally recruited during the follow-up surveys in 2016, 2017, 2019, and 2021 (Waves 2–5). Socio-demographic, lifestyle, and other health-related information were collected through face-to-face interviews at each survey. RuLAS was ethically approved by the Human Ethics Committee of the School of Life Sciences of Fudan University. All participants provided signed informed consent.

In the present study, we applied a two-step strategy to identify cognitive impairment-related metabolites (Fig. 1), including a discovery phase of the cross-sectional analysis and a validation phase of the prospective analysis. The cross-sectional analysis was conducted at Wave 4 (year 2019), where plasma metabolome was first measured. After excluding 557 individuals without data on plasma metabolomics, objective cognitive tests, or education level, a total of 1643 participants were included in the cross-sectional analysis. In the prospective analysis (nested case-control design), we further excluded participants who already had cognitive impairment at Wave 4 and those who did not complete cognitive assessments at Wave 5 (year 2021). Of the remaining participants, we identified 68 incident cases of cognitive impairment and matched them with 204 randomly selected cognitively normal controls by age (±5 years) and sex at a matching ratio of 1:3 (Fig. S1) [17].

Plasma metabolome

Fasting plasma samples were collected by trained nurses at Wave 4. Plasma metabolome was profiled by high-throughput untargeted nuclear magnetic resonance (NMR) spectroscopy. All NMR spectra were acquired at 310 K on a Bruker Advance III HD 600 MHz NMR spectrometer equipped with a 5 mm BBI probe (Bruker Biospin, Germany). Detailed information has been published elsewhere [18, 19].

Plasma metabolites with >80% of missing values were removed before analysis, and for the remaining metabolites, missing values were imputed using half of the lowest detected values. To approximate the normal distribution, the rank-based inverse normal transformation was applied to the metabolomics data [20]. In total, 9 total fractions, 8 VLDL fractions, 8 LDL fractions, 8 IDL fractions, 8 HDL fractions, 9 organic acids, 15 amino acids, 6 fatty acids, 2 inflammation markers, 5 other low-molecular-weight metabolites, 110 lipoprotein subfractions, and 123 lipoprotein fractions ratios were included in the analysis. Lipoprotein subfractions and lipoprotein fractions ratios were recognized as exploratory measures and the other 78 metabolites were considered as the main measures in the analyses.

To partially validate the accuracy of NMR spectroscopy quantitative results, spearman correlation coefficients of 6 metabolites simultaneously measured by both NMR spectroscopy in the plasma samples and clinical chemistry in the serum samples were calculated. Among them, 4 measures (triglycerides, cholesterol, LDL cholesterol, and HDL cholesterol) were highly correlated, with correlation coefficients ranging from 0.73 to 0.93 (Fig. S2). Correlation coefficients for glucose and creatinine were slightly lower but remained higher than 0.60.

Cognitive impairment

Two objective cognitive tests commonly used and well-validated in the Chinese population, the Hasegawa Dementia Scale (HDS) and Mini-Mental State Examination (MMSE), were interviewer-administered to evaluate the cognitive function of participants [21–23]. Apart from the shared cognitive domains of orientation, memory, and attention, HDS focused on the domains of abstract and naming while MMSE focused on the domains of language and visuospatial ability [24]. According to the education-specific cutoffs of HDS, cognitive impairment was defined as follows: ≤15 for illiterate, ≤19 for primary school, and ≤23 for middle school and above [21, 25]. According to the education-specific cutoffs of MMSE, cognitive impairment was defined as follows: ≤17 for illiterate, ≤20 for primary school, and ≤24 for middle school and above [22]. Participants identified as cognitively impaired by both HDS and MMSE may have deficits in more cognitive domains and exhibit more severe cognitive impairment compared to those identified by either HDS or MMSE alone [26]. Therefore, we classified participants into three groups: cognitively normal (not impaired as judged by HDS and MMSE), moderately impaired (impaired as judged by either HDS or MMSE), and severely impaired (impaired as judged by both HDS and MMSE).

Covariates

The following baseline characteristics (from Wave 4) were identified as covariates, mainly involving sociodemographic factors, lifestyles, health conditions, and *APOE* genotype. Sociodemographic factors included age (years), sex (male or female), education level (illiterate, primary school, or middle school and above), and marital status (married or others). Lifestyles comprised smoking status (never, former, or current), drinking status (never, former, or current), physical activity (>3 times/week or ≤3 times/week), and sleep quality (scores of Pittsburgh Sleep Quality Index [PSQI] ≤5, good or poor) [21]. Health conditions consisted of BMI (<18.5, 18.5–23.9, 24.0–27.9, or ≥28.0 kg/m^2), depressive symptoms (scores of Geriatric Depression Scale-15 [GDS-15] ≥5, yes or no) [27], history of hypertension, diabetes, cardiovascular disease, and cancer, and self-reported use of medications for hypertension, diabetes, cardiovascular disease, and high cholesterol. Participants with self-reported physician diagnosis, use of medications, systolic blood pressure ≥140 mmHg, or diastolic blood pressure ≥90 mmHg were identified as having hypertension. Those with self-reported physician diagnosis, use of medications, fasting blood glucose levels ≥126 mg/dL, or non-fasting blood glucose levels ≥200 mg/dL were identified as having diabetes. Those with self-reported physician diagnosis or use of medications were identified as having cardiovascular disease. Those with self-reported physician diagnosis were identified as having cancer. *APOE* rs429358 and rs7412 polymorphisms were used to determine the *APOE* genotype and participants with at least one *APOE* ε4 allele (ε2/ε4, ε3/ε4, or ε4/ε4) were identified as *APOE* ε4 carriers [28].

Statistical analyses

Baseline characteristics of the study participants grouped by different cognitive statuses were compared using Pearson's chi-square test (χ^2) for categorical variables and one-way analysis of variance (ANOVA) or Student's t-test for continuous variables. Missing values for categorical covariates were imputed to a separate missing category. Univariate fold change was calculated as the log2 transformation of the ratio between mean values of metabolites in the moderately impaired/severely impaired group relative to the cognitively normal group.

To identify cognitive impairment-related metabolites, we applied a two-step strategy consisting of an initial discovery phase based on the cross-sectional analysis and a further validation phase based on the prospective analysis [29]. In the discovery stage, multinomial logistic regression was performed to examine associations of per 1-SD increment in main measures with moderately impaired and severely impaired compared to cognitively normal (reference group), with adjustment for age, sex, education level, and *APOE* ε4 carrier status. The Benjamini-Hochberg method was used for multiple testing adjustment, and metabolites with false discovery rate (FDR) adjusted P values < 0.05 were considered as cognitive impairment-related metabolite candidates. Protective metabolite scores of severely impaired were calculated as the sum of all or subclasses of metabolite candidates, weighted by the inverse values of coefficients from the regression models abovementioned. For example, the amino acid score was calculated as the sum of leucine and valine, each weighted by the inverse values of respective coefficients of association with severely impaired cognition. Further, we validated the metabolite candidates and metabolite scores using the nested case-control design (68 incident cases and 204 healthy controls). In the validation stage, we combined participants with newly developed moderately and severely impaired cognition during the 2-year follow-up as the incident cognitive impairment group. Conditional logistic regression adjusting for age, education level, and *APOE* ε4 carrier status was used to investigate associations of metabolite candidates and metabolite scores with incident cognitive impairment. Metabolite candidates with raw P values < 0.05 in 1-SD increment were considered to be successfully validated and named as cognitive impairment-related metabolites. Restricted cubic spline (RCS)

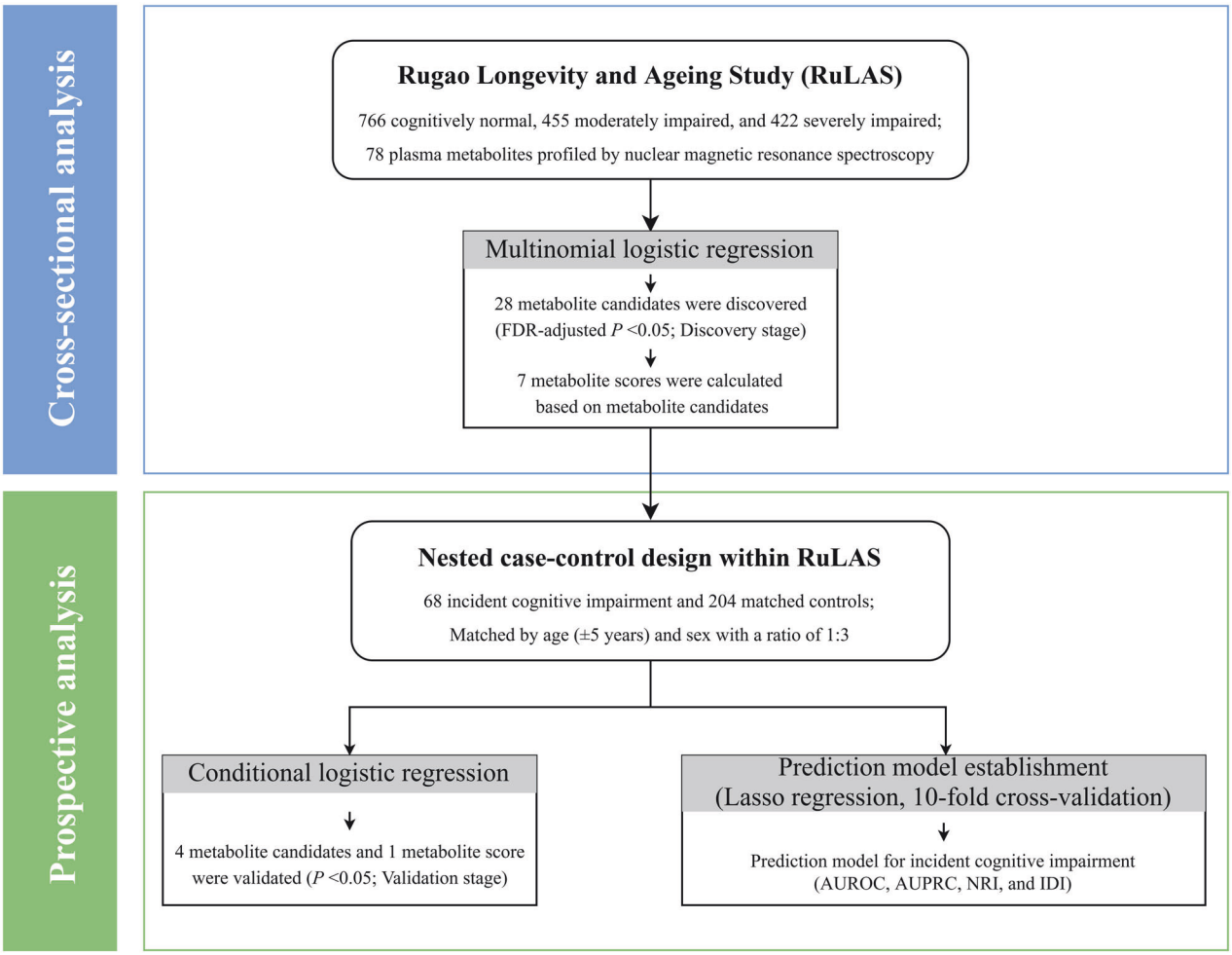

Fig. 1 Study design. To identify cognitive impairment-related metabolites, we applied a two-step strategy, including a discovery phase of the cross-sectional analysis and a validation phase of the prospective analysis. In the community-based Rugao Longevity and Ageing Study (RuLAS), plasma metabolome was profiled by nuclear magnetic resonance (NMR) spectroscopy. Participants were classified into the cognitively normal, moderately impaired, and severely impaired groups according to their performance in two objective cognitive tests. In the discovery stage, we used multinomial logistic regression to identify cognitive impairment-related metabolite candidates (FDR-adjusted $P < 0.05$), and metabolite scores were then calculated based on all or subclasses of metabolite candidates. In the validation stage, we matched 68 incident cases of cognitive impairment (moderately-to-severely impaired) during the 2-year follow-up with 204 cognitively normal controls by age (±5 years) and sex at a 1:3 ratio, and used conditional logistic regression to validate metabolite candidates and metabolite scores. Besides, we established the prediction models for incident cognitive impairment using Lasso regression with 10-fold cross-validation. AUROC Area under the receiver operating characteristic curve, AUPRC Area under the precision-recall curve, NRI Net reclassification index, IDI Integrated discrimination improvement.

was employed to explore potential non-linear relationships. We conducted the subgroup analysis stratified by age (<80 or ≥80), sex (male or female), and *APOE* ε4 carrier status (carrier or non-carrier). To test the robustness of results, we performed several sensitivity analyses. 1) We additionally adjusted for marital status, smoking status, drinking status, physical activity, sleep quality, BMI, and depressive symptoms. 2) We further adjusted for hypertension, diabetes, cardiovascular disease, cancer, antihypertensive medication, antidiabetic medication, medication for cardiovascular disease, cholesterol-lowering medication, and baseline cognitive function (in prospective analysis only). 3) Main measures and exploratory measures were all included in the cross-sectional analysis and the FDR adjustment was also applied.

Metabolite set enrichment analysis was performed to identify cognitive impairment-related metabolic pathways using the web-based MetaboAnalyst 5.0 platform [30]. Cognitive impairment-related metabolites that could also be matched within the Kyoto Encyclopedia of Genes and Genomes (KEGG) database were used to reveal altered metabolic pathways. In addition, principal component analysis (PCA) and partial least square discriminant analysis (PLS-DA) with 10-fold cross-validation were conducted to assess the ability of all of main measures in distinguishing incident cognitive impairment from cognitively normal [20]. The statistical significance of PLS-DA model was tested with the permutation test. Considering the feature of high correlations across the metabolomics data [31], we further used the least absolute shrinkage and selection operator (Lasso) regression with 10-fold cross-validation to select metabolite biomarkers from main measures [32]. According to the calculated sample size required for developing prediction models, we included all participants to develop the models and evaluate the model performance with 10-fold cross-validation to avoid risk of overfitting [33]. To evaluate the added predictive performance of selected metabolite biomarkers on incident cognitive impairment beyond traditional risk factors, we established three prediction models. Basic model incorporated traditional risk factors including age, sex, education level, marital status, smoking status, drinking status, physical activity, sleep quality, BMI, depressive symptoms, and *APOE* ε4 carrier status. Metabolites model included the selected metabolite biomarkers. Combined model integrated predictors from both the basic model and metabolites model. The DeLong test was used to compare the area under the receiver operating characteristic curve (AUROC) of combined model and basic model, while the bootstrap-based method was utilized to compare the area under the precision-recall curve (AUPRC) [20, 32]. Net reclassification index (NRI) and integrated discrimination improvement (IDI) of the combined model compared with basic model were also evaluated. Besides, performance metrics (i.e.,

Table 1. Baseline characteristics of participants included in the cross-sectional analysis ($N = 1643$).

Characteristic	Total	Cognitively normal	Moderately impaired	Severely impaired	*P* value
N	1643	766	455	422	
Age	78.9 ± 4.5	77.6 ± 4.1	79.4 ± 4.2	80.7 ± 4.6	<0.001
Sex					<0.001
Male	747 (45.5)	460 (60.1)	187 (41.1)	100 (23.7)	
Female	896 (54.5)	306 (39.9)	268 (58.9)	322 (76.3)	
Education level					<0.001
Illiterate	747 (45.5)	243 (31.7)	235 (51.6)	269 (63.7)	
Primary school	499 (30.4)	284 (37.1)	116 (25.5)	99 (23.5)	
Middle school and above	397 (24.2)	239 (31.2)	104 (22.9)	54 (12.8)	
Marital status					<0.001
Married	1047 (63.7)	542 (70.8)	288 (63.3)	217 (51.4)	
Others	596 (36.3)	224 (29.2)	167 (36.7)	205 (48.6)	
Drinking status					<0.001
Never	1059 (64.5)	444 (58.0)	302 (66.4)	313 (74.2)	
Former	134 (8.2)	74 (9.7)	39 (8.6)	21 (5.0)	
Current	450 (27.4)	248 (32.4)	114 (25.1)	88 (20.9)	
Smoking status					<0.001
Never	1236 (75.2)	519 (67.8)	350 (76.9)	367 (87.0)	
Former	149 (9.1)	90 (11.7)	37 (8.1)	22 (5.2)	
Current	258 (15.7)	157 (20.5)	68 (14.9)	33 (7.8)	
Physical activity					<0.001
>3 times/week	612 (38.1)	342 (45.0)	153 (34.2)	117 (29.3)	
≤3 times/week	994 (61.9)	418 (55.0)	294 (65.8)	282 (70.7)	
Sleep quality[a]					0.020
Good	809 (59.7)	405 (63.6)	212 (55.4)	192 (57.3)	
Poor	546 (40.3)	232 (36.4)	171 (44.6)	143 (42.7)	
BMI (kg/m^2)					<0.001
<18.5	97 (6.1)	29 (3.8)	27 (6.1)	41 (10.4)	
18.5–23.9	770 (48.4)	361 (47.9)	215 (48.5)	194 (49.1)	
24.0–27.9	544 (34.2)	284 (37.7)	140 (31.6)	120 (30.4)	
≥28.0	181 (11.4)	80 (10.6)	61 (13.8)	40 (10.1)	
Depressive symptoms[b]	223 (13.6)	80 (10.4)	63 (13.8)	80 (19.0)	<0.001
Comorbidity					
Hypertension	1176 (71.6)	563 (73.5)	317 (69.7)	296 (70.1)	0.268
Diabetes	238 (14.5)	106 (13.8)	67 (14.7)	65 (15.4)	0.753
Cardiovascular disease	324 (19.7)	148 (19.3)	91 (20.0)	85 (20.1)	0.929
Cancer	36 (2.2)	14 (1.8)	8 (1.8)	14 (3.3)	0.186
APOE ε4 carrier[c]					0.065
Yes	296 (18.0)	154 (20.1)	80 (17.6)	62 (14.7)	
No	1345 (82.0)	611 (79.9)	375 (82.4)	359 (85.3)	
Cognitive function scores					
HDS	21.56 ± 6.45	26.01 ± 4.48	20.91 ± 4.56	14.20 ± 3.50	<0.001
MMSE	19.71 ± 6.58	24.91 ± 3.48	17.20 ± 4.48	12.96 ± 4.87	<0.001

Values were mean ± sd for continuous variables and *n* (%) for categorical variables. *P* values were derived using Pearson's chi-square test (χ^2) for categorical variables and one-way analysis of variance for continuous variables.

[a]Individuals with PSQI scores ≤5 were regarded as having good sleep quality.

[b]Individuals with GDS-15 scores ≥5 were regarded as having depressive symptoms.

[c]Individuals with at least one *APOE* ε4 allele were regarded as *APOE* ε4 carriers.

BMI Body mass index, *APOE* ε4 Apolipoprotein E ε4, *HDS* Hasegawa Dementia Scale, *MMSE* Mini-Mental State Examination, *PSQI* Pittsburgh Sleep Quality Index, *GDS* Geriatric Depression Scale.

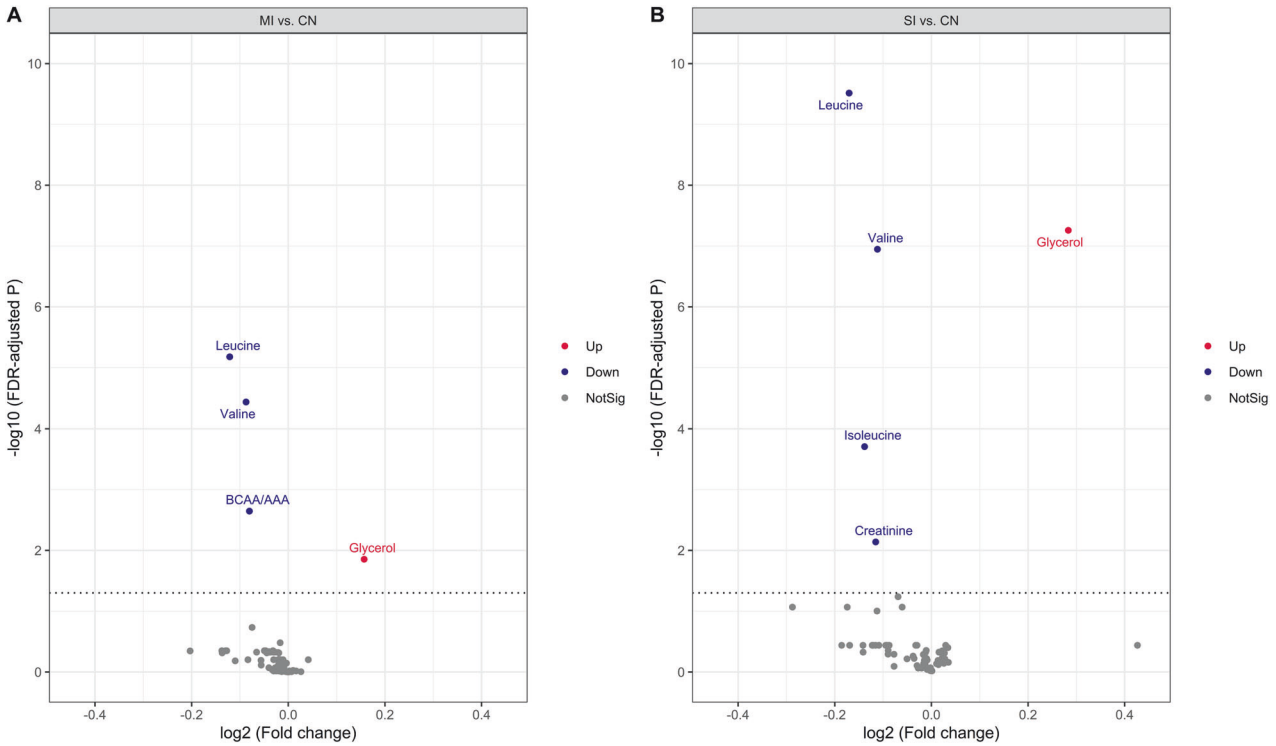

Fig. 2 Volcano plot showing the fold change of main metabolite measures between moderately impaired and severely impaired with cognitively normal. A Fold change between moderately impaired and cognitively normal. **B** Fold change between severely impaired and cognitively normal. The horizontal dashed line indicated FDR-adjusted *P* value of 0.05. MI Moderately impaired, SI Severely impaired, CN Cognitively normal, BCAA Branched-chain amino acid (sum of leucine, valine, and isoleucine), AAA Aromatic amino acid (sum of phenylalanine, tyrosine, and tryptophan).

sensitivity, specificity, AUROC, precision, recall, F1-score, AUPRC) of the combined model in the all participants and subgroups stratified by age, sex, and *APOE* ε4 carrier status were reported. All the statistical analyses except for the metabolite set enrichment analysis were performed using R (version 4.0.5).

RESULTS

Participant characteristics and metabolite-cognitive impairment associations in the discovery stage

Of the 1643 participants (age: 78.9 ± 4.5 years; female: 54.5%) included in the discovery stage (cross-sectional analysis), 46.6%, 27.7%, and 25.7% of them were cognitively normal, moderately impaired, and severely impaired, respectively (Table 1). A significant trend was observed in objective cognitive test scores across three groups, with mean HDS scores of 26.01 vs. 20.91 vs. 14.20 and mean MMSE scores of 24.91 vs. 17.20 vs. 12.96. Besides, participants in the severely impaired group were more likely to be older, female, and underweight, have lower education levels and physical activity levels, and have depressive symptoms. They were also less likely to be married, current drinkers, and current smokers.

Compared with cognitively normal individuals (Fig. 2), those with moderately impaired cognition had significantly higher circulating levels of glycerol and lower levels of leucine, valine, and branched-chain amino acid (leucine, valine, and isoleucine) to aromatic amino acid (phenylalanine, tyrosine, and tryptophan) ratio. Similarly, compared with cognitively normal individuals (Fig. 2), those with severely impaired cognition showed significantly higher circulating levels of glycerol and lower levels of leucine, valine, isoleucine, and creatinine. Notably, circulating levels of leucine, valine, and glycerol potentially altered in a severity-dependent manner across groups. When extending the analysis to both main measures and exploratory measures, leucine and valine remained statistically significant (FDR-adjusted $P < 0.05$), and another 11 metabolites from lipoprotein subfractions and lipoprotein fractions ratios were found to change in individuals with severely impaired cognition (Fig. S3).

In the multinomial logistic regression, 27 metabolites (3 total fractions, 8 VLDL fractions, 8 IDL fractions, 2 amino acids, 4 fatty acids, and 2 inflammation markers) were associated with lower odds of severely impaired cognition compared to cognitively normal, with ORs ranging from 0.75 to 0.84 (Fig. 3A, B; Table S1). In contrast, saturated fatty acid to total fatty acid ratio was related to higher odds (OR: 1.28, 1.11–1.47). In addition, 7 metabolite scores derived from all or subclasses of metabolite candidates were associated with lower odds of severely impaired cognition where the composite metabolite score and IDL fraction score were related to 28% (17%–37%) and 31% (20%–40%) decreased odds (Fig. 3B). When comparing moderately impaired cognition with cognitively normal, associations showed similar patterns but were attenuated to be statistically non-significant. After adjusting for additional covariates, results did not change materially (Table S2). When extending the analysis to both main measures and exploratory measures, 28 metabolite candidates remained statistically significant (FDR-adjusted $P < 0.05$), and another 48 metabolites from lipoprotein subfractions and lipoprotein fractions ratios were found to be associated with severely impaired cognition (Table S3).

Participant characteristics and metabolite-cognitive impairment associations in the validation stage

A total of 272 participants (age: 78.4 ± 3.7 years; female: 51.5%) without cognitive impairment at baseline were included in the validation stage (prospective analysis; Table S4). In the conditional logistic regression, associations of most metabolite candidates and metabolite scores with cognitive impairment showed the same directions as in the discovery stage, though the corresponding strengths were attenuated (Table 2). Among them, per SD

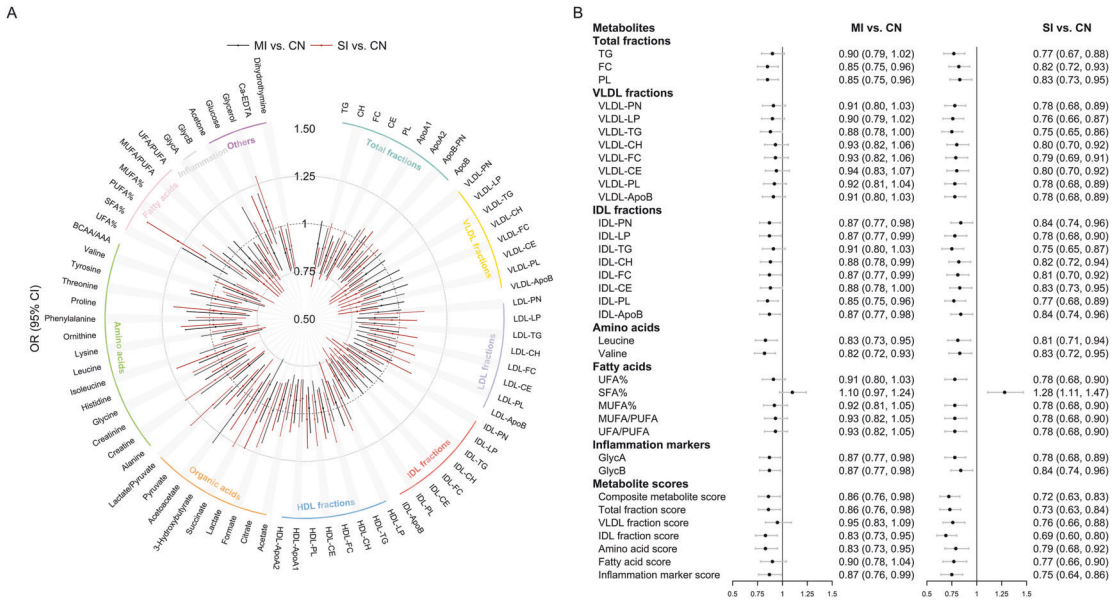

Fig. 3 Cross-sectional associations of main metabolite measures with moderately impaired and severely impaired. A Circular plot showing associations of main measures with moderately impaired (in dark) and severely impaired (in red) compared to cognitively normal. **B** Forest plot showing associations of metabolite candidates (FDR-adjusted $P < 0.05$) and metabolite scores with moderately impaired and severely impaired compared to cognitively normal. Values were ORs (95% CI) derived from multinomial logistic regression models adjusting for age, sex, education level, and *APOE* ε4 carrier status. Main measures with FDR-adjusted $P < 0.05$ were identified as metabolite candidates and metabolite scores were then calculated based on all or subclasses of metabolite candidates. MI Moderately impaired, SI Severely impaired, CN Cognitively normal, VLDL Very Low Density Lipoprotein, LDL Low Density Lipoprotein, IDL Intermediate Density Lipoprotein, HDL High Density Lipoprotein, TG Triglycerides, CH Cholesterol, FC Free Cholesterol, CE Cholesterol esters, PL Phospholipids, ApoA1 Apolipoprotein A1, ApoA2 Apolipoprotein A2, ApoB Apolipoprotein B100, PN Particle Number, LP Total lipids (sum of triglycerides, cholesterol, and phospholipids), BCAA Branched-chain amino acid (sum of leucine, valine, and isoleucine), AAA Aromatic amino acid (sum of phenylalanine, tyrosine, and tryptophan), UFA Unsaturated fatty acid, UFA% Unsaturated fatty acid to total fatty acid ratio, SFA Saturated fatty acid, SFA% Saturated fatty acid to total fatty acid ratio, PUFA Polyunsaturated fatty acid, PUFA% Polyunsaturated fatty acid to total fatty acid ratio, MUFA Monounsaturated fatty acid, MUFA% Monounsaturated fatty acid to total fatty acid ratio, GlycA N-acetylglucosamine/N-acetylgalactosamine-glycoproteins, GlycB N-acetylneuraminoyl-glycoproteins, EDTA Ethylene diamine tetraacetic acid.

increments of IDL particle number (OR: 0.72, 0.53–0.98), ApoB in IDL (OR: 0.72, 0.53–0.98), leucine (OR: 0.71, 0.51–0.99), valine (OR: 0.67, 0.48–0.93), and amino acid score (OR: 0.67, 0.48–0.94) were significantly associated with lower risk of cognitive impairment. In addition, several metabolites in the total fractions (i.e., free cholesterol [OR: 0.35, 0.15–0.82]), IDL fractions (i.e., IDL free cholesterol [OR: 0.20, 0.07–0.58], IDL cholesterol ester [OR: 0.26, 0.10–0.70]), and inflammation markers (i.e., GlycA [OR: 0.30, 0.12–0.74]) were also related to lower risk of cognitive impairment when comparing the third to first quartile. A significant non-linear association between GlycA levels and risk of cognitive impairment (P for non-linearity <0.05) was found (Fig. S4). After adjusting for additional covariates, results did not materially change (Table S5). Although no modification effects were observed in the subgroup analysis (all P for interaction >0.1), associations between IDL particle number and ApoB in IDL with cognitive impairment remained significant among those with age under 80 years and *APOE* ε4 non-carriers (Table S6). Similarly, the inverse relationship between valine and cognitive impairment persisted among females while relationships of leucine and valine remained significant among *APOE* ε4 non-carriers.

Metabolite set enrichment analysis and prediction models

We entered 2 metabolites (leucine and valine) into the metabolite set enrichment analysis and identified 3 enriched metabolic pathways, including valine, leucine, and isoleucine biosynthesis; valine, leucine, and isoleucine degradation; and aminoacyl-tRNA biosynthesis (Fig. S5). PCA demonstrated the poor discrimination ability (Fig. 4A) while PLS-DA indicated that the main measures profiles of individuals with incident cognitive impairment were significantly different from those of cognitively normal controls (Fig. 4B). Due to the high correlations across metabolites, 13 metabolite biomarkers (including IDL particle number, acetate, 3-hydroxybutyrate, acetoacetate, pyruvate, creatinine, histidine, tyrosine, valine, polyunsaturated fatty acid to total fatty acid ratio, acetone, glycerol, and Ca-EDTA) were selected from main measures using Lasso regression. Incorporating them into the basic model (traditional risk factors-based) significantly improved the predictability of incident cognitive impairment ($P < 0.001$ comparing AUROC of basic model [0.703, 0.632–0.774] with AUROC of combined model [0.839, 0.782–0.897], Fig. 4C; $P < 0.001$ comparing AUPRC of basic model [0.405, 0.308–0.529] with AUPRC of combined model [0.705, 0.594–0.811], Fig. 4D). Furthermore, with adding the selected metabolite biomarkers, the categorical NRI, continuous NRI, relative IDI, and absolute IDI were 0.412 (0.263–0.541), 0.735 (0.531–0.943), 2.538 (1.618–4.244), and 0.228 (0.169–0.286), respectively (data not shown). Performance metrics of the combined model were similar across subgroups (i.e., AUROC: 0.852 in males vs. 0.826 in females; AUPRC: 0.710 in males vs. 0.715 in females, Table S7).

DISCUSSION

In this community-based cohort of older adults in China, we identified 2 branched-chain amino acids (leucine and valine) and 2 IDL fractions (IDL particle number and ApoB in IDL) that were inversely associated with risk of cognitive impairment. Besides, 3 metabolic pathways (valine, leucine, and isoleucine biosynthesis; valine, leucine, and isoleucine degradation; and aminoacyl-tRNA biosynthesis) were found to be potentially related to cognitive impairment. The utilization of selected metabolite biomarkers significantly improved the prediction ability of incident cognitive impairment beyond traditional risk factors.

Table 2. Prospective associations of metabolite candidates and metabolite scores with cognitive impairment.

Metabolites	Per 1-SD increment	Quartile			
		Q1	Q2	Q3	Q4
Total fractions					
TG	0.93 (0.67, 1.29)	1 (ref.)	0.60 (0.27, 1.34)	0.60 (0.27, 1.37)	0.73 (0.32, 1.68)
FC	0.89 (0.64, 1.23)	1 (ref.)	0.54 (0.24, 1.22)	**0.35 (0.15, 0.82)**	0.70 (0.30, 1.64)
PL	0.90 (0.65, 1.24)	1 (ref.)	0.91 (0.41, 2.02)	0.43 (0.18, 1.03)	0.77 (0.32, 1.87)
VLDL fractions					
VLDL-PN	0.95 (0.68, 1.32)	1 (ref.)	0.57 (0.25, 1.31)	0.56 (0.25, 1.29)	0.85 (0.37, 1.98)
VLDL-LP	0.96 (0.69, 1.33)	1 (ref.)	0.63 (0.27, 1.47)	0.81 (0.36, 1.82)	0.83 (0.36, 1.89)
VLDL-TG	0.95 (0.69, 1.32)	1 (ref.)	0.61 (0.26, 1.39)	0.63 (0.27, 1.46)	1.02 (0.45, 2.29)
VLDL-CH	0.94 (0.69, 1.29)	1 (ref.)	0.88 (0.40, 1.95)	0.91 (0.40, 2.09)	1.04 (0.45, 2.40)
VLDL-FC	0.98 (0.71, 1.36)	1 (ref.)	0.69 (0.31, 1.53)	0.83 (0.37, 1.88)	0.94 (0.41, 2.16)
VLDL-CE	0.92 (0.67, 1.25)	1 (ref.)	0.73 (0.32, 1.63)	0.70 (0.30, 1.63)	1.11 (0.49, 2.50)
VLDL-PL	1.00 (0.73, 1.38)	1 (ref.)	0.74 (0.33, 1.65)	0.92 (0.40, 2.10)	0.90 (0.39, 2.06)
VLDL-ApoB	0.95 (0.68, 1.32)	1 (ref.)	0.57 (0.25, 1.31)	0.56 (0.25, 1.29)	0.85 (0.37, 1.98)
IDL fractions					
IDL-PN	**0.72 (0.53, 0.98)**	1 (ref.)	0.67 (0.30, 1.49)	**0.38 (0.16, 0.88)**	0.46 (0.20, 1.06)
IDL-LP	0.84 (0.61, 1.15)	1 (ref.)	0.85 (0.40, 1.80)	**0.40 (0.16, 0.98)**	0.83 (0.36, 1.93)
IDL-TG	0.95 (0.68, 1.31)	1 (ref.)	0.76 (0.34, 1.68)	0.65 (0.28, 1.50)	0.91 (0.39, 2.11)
IDL-CH	0.74 (0.53, 1.02)	1 (ref.)	0.89 (0.42, 1.88)	**0.34 (0.14, 0.86)**	0.60 (0.26, 1.37)
IDL-FC	0.76 (0.55, 1.05)	1 (ref.)	0.93 (0.43, 2.02)	**0.20 (0.07, 0.58)**	0.84 (0.36, 1.97)
IDL-CE	0.73 (0.53, 1.01)	1 (ref.)	0.98 (0.47, 2.05)	**0.26 (0.10, 0.70)**	0.59 (0.26, 1.38)
IDL-PL	0.86 (0.63, 1.18)	1 (ref.)	1.13 (0.53, 2.39)	**0.29 (0.11, 0.77)**	1.09 (0.47, 2.53)
IDL-ApoB	**0.72 (0.53, 0.98)**	1 (ref.)	0.67 (0.30, 1.49)	**0.38 (0.16, 0.88)**	0.46 (0.20, 1.06)
Amino acids					
Leucine	**0.71 (0.51, 0.99)**	1 (ref.)	0.89 (0.42, 1.88)	0.69 (0.30, 1.60)	0.45 (0.18, 1.11)
Valine	**0.67 (0.48, 0.93)**	1 (ref.)	1.50 (0.71, 3.18)	0.73 (0.31, 1.73)	0.60 (0.26, 1.37)
Fatty acids					
UFA%	0.91 (0.66, 1.25)	1 (ref.)	0.83 (0.39, 1.81)	0.59 (0.25, 1.37)	0.90 (0.39, 2.04)
SFA%	1.10 (0.80, 1.52)	1 (ref.)	0.66 (0.28, 1.54)	0.93 (0.41, 2.12)	1.11 (0.49, 2.53)
MUFA%	1.06 (0.77, 1.46)	1 (ref.)	1.82 (0.81, 4.08)	0.96 (0.40, 2.30)	1.57 (0.63, 3.89)
MUFA/PUFA	1.10 (0.80, 1.51)	1 (ref.)	1.74 (0.78, 3.88)	0.77 (0.32, 1.86)	1.41 (0.60, 3.35)
UFA/PUFA	1.10 (0.80, 1.51)	1 (ref.)	1.74 (0.78, 3.88)	0.77 (0.32, 1.86)	1.41 (0.60, 3.35)
Inflammation markers					
GlycA	0.82 (0.60, 1.12)	1 (ref.)	0.69 (0.31, 1.53)	**0.30 (0.12, 0.74)**	0.67 (0.28, 1.60)
GlycB	0.76 (0.57, 1.02)	1 (ref.)	0.50 (0.23, 1.12)	**0.42 (0.19, 0.94)**	0.68 (0.30, 1.53)
Metabolite scores					
Composite metabolite score	0.85 (0.62, 1.17)	1 (ref.)	0.56 (0.24, 1.28)	0.64 (0.28, 1.45)	1.00 (0.44, 2.26)
Total fraction score	0.76 (0.54, 1.07)	1 (ref.)	0.56 (0.24, 1.27)	0.44 (0.19, 1.03)	0.65 (0.27, 1.58)
VLDL fraction score	0.99 (0.71, 1.39)	1 (ref.)	0.80 (0.35, 1.80)	0.83 (0.37, 1.88)	1.02 (0.44, 2.34)
IDL fraction score	0.93 (0.67, 1.30)	1 (ref.)	0.99 (0.45, 2.17)	0.75 (0.31, 1.80)	1.16 (0.50, 2.71)
Amino acid score	**0.67 (0.48, 0.94)**	1 (ref.)	0.89 (0.42, 1.88)	0.64 (0.30, 1.39)	0.43 (0.18, 1.06)
Fatty acid score	0.96 (0.67, 1.37)	1 (ref.)	0.94 (0.35, 2.54)	0.68 (0.32, 1.46)	0.81 (0.37, 1.80)
Inflammation marker score	0.80 (0.57, 1.11)	1 (ref.)	**0.27 (0.11, 0.65)**	0.63 (0.31, 1.28)	0.30 (0.07, 1.23)

In the prospective analysis, we matched 68 incident casesof cognitive impairment (moderately-to-severely impaired) during the 2-year follow-up with 204 cognitively normal controls by age and sex at a 1:3 ratio. Values were ORs (95% CI) derived from conditional logistic regression models adjusting for age, education level, and *APOE* ε4 carrier status.

Bold indicated raw $P < 0.05$.

VLDL Very Low Density Lipoprotein, IDL Intermediate Density Lipoprotein, TG Triglycerides, CH Cholesterol, FC Free Cholesterol, CE Cholesterol esters, PL Phospholipids, ApoB Apolipoprotein B100, PN Particle Number, LP Total lipids (sum of triglycerides, cholesterol, and phospholipids), UFA Unsaturated fatty acid, UFA% Unsaturated fatty acid to total fatty acid ratio, SFA Saturated fatty acid, SFA% Saturated fatty acid to total fatty acid ratio, MUFA Monounsaturated fatty acid, MUFA% Monounsaturated fatty acid to total fatty acid ratio, PUFA Polyunsaturated fatty acid, GlycA N-acetylglucosamine/N-acetylgalactosamine-glycoproteins, GlycB N-acetylneuraminoyl-glycoproteins.

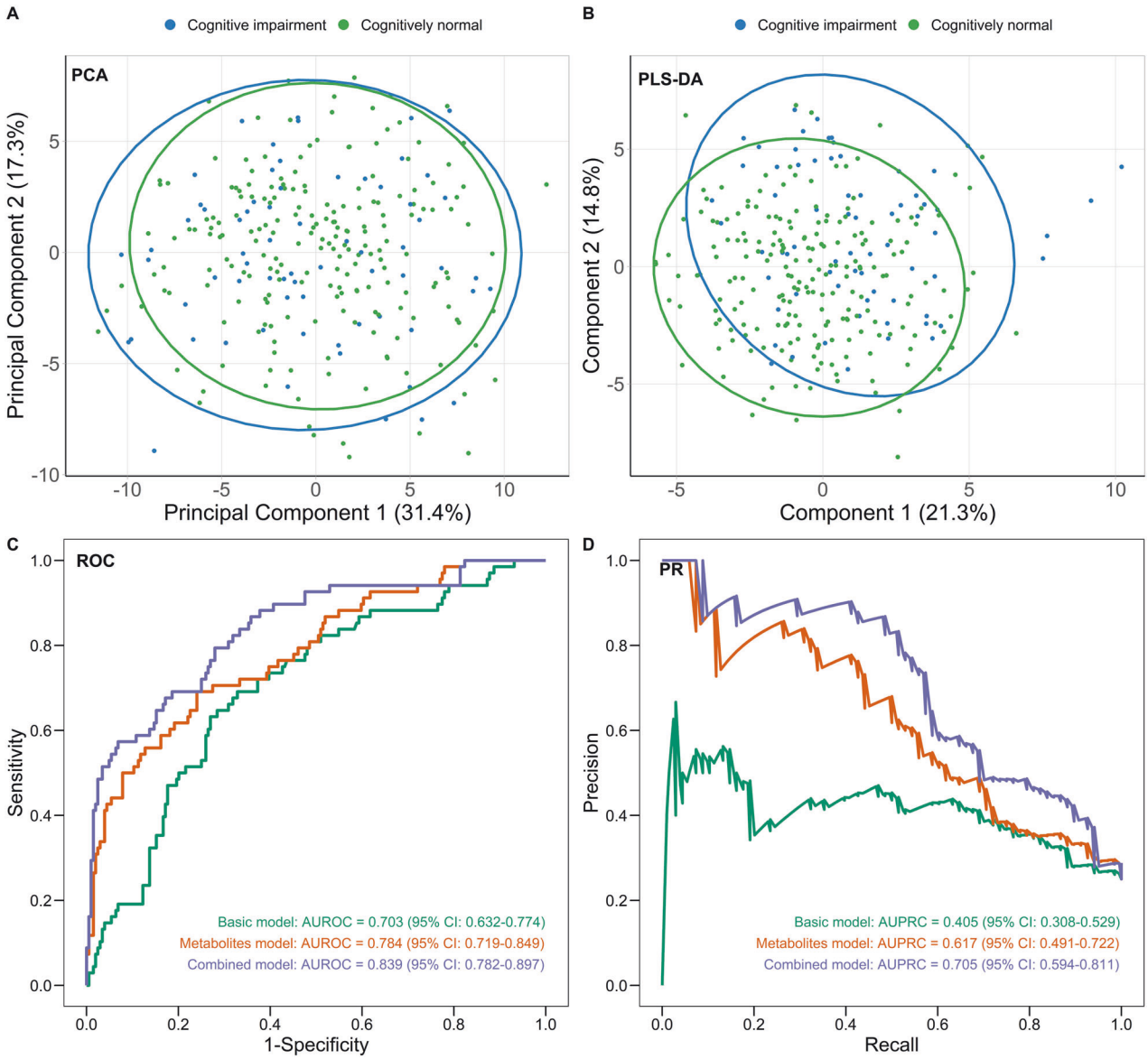

Fig. 4 Prediction models for classifying the incident cognitive impairment and cognitively normal controls. A Principal component analysis (PCA) model with all main measures used. **B** Partial least square discriminant analysis (PLS-DA) model with all main measures used. The permutation test indicated the statistical significance of PLS-DA model ($P = 0.024$). **C** ROC curves of the prediction models. Basic model included traditional risk factors. Metabolites model included 13 selected metabolite biomarkers using Lasso regression. Combined model integrated predictors from basic model and metabolites model. The DeLong test indicated the significant difference between AUROC of the combined model and basic model ($P < 0.001$). **D** PR curves of the prediction models. The bootstrap-based method indicated the significant difference between AUPRC of the combined model and basic model ($P < 0.001$). AUROC Area under the receiver operating characteristic curve, AUPRC Area under the precision-recall curve.

Several prospective studies have reported the inverse association between branched-chain amino acids (BCAA) and risk of cognitive impairment. A study combining data from 22,623 participants with the mean age of 58.4 years found that valine, leucine, and isoleucine were related to 16%, 17%, and 13% lower risk of dementia during an average follow-up of 10.9 years [8]. Another study, conducted among 110,655 participants (age: 56.5 ± 8.1 years) in the UK biobank, reported that the corresponding hazard ratios of dementia for valine, leucine, and isoleucine were 0.86, 0.87, and 0.90 during a median follow-up of 12.2 years [34]. Nevertheless, these studies were conducted in the Western population and a large proportion of the included participants were middle-aged. A study of 1440 Chinese participants (age: 70.7 ± 6.8 years) observed that baseline valine and leucine levels were lower in individuals who developed dementia compared to non-converters. However, no significant associations of BCAA were found in the multivariate Cox regression analysis [11]. In our study, circulating valine and leucine levels were associated with 33% and 29% lower risk of cognitive impairment, confirming the inverse relationships of BCAA (except for isoleucine) in the Asian population. Although isoleucine was not identified as cognitive impairment-related metabolite, the odds ratio for severely impaired cognition was 0.87 (0.75, 1.00) and the one for incident cognitive impairment was 0.77 (0.55, 1.08). Given that BCAA are essential amino acids and their levels are largely determined by diet, these findings may inform dietary intervention to prevent cognitive impairment. In addition to the metabolic pathways of BCAA biosynthesis and degradation, aminoacyl-tRNA biosynthesis was found to be another enriched pathway in our study. Similarly, aminoacyl-tRNA biosynthesis pathway was reported to be

downregulated in the postmortem hippocampus samples of AD patients [35]. Downregulation of aminoacyl-tRNA biosynthesis may affect protein synthesis, as the brains of AD and MCI individuals were found to show a decreased rate and capability for protein synthesis [36, 37]. Further studies are warranted to investigate whether these metabolic pathways contribute to the underlying mechanisms of cognitive impairment.

Additionally, IDL particle number and ApoB in IDL were both associated with 28% lower risk of cognitive impairment in our study. Although previous prospective studies didn't observe consistent associations, the corresponding hazard ratios of IDL particle concentration for dementia were 0.91 (0.80, 1.03) and 0.95 (0.90, 1.01) in the two mentioned Western studies [8, 34]. We also found that triglycerides, cholesterol, and phospholipids from VLDL1 to VLDL4 were inversely related to severely impaired cognition while those from VLDL5 showed no association, suggesting potentially different roles of diverse lipoprotein subfractions. However, findings from our study and previous studies are inconclusive, and relationships between lipoprotein subfractions and cognitive impairment need further research to clarify. RCS showed that GlycA was non-linearly related to risk of cognitive impairment (P for non-linearity <0.05), with the OR of 0.30 (0.12, 0.74) when comparing the third with first quartile. Studies with larger sample sizes are required to further uncover the potential non-linear relationships.

Recent studies have increasingly utilized blood-based metabolite biomarkers to improve the prediction ability of cognitive impairment beyond traditional risk factors [4, 11, 34, 38, 39]. In our study of older Chinese adults, the traditional risk factors-based prediction model yielded an AUROC of 0.703 and AUPRC of 0.405. Incorporating selected metabolite biomarkers into the model substantially increased the AUROC to 0.839 and AUPRC to 0.705, with significant improvement in both NRI and IDI. The positive NRI and IDI indicated the increased proportion of participants assigned to the correct group when comparing the combined model with basic model. Another prediction model of dementia targeting older Chinese adults didn't compare the accuracy between the covariates model (AUROC not provided) and the model combining covariates with metabolites (AUROC: 0.900) [11]. As no external validation was conducted in the current study, further validation of prediction models is required. In addition, predictive performance of metabolomics-based models built with different assumptions (i.e., Lasso, random forest, support vector machine, XGBoost) should be compared in the future studies.

The strengths of our study include the application of a two-step strategy of cross-sectional discovery followed by prospective validation and inclusion of the population at high-risk of developing cognitive impairment from a community-based cohort in China. Nevertheless, several limitations should be considered when interpreting the results of our study. The primary limitation is the nature of an observational study design where the observed associations may be impacted by residual and unmeasured confounding, although the adjustment of multiple covariates have partially mitigated this issue. Second, as in most previous studies, plasma metabolome was measured only once at baseline, thus whether it was representative of long-term exposure status was unclear. Another limitation is the relatively low follow-up rate (52%) in 2021 (Wave 5) due to the COVID-19 pandemic, which may reduce the statistical power of prospective analysis. Fourthly, we failed to find other cohorts with a similar study design to externally validate the identified cognitive impairment-related metabolites and prediction models. Hence, the results should be interpreted cautiously and further independent external validations are required.

In conclusion, our study identified specific plasma metabolites (leucine, valine, IDL particle number, and ApoB in IDL) and potential enriched metabolic pathways (valine, leucine, and isoleucine biosynthesis; valine, leucine, and isoleucine degradation; and aminoacyl-tRNA biosynthesis) related to cognitive impairment. Incorporation of 13 metabolite biomarkers significantly improved the prediction performance for cognitive impairment beyond traditional risk factors. Our findings may contribute to understanding the underlying etiology of cognitive impairment and to identifying high-risk individuals early.

DATA AVAILABILITY

Data described in the article, code book, and code will be made available upon reasonable request.

REFERENCES

1. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health. 2022;7:e105-e125.
2. Bai W, Chen P, Cai H, Zhang Q, Su Z, Cheung T et al. Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies. Age Ageing. 2022;51:afac173
3. Golde TE. Disease-modifying therapies for Alzheimer's disease: more questions than answers. Neurotherapeutics. 2022;19:209–27.
4. Hansson O, Edelmayer RM, Boxer AL, Carrillo MC, Mielke MM, Rabinovici GD, et al. The Alzheimer's Association appropriate use recommendations for blood biomarkers in Alzheimer's disease. Alzheimers Dement. 2022;18:2669–86.
5. Qiu S, Cai Y, Yao H, Lin C, Xie Y, Tang S, et al. Small molecule metabolites: discovery of biomarkers and therapeutic targets. Signal Transduct Target Ther. 2023;8:132.
6. Beger RD, Dunn W, Schmidt MA, Gross SS, Kirwan JA, Cascante M, et al. Metabolomics enables precision medicine: "a white paper, community perspective. Metabolomics. 2016;12:149.
7. Chouraki V, Preis SR, Yang Q, Beiser A, Li S, Larson MG et al. Association of amine biomarkers with incident dementia and Alzheimer's disease in the Framingham Study. Alzheimers Dement 2017;13:1327–36.
8. Tynkkynen J, Chouraki V, van der Lee SJ, Hernesniemi J, Yang Q, Li S et al. Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: a prospective study in eight cohorts. Alzheimers Dement 2018;14:723–33.
9. van der Lee SJ, Teunissen CE, Pool R, Shipley MJ, Teumer A, Chouraki V et al. Circulating metabolites and general cognitive ability and dementia: evidence from 11 cohort studies. Alzheimers Dement 2018;14:707–22.
10. He S, Granot-Hershkovitz E, Zhang Y, Bressler J, Tarraf W, Yu B et al. Blood metabolites predicting mild cognitive impairment in the study of Latinos-investigation of neurocognitive aging (HCHS/SOL). Alzheimers Dement 2022;14:e12259.
11. Cui M, Jiang Y, Zhao Q, Zhu Z, Liang X, Zhang K, et al. Metabolomics and incident dementia in older Chinese adults: the Shanghai Aging Study. Alzheimers Dement. 2020;16:779–88.
12. Ikeuchi T, Kanda M, Kitamura H, Morikawa F, Toru S, Nishimura C, et al. Decreased circulating branched-chain amino acids are associated with development of Alzheimer's disease in elderly individuals with mild cognitive impairment. Front Nutr. 2022;9:1040476.
13. Chen L, Zhernakova DV, Kurilshikov A, Andreu-Sánchez S, Wang D, Augustijn HE, et al. Influence of the microbiome, diet and genetics on inter-individual variation in the human plasma metabolome. Nat Med. 2022;28:2333–43.
14. Liu Z, Wang Y, Zhang Y, Chu X, Wang Z, Qian D, et al. Cohort profile: the rugao longevity and ageing study (RuLAS). Int J Epidemiol. 2016;45:1064–73.
15. Zhang H, Hao M, Hu Z, Li Y, Jiang X, Wang J, et al. Association of immunity markers with the risk of incident frailty: the Rugao longitudinal aging study. Immun Ageing. 2022;19:1.
16. Pu Y, Sun Z, Zhang H, Huang Q, Wang Z, Mei Z et al. Gut microbial features and circulating metabolomic signatures of frailty in older adults. Nat Aging 2024;4:1249–62
17. Ernster VL. Nested case-control studies. Prev Med. 1994;23:587–90.
18. Chen Y-M, Zheng Y, Yu Y, Wang Y, Huang Q, Qian F, et al. Blood molecular markers associated with COVID-19 immunopathology and multi-organ damage. EMBO J. 2020;39:e105896.
19. Wu Q, Huang Q-X, Zeng H-L, Ma S, Lin H-D, Xia M-F, et al. Prediction of metabolic disorders Using NMR-based metabolomics: the Shanghai Changfeng study. Phenomics. 2021;1:186–98.
20. Mei Z, Dong X, Qian Y, Hong D, Xie Z, Yao G, et al. Association between the metabolome and bone mineral density in a Chinese population. EBioMedicine. 2020;62:103111.

21. Chen X, Yang J, Zhang H, Huang Y, Cao Y, Yan S, et al. Plasma folate levels in relation to cognitive impairment: a community-based cohort of older adults in China. Eur J Nutr. 2022;61:2837–45.
22. Tang H-D, Zhou Y, Gao X, Liang L, Hou M-M, Qiao Y, et al. Prevalence and risk factor of cognitive impairment were different between urban and rural population: a community-based study. J Alzheimers Dis. 2016;49:917–25.
23. Cui GH, Yao YH, Xu RF, Tang HD, Jiang GX, Wang Y et al. Cognitive impairment using education-based cutoff points for CMMSE scores in elderly Chinese people of agricultural and rural Shanghai China. Acta Neurol Scand 2011;124:361–7.
24. Huo Z, Lin J, Bat BKK, Chan JYC, Tsoi KKF, Yip BHK. Diagnostic accuracy of dementia screening tools in the Chinese population: a systematic review and meta-analysis of 167 diagnostic studies. Age Ageing. 2021;50:1093–101.
25. Cai G, Zhang M, Ren F, Qin S Validity and reliability of Hasegawa Dementia Scale (HDS) in the screening of senile dementia. [Chinese]. 1991;(01):21–3.
26. Brambati SM, Belleville S, Kergoat M-J, Chayer C, Gauthier S, Joubert S. Single- and multiple-domain amnestic mild cognitive impairment: two sides of the same coin? Dement Geriatr Cogn Disord. 2009;28:541–9.
27. Pocklington C, Gilbody S, Manea L, McMillan D. The diagnostic accuracy of brief versions of the Geriatric Depression Scale: a systematic review and meta-analysis. Int J Geriatr Psychiatry. 2016;31:837–57.
28. Jin X, He W, Zhang Y, Gong E, Niu Z, Ji J, et al. Association of APOE ε4 genotype and lifestyle with cognitive function among Chinese adults aged 80 years and older: a cross-sectional study. PLoS Med. 2021;18:e1003597.
29. Drouin-Chartier J-P, Hernández-Alonso P, Guasch-Ferré M, Ruiz-Canela M, Li J, Wittenbecher C, et al. Dairy consumption, plasma metabolites, and risk of type 2 diabetes. Am J Clin Nutr. 2021;114:163–74.
30. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 2021;49:W388–W396.
31. Steuer R. Review: on the analysis and interpretation of correlations in metabolomic data. Brief Bioinform. 2006;7:151–8.
32. Wang Y, Huang Y, Wu P, Ye Y, Sun F, Yang X, et al. Plasma lipidomics in early pregnancy and risk of gestational diabetes mellitus: a prospective nested case-control study in Chinese women. Am J Clin Nutr. 2021;114:1763–73.
33. Riley RD, Ensor J, Snell KIE, Harrell FE Jr, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020;368:m441.
34. Zhang X, Hu W, Wang Y, Wang W, Liao H, Zhang X, et al. Plasma metabolomic profiles of dementia: a prospective study of 110,655 participants in the UK Biobank. BMC Med. 2022;20:252.
35. Maffioli E, Murtas G, Rabattoni V, Badone B, Tripodi F, Iannuzzi F, et al. Insulin and serine metabolism as sex-specific hallmarks of Alzheimer's disease in the human hippocampus. Cell Rep. 2022;40:111271.
36. Ibba M, Curnow AW, Söll D. Aminoacyl-tRNA synthesis: divergent routes to a common goal. Trends Biochem Sci. 1997;22:39–42.
37. Nunomura A, Perry G. RNA and oxidative stress in Alzheimer's disease: focus on microRNAs. Oxid Med Cell Longev. 2020;2020:2638130.
38. Machado-Fragua MD, Landré B, Chen M, Fayosse A, Dugravot A, Kivimaki M, et al. Circulating serum metabolites as predictors of dementia: a machine learning approach in a 21-year follow-up of the Whitehall II cohort study. BMC Med. 2022;20:334.
39. Buergel T, Steinfeldt J, Ruyoga G, Pietzner M, Bizzarri D, Vojinovic D, et al. Metabolomic profiles predict individual multidisease outcomes. Nat Med. 2022;28:2309–20.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No. 8210120183). The funder of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report or involved in the decision to submit the paper for publication. The authors would like to express genuine gratitude to the participants and staff of the Rugao Longevity and Ageing Study, who contributed greatly to the academic community and made this study possible.

AUTHOR CONTRIBUTIONS

CY, YZ, XJ, XW, and YH designed the study; XS, HZ, XZ, QMH, and YH collected the data; QXH and HT acquired the metabolomics data; YH and QMH performed the statistical analysis; YH drafted the manuscript. All authors interpreted the data and further revised the manuscript. CY, YZ, XJ, and XW supervised the data analysis and interpretation; CY had the primary responsibility for the study final content. All authors critically reviewed the manuscript and approved the final draft.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-024-03147-9.

Correspondence and requests for materials should be addressed to Xiaoyan Jiang, Yan Zheng or Changzheng Yuan.