Principal components from untargeted cerebrospinal fluid metabolomics associated with Alzheimer's disease biomarkers

Ruocheng Dong[1], Diandra N. Denier-Fields[1,2], Carol A. Van Hulle[3,4], Ivonne Suridjan[6], Gwendlyn Kollmorgen[5], Norbert Wild[6], Qiongshi Lu[7], Rozalyn M. Anderson[3,8], Henrik Zetterberg[9-13], Kaj Blennow[9,10], Cynthia M. Carlsson[3,4,8], Sterling C. Johnson[3,4,8,14], Corinne D. Engelman[1,4,14]*

[1]Department of Population Health Sciences, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, 53726, USA
[2]Department Nutrition Science, University of Wisconsin-Madison, Madison, WI, 53706, USA
[3]Department of Medicine, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, 53792, USA
[4]Wisconsin Alzheimer's Disease Research Center, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, 53792, USA
[5]Roche Diagnostics International Ltd, Rotkreuz, CH-6346, Switzerland
[6]Roche Diagnostics GmbH, Penzberg, 82377, Germany
[7]Department of Biostatistics and Medical Informatics, School of Medicine and Public Health, University of Wisconsin, Madison, WI, 53792, USA
[8]Geriatric Research Education and Clinical Center, Wm. S. Middleton Memorial VA Hospital, Madison, WI, 53705, USA
[9]Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, The Sahlgrenska Academy at University of Gothenburg, Mölndal, S-43180, Sweden
[10]Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, S-43180, Sweden
[11]UK Dementia Research Institute at UCL, London, WC1E 6BT, UK
[12]Department of Neurodegenerative Disease, UCL Institute of Neurology, London, WC1H 0AL, UK
[13]Hong Kong Center for Neurodegenerative Diseases, Clear Water Bay, Hong Kong, China
[14]Wisconsin Alzheimer's Institute, University of Wisconsin School of Medicine and Public Health, Madison, WI, 53719, USA

* Correspondence:
Corinne D. Engelman
608-265-5491
cengelman@wisc.edu

Abstract

Studying the correlation between cerebrospinal fluid (CSF) metabolites and the Alzheimer's Disease (AD) biomarkers may offer a window to the alterations of the brain metabolome and unveil potential biological mechanisms underlying AD. In this analysis, 308 CSF metabolites from 338 individuals of Wisconsin Registry for Alzheimer's Prevention and Wisconsin Alzheimer's Disease Research Center were included in a principal component analysis (PCA). The resulted principal components (PCs) were tested for association with CSF total tau (t-tau), phosphorylated tau (p-tau), amyloid β 42 (Aβ42), and Aβ42/40 ratio using linear regression models. Significant PCs were further tested with other CSF NeuroToolKit (NTK) and imaging biomarkers. Using a Bonferroni corrected $p < 0.05$, 5 PCs were significantly associated with CSF p-tau and t-tau and 3 PCs were significantly associated with CSF Aβ42. Pathway analysis suggested that these PCS were enriched in 6 pathways, including metabolism of caffeine and nicotinate and nicotinamide. This study provides evidence that CSF metabolites are associated with AD pathology through core AD biomarkers and other NTK markers and suggests potential pathways to follow up in future studies.

Keywords: Plasma metabolites, MIND diet, physical activity, smoking, caffeine, CSF NeuroToolKit biomarkers, cognition, mediation

1. Introduction

Alzheimer's disease (AD) pathology begins 20 years or more before symptoms arise, and genetic and modifiable risk factors contribute to the pathological initiation and/or progression. When compared to genetic risk factors, most modifiable factors have the advantage that they can be intervened on. Extensive research has suggested that modifiable risk factors, such as diet, physical activity, smoking, and caffeine intake, influence the risk for dementia and AD, and behavioral modification could reduce disease risk[1–8]. Thus, modifiable risk factors provide a potential window for dementia and AD prevention and intervention.

Previous studies have suggested that adherence to specific dietary patterns, including the Mediterranean and Dietary Approaches to Stop Hypertension (DASH), is associated with a reduced risk of dementia and AD. Morris *et al.* introduced a hybrid diet, the Mediterranean-DASH Intervention for Neurodegenerative Delay (MIND) diet, designed to focus on brain health and preventing cognitive decline[1]. The study investigated the MIND diet-AD[1] and MIND diet-cognitive decline[9] relationships and recommended that high and moderate adherence to the MIND diet would slow age-related cognitive decline and lower AD risk. Besides the MIND diet, previous studies have shown that physical activity interventions that include aerobic exercise have a beneficial effect on cognitive function in patients with dementia[10]. Engagement in moderate physical activity was also associated with higher amyloid β (Aβ) 42, lower total tau (T-tau)/Aβ42, and lower phosphorylated tau (P-tau)/Aβ42, which further confirms the protective role of physical activity in AD development. A systematic review[7] of caffeine found evidence supporting favorable effects of coffee consumption on AD risk

and cognitive decline. For example, the Cardiovascular Risk Factors, Aging and Dementia study followed 1409 participants for 21 years and found that daily coffee consumption of 3-5 cups was associated with a 65% decrease in dementia and a 64% decrease in AD[11]. On the contrary, smoking was suggested to have a negative effect on AD risk. A longitudinal study by Rusanen *et al.* found that heavy smoking in midlife was associated with an increased risk for AD and vascular dementia[12]. However, these studies only showed the associations without further evidence of biological mechanisms.

AD is now recognized as a systemic disease that starts influencing peripheral tissues outside the central nervous system from the beginning stages of the disease rather than a previously regarded brain-only disease[13], which has led to more research studying AD pathology in the blood[14,15]. The blood is also undoubtedly useful for studying nutritionally and environmentally influenced metabolomics. The blood is a rich source of transiting nutrients and metabolites between different organs, and the concentration of metabolites is relatively higher in blood when compared to other sources such as urine and saliva[16]. Thus, linking plasma metabolites with MIND diet, physical activity, caffeine intake, and smoking could provide potential evidence of the biological response inside the human body and how this response might influence disease pathology.

In this study, we aim to identify the plasma metabolites that are associated with modifiable risk factors for AD and then determine if these metabolites serve as mediators between the modifiable risk factors and AD endophenotypes in the Wisconsin Registry for Alzheimer's Prevention (WRAP) cohort. The results suggest a range of plasma metabolites associated with each modifiable risk factor. Some of these

metabolites were significant mediators of modifiable risk factor-AD endophenotype (*e.g.,* cognitive function) relationships, providing potential evidence for a biological mechanism.

2. Methods

2.1 Participants

WRAP is an ongoing longitudinal cohort study that started in 2001, with an initial follow-up four years after baseline and subsequent follow-up every two years[17,18]. WRAP is comprised of initially cognitively unimpaired, asymptomatic, English-speaking, middle-aged (between 40 and 65 at baseline) adults enriched for genetic risk due to parental history of clinical AD. At each visit, participants go through a comprehensive medical and cognitive evaluation, complete questionnaires on potential risk factors related to AD and cognitive function, and provide a blood sample. All willing participants also undergo a lumbar puncture for the collection of cerebrospinal fluid. This study was conducted with the approval of the University of Wisconsin Institutional Review Board, and all participants provided signed informed consent before participation.

2.2 Plasma metabolite analyses and quality control (QC)

Plasma metabolomic analyses and quantification in WRAP were performed in one batch by Metabolon (Durham, NC) using an untargeted approach, based on an Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectrometry platform (UPLC-MS/MS)[19]. The raw data were extracted, peak-identified, and QC processed using Metabolon's hardware and software. Metabolites within eight super pathways were identified: amino acids, carbohydrates, cofactors and vitamins, energy, lipids,

nucleotides, peptides, and xenobiotics. Details of the metabolomic profiling were described in an earlier study[20].

For plasma metabolite quality control, each metabolite value was first scaled, so the median was equal to one across all samples. Metabolites with missingness >80% were excluded (n=167). Metabolites with zero variability between individuals (n=0), or an interquartile range of zero (n=178), were also excluded. After this, 757 metabolites with known biochemical names remained in the study. The missing percentage of each metabolite is reported in Supplemental Table 1. The biochemical name, sub-pathway, and super-pathway of each remaining metabolite can be found in Supplemental Table 2. Log 10 transformation was employed to normalize the data.

2.3 Modifiable risk factors

MIND diet measures. The MIND diet is a hybrid of the Mediterranean and DASH diets that emphasizes the dietary components and servings linked to neuroprotection and dementia prevention[1]. The MIND diet includes ten types of healthy foods, green leafy vegetables, other vegetables, nuts, berries, beans, whole grains, fish, poultry, olive oil, and wine, along with five types of unhealthy food, red meats, butter and stick margarine, cheese, pastries and sweets, and fried/fast food. Participants from WRAP completed a 15-item self-reported diet questionnaire developed by Morris *et al.*[1]. In the questionnaire, each item was assigned a score (0, 0.5, and 1) based on the frequency of consumption (higher score for higher consumption of healthy food but lower score for higher consumption of unhealthy food), and the total MIND diet score was computed by summing over all 15 of the component scores. In this analysis, we identified 924

individuals with 979 longitudinal samples of plasma metabolomics and corresponding MIND diet scores.

Physical activity. The WRAP participants completed a self-reported physical activity questionnaire. For each individual, the total metabolic equivalent (MET) hours per week was created by summing up each activity's corresponding MET-hours per week, which was calculated based on the MET values of physical activity developed by McTiernan *et al*[21]. A MET value was assigned to a specific activity according to its intensity classification (*e.g.,* 3.0 for average walking with a speed of 2-3 mph), and the MET-hours per week were calculated by multiplying the MET value by the hours per week spent doing the activity. In this analysis, we identified 1146 individuals with 2352longitudinal samples of plasma metabolomics and corresponding physical activity data.

Smoking. The smoking measure was defined by the question "In the past month, have you smoked cigarettes?" and the responses "0 = No, 1 = Yes, 2 = Don't Know" for the WRAP participants, and we manually set the "Don't know" to missing. In this analysis, we identified 1122 individuals with 2194 longitudinal samples of plasma metabolomics and corresponding smoking data.

Caffeine intake. Caffeine intake was defined by the question "During the past month, how often did you drink any caffeinated beverages? (*e.g.,* coffee, tea, soft drinks)". The response categories were "1 = less than once per day, 2 = 1-2 per day, 3 = 3-5 per day, 4 = 6 or more per day, 999 = missing value". In this analysis, we identified 1146 individuals with 2350 longitudinal samples of plasma metabolomics and corresponding caffeine intake data.

2.4 AD endophenotypes

Cerebrospinal fluid NeuroToolKit biomarkers. All Cerebrospinal fluid (CSF) samples were assayed for biomarkers in the NeuroToolKit (NTK) at the Clinical Neurochemistry Laboratory, University of Gothenburg, using the same batch of reagents, under strict quality control procedures. The immunoassays of Elecsys® β-amyloid(1-42), Phospho-Tau (181P) and Total-Tau, as well as S100 calcium-binding protein B (S100b) and interleukin-6 (IL-6) were performed on a cobas e 601 analyzer[22]. The remaining NTK panel was assayed on a cobas e 411 analyzer including β-amyloid (1-40), α-synuclein, glial fibrillary acidic protein (GFAP), chitinase-3-like protein 1 (YKL-40), soluble triggering receptor expressed on myeloid cells 2 (sTREM2), neurofilament light protein (NfL), and neurogranin[22]. In this analysis, we identified about 165 individuals with 293 longitudinal samples of these NTK biomarkers (the number for each biomarker will be slightly different).

Measures of cognitive function. The standardized widely used clinical neuropsychological tests that provide a comprehensive estimate of cognitive abilities, especially for early-stage AD, were conducted in the WRAP cohort. Three cognitive composite scores: Immediate Learning, Delayed Recall and Executive Function, and a global score, Preclinical Alzheimer Cognitive Composite (PACC)[23] 4-test version[24], were used for the analysis. The tests and measures included in the three composite scores were: (1) Immediate Learning (Rey Auditory Verbal Learning Test [RAVLT] total trials 1–5, Wechsler Memory Scale–Revised Logical Memory subtest [WMS-RLM] immediate recall, and Brief Visuospatial Memory Test [BVMT-R] immediate recall), (2) Delayed Recall (RAVLT long-delay free recall, WMS-RLM delayed recall, and BVMT-R delayed

recall), and (3) Executive Function (Trail Making Test Part B [TMT B] total time to completion, Stroop Neuropsychological Screening Test color-word interference total items completed in 120 seconds, and Wechsler Abbreviated Intelligence Scale–Revised Digit Symbol Coding total items completed in 90 seconds). Other details of the composite scores can be found in a previous study[25].

2.5 Statistical analysis

We first tested the associations between each plasma metabolite and MIND diet, physical activity, smoking and caffeine intake, respectively, by using a linear mixed-effects regression model. All models were adjusted for age, sex, years of education, and body mass index (BMI), with random intercepts to account for the correlation of repeated measures and family (sibling) relationships. Since these four modifiable risk factors may be correlated with each other and potential confounders for each other, we calculated their bivariate associations (Supplemental Table 3.) to decide which modifiable risk factors should be adjusted for in each model. Besides the covariates mentioned above, the MIND diet model was additionally adjusted for physical activity. Similarly, for physical activity, the MIND diet and smoking were included in the model. When testing for smoking, physical activity and caffeine intake were adjusted for in the model. Although the bivariate association test results did not support a strong association between MIND diet and smoking in our data, previous epidemiological research has found an association between smoking and diet, so we performed a sensitivity analysis for the MIND diet by further adjusting for smoking and vice versa. For caffeine intake, smoking was additionally adjusted for. Both Bonferroni and false discovery rate (FDR) methods were applied to adjust p values for multiple testing. After

that, the significant plasma metabolites with Bonferroni adjusted p-value<0.05 were tested for an association with each of the AD endophenotypes, including twelve CSF NTK biomarkers and four cognitive composite scores. If Bonferroni-adjusted significant associations for a plasma metabolite were detected in all previous steps, we conducted a causal mediation analysis to calculate the effect of the modifiable risk factor on the AD endophenotype that goes through the metabolite. The mediation analyses were conducted using the "Mediation" package in R based on the same models mentioned above. However, the mediation function from this package can only include one random intercept, so the intercept for family was not included since the number of related individuals is relatively small.

3. Results

3.1 Participant characteristics

Descriptive characteristics of the WRAP participants were calculated based on 1146 individuals who had data for at least one of the modifiable risk factors of interest and are displayed in Table 1. The most recent measure for each participant was used. For WRAP participants, the mean age was 61, and 69% of the participants were female. The mean BMI was approximately 29, and the mean years of education were 16. The mean MET hours per week was 17 (roughly equivalent to 5-6 hours of normal walking at a speed of 2-3 mph per week), and the mean MIND diet score was 9.2 (moderate adherence to the dietary pattern, maximum value = 15). Just around 5% of participants smoked in the past month, and about 46% drank 1-2 caffeinated beverages per day.

3.2 Plasma metabolites associated with modifiable risk factors and AD endophenotypes

We detected 27 plasma metabolites that were significantly associated with the MIND diet score at the Bonferroni-corrected $p < 0.05$ level (Table 2; 81 significant metabolites at the FDR $q < 0.05$ level [Supplemental Table 4]). The top five metabolites were carotene diol (1), carotene diol (2), S-methylcysteine sulfoxide, indolepropionate, and docosahexaenoate (DHA; 22:6n3). The results of the sensitivity analysis with additional adjustment for smoking are provided in the Supplemental Table 5. For physical activity, 3 metabolites were significantly associated at the Bonferroni-corrected $p < 0.05$ level (11 significant at the FDR $q < 0.05$ level [Supplemental Table 6]). The significant metabolites were imidazole lactate, alpha-hydroxyisovalerate, and glutamate. Twenty-three and 90 metabolites were significantly associated with smoking at the Bonferroni-corrected $p < 0.05$ and FDR $q < 0.05$ levels (Supplemental Table 7), respectively. The top five metabolites were o-cresol sulfate, 3-methyl catechol sulfate (1), 4-vinylguaiacol sulfate, 3-hydroxypyridine sulfate, and N-(2-furoyl)glycine. The results of a sensitivity analysis with additional adjustment for MIND diet are provided in the Supplemental Table 8. Finally, for caffeine intake, 24 and 72 metabolites were significant at the Bonferroni-corrected $p < 0.05$ and FDR $q < 0.05$ levels (Supplemental Table 9), respectively. The top five metabolites were 5-acetylamino-6-amino-3-methyluracil, 1,7-dimethylurate, theophylline, 1,3-dimethylurate and 1-methylxanthine.

From all Bonferroni-corrected statistically significant metabolites associated with the MIND diet score, heptenedioate (C7:1-DC) was significantly associated with P-tau and T-tau, beta-cryptoxanthin was significantly associated with PACC, and hippurate was associated with Immediate Learning. (Table 3). For physical activity, glutamate was associated with NfL. Among metabolites associated with smoking, beta-cryptoxanthin

was associated with Immediate Learning. The full results of all association tests are shown in Supplemental Table 10-17.

3.3 Mediation effect of plasma metabolites

Significant mediation effects were detected for beta-cryptoxanthin in the association between MIND diet score and PACC, hippurate between MIND diet score and Immediate Learning, glutamate between physical activity and NfL, and beta-cryptoxanthin between smoking and PACC. The estimated mediation effect of each model is reported in Figure 1. Other tested mediation analysis results are provided in Supplemental Table 18.

4. Discussion

Using longitudinal data from the WRAP cohort, we were able to identify several plasma metabolites that were associated with the MIND diet, physical activity, smoking, and caffeine intake. Several of these were associated with AD endophenotypes at a strict Bonferroni-adjusted level of significance. Mediation effects were detected for beta-cryptoxanthin in the association between MIND diet and PACC, hippurate in the association between MIND diet and Immediate Learning, glutamate between physical activity and NfL, and beta-cryptoxanthin between smoking and PACC.

Among 27 plasma metabolites that were significantly associated with the MIND diet score, most are derived from specific food groups which compose the MIND diet. Fruit and vegetable consumption in the MIND diet can be directly linked to metabolomics related to antioxidants, specifically vitamin A and polyphenols, and gut microbiome by-products from high fiber foods. For example, carotenoids are found in vegetables and fruits that are orange in color and are also present in olive oil[26]. Our

results showed that a higher MIND diet score is associated with higher vitamin A in the form of carotene diol (1), carotene diol (2), carotene diol (3), and beta-cryptoxanthin, a vitamin A precursor. S-Methylcysteine sulfoxide is also positively associated with the MIND diet score, and its major food sources are Brassicas (cruciferous vegetables; *e.g.*, Brussels sprout)[27]. Seafood is a common staple in the Mediterranean diet and is designated as a positive component of the MIND diet. Both marine and shellfish are rich sources of a variety of long-chain omega-3 fatty acids[28,29]. Specific fatty acids that were positively and significantly related to MIND diet scores in our study were docosahexaenoate (DHA; 22:6n3), eicosapentaenoate (EPA; 20:5n3), and stearidonate (stearidonic acid, SDA; 18:4n3), which are all abundant in seafood[28,29]. Research suggests that DHA protects against AD and other dementias and is also beneficial for human cognitive function[30]. In addition, some vegetables, berries, and red wine also contain polyphenols, another class of antioxidants that includes 3-phenylpropionate (hydrocinnamate)[31]. The major benefit of antioxidants is that they offer protection against free radical damage, which can be beneficial for those at risk for neurodegenerative diseases[32]. Higher consumption of fiber-rich foods, such as whole grains, provides a source of insoluble fiber that is utilized by the gut microbiome to create compounds, including indolepropionate[33]. Overall, these metabolites also provide some biological evidence of how the MIND diet influences the human body through food and potentially protects people from cognitive decline and AD.

Among the metabolites associated with physical activity, glutamate, which refers to the anion of glutamic acid, was previously reported to decrease in plasma with prolonged physical exercise[34,35]. Alpha-hydroxyisovalerate has been reported to be

negatively associated with physical activity energy expenditure in a blood metabolite study based on a Chinese cohort[36].

As for smoking, we were able to replicate several findings from previous studies for metabolites that were associated with smoking status, specifically o-cresol sulfate, 3-methyl catechol sulfate (1), N-(2-furoyl)glycine, oxalate (ethanedioate), 4-vinylphenol sulphate, piperine, threonate and indolepropionate [37,38]. Among them, the o-cresol sulfate and N-(2-furoyl)glycine are marked as respiratory irritants or potential health hazards by the PubChem database, which could be related to pesticide residue and chemical preservatives that are present in cigarettes.

For caffeine intake, six significant metabolites, paraxanthine, 1-methylxanthine, caffeine, 1-methylurate, 1,7-dimethylurate, and 5-acetylamino-6-formylamino-3-methyluracil (AAMU), belong to the caffeine metabolism pathway. Metabolites associated with caffeine intake in our study, such as AAMU, theophylline, paraxanthine, caffeine, quinate, trigonelline, 1-methylxanthine, and 1-methylurate, replicated previous findings[39,40]. Methylxanthines are a class of phytochemicals, including caffeine, theophylline, and 1-methylxanthine, that act as vasodilators and can reduce the risk of stroke, preventing further neurological damage[41].

We identified four significant mediation effects. The MIND diet was associated with higher levels of beta-cryptoxanthin and a higher PACC score. As mentioned above, food components of the MIND diet, like vegetables and fruits, are rich in beta-cryptoxanthin. Beta-cryptoxanthin is an antioxidant protecting organs and tissues from oxidative damage and is the precursor of vitamin A[42]. Previous studies have suggested that vitamin A is crucial for maintaining a higher central nervous system function in

older people[43], and lower levels of vitamin A were usually associated with cognitive decline and higher risk of AD[44–46]. On the contrary, beta-cryptoxanthin was negatively associated with smoking and a significant mediator of the negative association between smoking and PACC. Possible explanations for this mediation effect may be that metabolites or metabolism pathways related to smoking are damaging or decreasing the level of beta-cryptoxanthin[47] or that the lower level of beta-cryptoxanthin is due to a relatively unhealthy diet (e.g., having fewer vegetables) among smokers. A mediation effect of hippurate (hippuric acid) was detected between MIND diet and Immediate Learning. Previous research suggested that hippurate is strongly associated with the consumption of polyphenol-rich foods such as berries and coffee[48,49]. A recent randomized control study conducted by Rutledge *et al.*[49] in 2021 showed that the intake of additional blueberries increased the level of blood hippurate and was associated with improvements in cognition. Our results also suggested that sustained physical activity may reduce the blood level of glutamate, thus resulting in lower CSF NfL; lower levels of CSF NfL are usually detected in cognitively normal individuals when compared to those with mild cognitive impairment and AD[50]. Glutamate is a neurotransmitter and has an important relationship with NfL[51], such as excitotoxicity, which is a cell death mechanism caused by excessive glutamate release from neurons as well as glial cells[52].

The WRAP cohort is an ongoing longitudinal cohort and has multiple time points of modifiable risk factor and metabolomics data per participant. The sample size is fairly large for each modifiable risk factor (around 900-1000). Although we were not able to replicate the findings in an independent cohort, the significant plasma metabolites reported to be associated with modifiable risk factors in our study are based on the most conservative Bonferroni corrected p values. The results also replicate those from

previous findings and are supported by nutritional and biological theories. By adjusting the correlated modifiable risk factors in each model, we are more confident that these significant metabolites are associated with each modifiable risk factor independently rather than driven by other correlated behaviors when compared to previous studies. However, there are some limitations. When we tested the metabolites with CSF biomarkers, especially in the mediation analysis, which also included the modifiable risk factors, the sample size became smaller (around 240-290, depending on the model), reducing the power to detect significant associations. Future research to replicate our results and test for mediation in a larger independent sample would be informative. Since the biology of diet, food, and drink consumption is complex, a more refined mediation model of multiple mediators or multiple pathways would be a logical next step for future research. Another limitation of this analysis is that it is only based on non-Hispanic white individuals due to a small sample size for other racial/ethnic groups. Thus, the results may not be generalizable to other groups.

In summary, our study identified several plasma metabolites that are associated with the MIND diet score, physical activity, smoking, and caffeine intake. These significant plasma metabolites may be employed as biomarkers to track the activity of each modifiable risk factor. For example, beta-cryptoxanthin, DHA, and EPA could be used as biomarkers of the MIND diet. At the same time, these metabolites may also unveil potential biological mechanisms of how modifiable risk factors influence the human body. Testing these hypotheses in an intervention studies are important next steps.

Acknowledgments

The authors especially thank the WRAP participants and staff for their contributions to the studies. Without their efforts, this research would not be possible. This study was supported by the National Institutes of Health (NIH) grants [R01AG27161 (Wisconsin Registry for Alzheimer Prevention: Biomarkers of Preclinical AD), R01AG054047 (Genomic and Metabolomic Data Integration in a Longitudinal Cohort at Risk for Alzheimer's Disease), and R21AG067092 (Identifying Metabolomic Risk Factors in Plasma and Cerebrospinal Fluid for Alzheimer's Disease)], the Helen Bader Foundation, Northwestern Mutual Foundation, Extendicare Foundation, State of Wisconsin, the Clinical and Translational Science Award (CTSA) program through the NIH National Center for Advancing Translational Sciences (NCATS) grant [UL1TR000427], and the University of Wisconsin-Madison Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation. Computational resources were supported by core grants to the Center for Demography and Ecology [P2CHD047873] and the Center for Demography of Health and Aging [P30AG017266].

HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council (#2018-02532), the European Research Council (#681712), Swedish State Support for Clinical Research (#ALFGBG-720931), the Alzheimer Drug Discovery Foundation (ADDF), USA (#201809-2016862), the AD Strategic Fund and the Alzheimer's Association (#ADSF-21-831376-C, #ADSF-21-831381-C and #ADSF-21-831377-C), the Olav Thon Foundation, the Erling-Persson Family Foundation, Stiftelsen för Gamla Tjänarinnor, Hjärnfonden, Sweden (#FO2019-0228), the European Union's

Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860197 (MIRIADE), European Union Joint Program for Neurodegenerative Disorders (JPND2021-00694), and the UK Dementia Research Institute at UCL.

KB is supported by the Swedish Research Council (#2017-00915), ADDF, USA [#RDAPB-201809-2016615], the Swedish Alzheimer Foundation [#AF-742881], Hjärnfonden, Sweden [#FO2017-0243], the Swedish state under the agreement between the Swedish government and the County Councils, the ALF-agreement [#ALFGBG-715986], and European Union Joint Program for Neurodegenerative Disorders [JPND2019-466-236].

Thank Roche for providing the NTK kits for this study. COBAS, COBAS E and ELECSYS are trademarks of Roche.

Conflicts of interest

HZ has served at scientific advisory boards and/or as a consultant for Abbvie, Alector, Annexon, Artery Therapeutics, AZTherapies, CogRx, Denali, Eisai, Nervgen, Pinteon Therapeutics, Red Abbey Labs, Passage Bio, Roche, Samumed, Siemens Healthineers, Triplet Therapeutics, and Wave, has given lectures in symposia sponsored by Cellectricon, Fujirebio, Alzecure, Biogen, and Roche, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program (outside submitted work).

Reference

1. Morris MC, Tangney CC, Wang Y, Sacks FM, Bennett DA, Aggarwal NT. MIND diet associated with reduced incidence of Alzheimer's disease. *Alzheimers Dement*. 2015;11(9):1007-1014. doi:10.1016/j.jalz.2014.11.009

2. Hardman RJ, Kennedy G, Macpherson H, Scholey AB, Pipingas A. Adherence to a Mediterranean-Style Diet and Effects on Cognition in Adults: A Qualitative Evaluation and Systematic Review of Longitudinal and Prospective Trials. *Front Nutr*. 2016;3. doi:10.3389/fnut.2016.00022

3. Stephen R, Hongisto K, Solomon A, Lönnroos E. Physical Activity and Alzheimer's Disease: A Systematic Review. *J Gerontol Ser A*. 2017;72(6):733-739. doi:10.1093/gerona/glw251

4. Tan ZS, Spartano NL, Beiser AS, et al. Physical Activity, Brain Volume, and Dementia Risk: The Framingham Study. *J Gerontol A Biol Sci Med Sci*. Published online July 15, 2016:glw130. doi:10.1093/gerona/glw130

5. Lee PN. Smoking and Alzheimer's Disease: A Review of the Epidemiological Evidence. *Neuroepidemiology*. 1994;13(4):131-144. doi:10.1159/000110372

6. Ott A, Slooter A, Hofman A, et al. Smoking and risk of dementia and Alzheimer's disease in a population-based cohort study: the Rotterdam Study. *The Lancet*. 1998;351(9119):1840-1843. doi:10.1016/S0140-6736(97)07541-7

7. Panza F, Solfrizzi V, Barulli MR, et al. Coffee, tea, and caffeine consumption and prevention of late-life cognitive decline and dementia: A systematic review. *J Nutr Health Aging*. 2015;19(3):313-328. doi:10.1007/s12603-014-0563-8

8. Arendash GW, Cao C. Caffeine and Coffee as Therapeutics Against Alzheimer's Disease. Cunha RA, de Mendonça A, eds. *J Alzheimers Dis*. 2010;20(s1):S117-S126. doi:10.3233/JAD-2010-091249

9. Morris MC, Tangney CC, Wang Y, et al. MIND diet slows cognitive decline with aging. *Alzheimers Dement*. 2015;11(9):1015-1022. doi:10.1016/j.jalz.2015.04.011

10. Groot C, Hooghiemstra AM, Raijmakers PGHM, et al. The effect of physical activity on cognitive function in patients with dementia: A meta-analysis of randomized control trials. *Ageing Res Rev*. 2016;25:13-23. doi:10.1016/j.arr.2015.11.005

11. Eskelinen MH, Ngandu T, Tuomilehto J, Soininen H, Kivipelto M. Midlife Coffee and Tea Drinking and the Risk of Late-Life Dementia: A Population-Based CAIDE Study. *J Alzheimers Dis*. 2009;16(1):85-91. doi:10.3233/JAD-2009-0920

12. Rusanen M, Kivipelto M, Quesenberry CP, Zhou J, Whitmer RA. Heavy Smoking in Midlife and Long-term Risk of Alzheimer Disease and Vascular Dementia. *Arch Intern Med*. 2011;171(4). doi:10.1001/archinternmed.2010.393

13. Khan TK, Alkon DL. Peripheral Biomarkers of Alzheimer's Disease. *J Alzheimers Dis*. 2015;44(3):729-744. doi:10.3233/JAD-142262

14. Shi L, Baird AL, Westwood S, et al. A Decade of Blood Biomarkers for Alzheimer's Disease Research: An Evolving Field, Improving Study Designs, and the Challenge of Replication. Perry G, Avila J, Tabaton M, Zhu X, eds. *J Alzheimers Dis*. 2018;62(3):1181-1198. doi:10.3233/JAD-170531

15. Zetterberg H. Blood-based biomarkers for Alzheimer's disease—An update. *J Neurosci Methods*. 2019;319:2-6. doi:10.1016/j.jneumeth.2018.10.025

16. Gibney MJ, Walsh M, Brennan L, Roche HM, German B, van Ommen B. Metabolomics in human nutrition: opportunities and challenges. *Am J Clin Nutr*. 2005;82(3):497-503. doi:10.1093/ajcn/82.3.497

17. Rue AL, Hermann B, Jones JE, Johnson S, Asthana S, Sager MA. Effect of parental family history of Alzheimer's disease on serial position profiles. *Alzheimers Dement*. 2008;4(4):285-290. doi:10.1016/j.jalz.2008.03.009

18. Johnson SC, Koscik RL, Jonaitis EM, et al. The Wisconsin Registry for Alzheimer's Prevention: A review of findings and current directions. *Alzheimers Dement Diagn Assess Dis Monit*. 2018;10:130-142. doi:10.1016/j.dadm.2017.11.007

19. Bridgewater BR EA. High Resolution Mass Spectrometry Improves Data Quantity and Quality as Compared to Unit Mass Resolution Mass Spectrometry in High-Throughput Profiling Metabolomics. *J Postgenomics Drug Biomark Dev*. 2014;04(02). doi:10.4172/2153-0769.1000132

20. Darst BF, Lu Q, Johnson SC, Engelman CD. Integrated analysis of genomics, longitudinal metabolomics, and Alzheimer's risk factors among 1,111 cohort participants. *bioRxiv*. Published online October 5, 2018. doi:10.1101/436923

21. McTiernan A, Kooperberg C, White E, et al. Recreational Physical Activity and the Risk of Breast Cancer in Postmenopausal WomenThe Women's Health Initiative Cohort Study. *JAMA*. 2003;290(10):1331-1336. doi:10.1001/jama.290.10.1331

22. Van Hulle C, Jonaitis EM, Betthauser TJ, et al. An examination of a novel multipanel of CSF biomarkers in the Alzheimer's disease clinical and pathological continuum. *Alzheimers Dement J Alzheimers Assoc*. Published online December 18, 2020. doi:10.1002/alz.12204

23. Donohue MC, Sperling RA, Salmon DP, et al. The Preclinical Alzheimer Cognitive Composite: Measuring Amyloid-Related Decline. *JAMA Neurol*. 2014;71(8):961-970. doi:10.1001/jamaneurol.2014.803

24. Jonaitis EM, Koscik RL, Clark LR, et al. Measuring longitudinal cognition: Individual tests versus composites. *Alzheimers Dement Diagn Assess Dis Monit*. 2019;11:74-84. doi:10.1016/j.dadm.2018.11.006

25. Clark LR, Racine AM, Koscik RL, et al. Beta-amyloid and cognitive decline in late middle age: Findings from the Wisconsin Registry for Alzheimer's Prevention study. *Alzheimers Dement*. 2016;12(7):805-814. doi:10.1016/j.jalz.2015.12.009

26. Maiani G, Castón MJP, Catasta G, et al. Carotenoids: Actual knowledge on food sources, intakes, stability and bioavailability and their protective role in humans. *Mol Nutr Food Res*. 2009;53(S2):S194-S218. doi:10.1002/mnfr.200800053

27. Marks HS, Hilson JA, Leichtweis HC, Stoewsand GS. S-Methylcysteine sulfoxide in Brassica vegetables and formation of methyl methanethiosulfinate from Brussels sprouts. *J Agric Food Chem*. 1992;40(11):2098-2101. doi:10.1021/jf00023a012

28. Whelan J. Dietary Stearidonic Acid Is a Long Chain (n-3) Polyunsaturated Fatty Acid with Potential Health Benefits. *J Nutr*. 2009;139(1):5-10. doi:10.3945/jn.108.094268

29. Zhang T-T, Xu J, Wang Y-M, Xue C-H. Health benefits of dietary marine DHA/EPA-enriched glycerophospholipids. *Prog Lipid Res*. 2019;75:100997. doi:10.1016/j.plipres.2019.100997

30. Cunnane SC, Plourde M, Pifferi F, Bégin M, Féart C, Barberger-Gateau P. Fish, docosahexaenoic acid and Alzheimer's disease. *Prog Lipid Res*. 2009;48(5):239-256. doi:10.1016/j.plipres.2009.04.001

31. El-Seedi HR, El-Said AMA, Khalifa SAM, et al. Biosynthesis, Natural Sources, Dietary Intake, Pharmacokinetic Properties, and Biological Activities of Hydroxycinnamic Acids. *J Agric Food Chem*. 2012;60(44):10877-10895. doi:10.1021/jf301807g

32. Ganesan K, Xu B. A Critical Review on Polyphenols and Health Benefits of Black Soybeans. *Nutrients*. 2017;9(5):455. doi:10.3390/nu9050455

33. Tuomainen M, Lindström J, Lehtonen M, et al. Associations of serum indolepropionic acid, a gut microbiota metabolite, with type 2 diabetes and low-grade inflammation in high-risk individuals. *Nutr Diabetes*. 2018;8:35. doi:10.1038/s41387-018-0046-9

34. Henriksson J. Effect of exercise on amino acid concentrations in skeletal muscle and plasma. *J Exp Biol*. 1991;160:149-165.

35. Agostini F, Biolo G. Effect of physical activity on glutamine metabolism. *Curr Opin Clin Nutr Metab Care*. 2010;13(1):58–64. doi:10.1097/MCO.0b013e328332f946

36. Xiao Q, Moore SC, Keadle SK, et al. Objectively measured physical activity and plasma metabolomics in the Shanghai Physical Activity Study. *Int J Epidemiol*. 2016;45(5):1433-1444. doi:10.1093/ije/dyw033

37. Gu F, Derkach A, Freedman ND, et al. Cigarette smoking behaviour and blood metabolomics. *Int J Epidemiol*. 2016;45(5):1421-1432. doi:10.1093/ije/dyv330

38. Cross AJ, Boca S, Freedman ND, et al. Metabolites of tobacco smoking and colorectal cancer risk. *Carcinogenesis*. 2014;35(7):1516-1522. doi:10.1093/carcin/bgu071

39. Rothwell JA, Keski-Rahkonen P, Robinot N, et al. A Metabolomic Study of Biomarkers of Habitual Coffee Intake in Four European Countries. *Mol Nutr Food Res*. 2019;63(22):1900659. doi:10.1002/mnfr.201900659

40. Langenau J, Oluwagbemigun K, Brachem C, et al. Blood Metabolomic Profiling Confirms and Identifies Biomarkers of Food Intake. *Metabolites*. 2020;10(11):468. doi:10.3390/metabo10110468

41. Camandola S, Plick N, Mattson MP. Impact of Coffee and Cacao Purine Metabolites on Neuroplasticity and Neurodegenerative Disease. *Neurochem Res*. 2019;44(1):214-227. doi:10.1007/s11064-018-2492-0

42. Burri BJ. Beta-cryptoxanthin as a source of vitamin A. *J Sci Food Agric*. 2015;95(9):1786-1794. doi:10.1002/jsfa.6942

43. Lee H-P, Casadesus G, Zhu X, et al. All-trans retinoic acid as a novel therapeutic strategy for Alzheimer's disease. *Expert Rev Neurother*. 2009;9(11):1615-1621. doi:10.1586/ern.09.86

44. Ono K, Yamada M. Vitamin A and Alzheimer's disease. *Geriatr Gerontol Int*. 2012;12(2):180-188. doi:10.1111/j.1447-0594.2011.00786.x

45. Wołoszynowska-Fraser MU, Kouchmeshky A, McCaffery P. Vitamin A and Retinoic Acid in Cognition and Cognitive Disease. *Annu Rev Nutr*. 2020;40:247-272. doi:10.1146/annurev-nutr-122319-034227

46. Lopes da Silva S, Vellas B, Elemans S, et al. Plasma nutrient status of patients with Alzheimer's disease: Systematic review and meta-analysis. *Alzheimers Dement*. 2014;10(4):485-502. doi:10.1016/j.jalz.2013.05.1771

47. Lagiou, Benetou, Tebelis, Papas, Naska, Trichopoulou. Plasma Carotenoid Levels in Relation to Tobacco Smoking and Demographic Factors. *Int J Vitam Nutr Res*. 2003;73(3):226-231. doi:10.1024/0300-9831.73.3.226

48. Pallister T, Jackson MA, Martin TC, et al. Hippurate as a metabolomic marker of gut microbiome diversity: Modulation by diet and relationship to metabolic syndrome. *Sci Rep*. 2017;7. doi:10.1038/s41598-017-13722-4

49. A. Rutledge G, K. Sandhu A, G. Miller M, Edirisinghe I, B. Burton-Freeman B, Shukitt-Hale B. Blueberry phenolics are associated with cognitive enhancement in

supplemented healthy older adults. *Food Funct*. 2021;12(1):107-118. doi:10.1039/D0FO02125C

50. Dhiman K, Gupta VB, Villemagne VL, et al. Cerebrospinal fluid neurofilament light concentration predicts brain atrophy and cognition in Alzheimer's disease. *Alzheimers Dement Diagn Assess Dis Monit*. 2020;12(1):e12005. doi:https://doi.org/10.1002/dad2.12005

51. Gafson AR, Barthélemy NR, Bomont P, et al. Neurofilaments: neurobiological foundations for biomarker applications. *Brain*. 2020;143(7):1975-1998. doi:10.1093/brain/awaa098

52. Belov Kirdajova D, Kriska J, Tureckova J, Anderova M. Ischemia-Triggered Glutamate Excitotoxicity From the Perspective of Glial Cells. *Front Cell Neurosci*. 2020;0. doi:10.3389/fncel.2020.00051

Table

Table 1. Sample characteristics of WRAP participants' most recent measures.

Characteristics	N=1146	
	Mean	SD
Age(years)	63.5	6.7
Years of education(years)	15.8	2.2
BMI	29.2	6.2
MIND diet score[$]	9.3	2.0
MET hours(hours)	17.4	15.5
	N	%
Gender		
Male	355	31.0
Female	791	69.0
Smoke in past month[*]		
Yes	54	4.7
No	1068	93.2
Caffeine drink[#]		
less than 1/day	249	21.8
1-2/day	556	48.6
3-5/day	300	26.3
6 or more/day	38	3.3

The sample characteristics were calculated based on the most recent measure of participants who have data at least for one modifiable risk factor.

[$] 1078 participants who have non-missing MIND diet score for their most recent measure.

*1122 participants who have non-missing smoking for their most recent measure.

1143 participants who have non-missing caffeine intake for their most recent measure.

Table 2. The significant results of the associations between plasma metabolites and modifiable risk factors.

Outcome	Biochemical names	N	Estimate	P	Adjusted p	FDR q	Super pathway
MIND diet	carotene diol (1)	923	0.0228	1.36E-10	1.03E-07	1.03E-07	Cofactors and Vitamins
	carotene diol (2)	921	0.0244	5.92E-10	4.48E-07	2.24E-07	Cofactors and Vitamins
	S-methylcysteine sulfoxide	923	0.0266	1.59E-08	1.20E-05	3.74E-06	Amino Acid
	indolepropionate	916	0.0310	2.30E-08	1.74E-05	3.74E-06	Amino Acid
	docosahexaenoate (DHA; 22:6n3)	923	0.0228	2.47E-08	1.87E-05	3.74E-06	Lipid
	eicosapentaenoate (EPA; 20:5n3)	923	0.0268	4.93E-08	3.73E-05	6.22E-06	Lipid
	heptenedioate (C7:1-DC)*	912	-0.0180	6.40E-08	4.84E-05	6.48E-06	Lipid
	3-phenylpropionate (hydrocinnamate)	788	0.0414	7.44E-08	5.63E-05	6.48E-06	Xenobiotics
	catechol sulfate	923	0.0263	7.70E-08	5.83E-05	6.48E-06	Xenobiotics
	beta-cryptoxanthin	920	0.0292	1.15E-07	8.73E-05	8.73E-06	Cofactors and Vitamins
	3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)	923	0.0508	3.06E-07	2.32E-04	2.11E-05	Lipid
	cinnamoylglycine	857	0.0466	1.01E-06	7.67E-04	6.39E-05	Xenobiotics
	stearidonate (18:4n3)	922	0.0215	2.68E-06	2.03E-03	1.56E-04	Lipid
	quinate	911	0.0558	4.68E-06	3.54E-03	2.53E-04	Xenobiotics
	sphingosine	864	-0.0169	7.81E-06	5.91E-03	3.94E-04	Lipid
	carotene diol (3)	893	0.0192	1.49E-05	1.13E-02	6.71E-04	Cofactors and Vitamins
	myristoylcarnitine (C14)	923	-0.0126	1.53E-05	1.16E-02	6.71E-04	Lipid
	trigonelline (N'-methylnicotinate)	923	0.0314	1.66E-05	1.26E-02	6.71E-04	Cofactors and Vitamins
	N6-carbamoylthreonyladenosine	922	-0.0088	1.69E-05	1.28E-02	6.71E-04	Nucleotide
	1-stearoyl-2-docosahexaenoyl-GPC (18:0/22:6)	923	0.0114	1.82E-05	1.38E-02	6.88E-04	Lipid
	1-palmitoyl-2-docosahexaenoyl-GPC (16:0/22:6)	923	0.0084	2.96E-05	2.24E-02	1.07E-03	Lipid
	sphingomyelin (d18:2/24:1, d18:1/24:2)*	923	0.0057	4.07E-05	3.08E-02	1.40E-03	Lipid
	1-stearoyl-2-oleoyl-GPG (18:0/18:1)	917	0.0107	4.90E-05	3.71E-02	1.61E-03	Lipid
	margaroylcarnitine (C17)*	914	-0.0118	5.15E-05	3.90E-02	1.62E-03	Lipid
	hippurate	923	0.0250	5.99E-05	4.53E-02	1.81E-03	Xenobiotics
	stearoylcarnitine (C18)	923	-0.0092	6.43E-05	4.86E-02	1.83E-03	Lipid
	4-allylphenol sulfate	918	0.0293	6.54E-05	4.95E-02	1.83E-03	Xenobiotics
Physical activity	imidazole lactate	919	0.0013	1.91E-05	1.44E-02	1.20E-02	Amino Acid
	alpha-hydroxyisovalerate	919	0.0017	4.13E-05	3.12E-02	1.20E-02	Amino Acid
	glutamate	919	-0.0016	5.74E-05	4.35E-02	1.20E-02	Amino Acid
Smoking	o-cresol sulfate	809	0.6441	6.09E-43	4.61E-40	4.61E-40	Xenobiotics
	3-methyl catechol sulfate (1)	1110	0.4255	4.01E-19	3.04E-16	1.52E-16	Xenobiotics
	4-vinylguaiacol sulfate	434	0.5665	3.30E-18	2.50E-15	8.33E-16	Xenobiotics
	3-hydroxypyridine sulfate	1113	0.3644	1.32E-12	1.00E-09	2.50E-10	Xenobiotics
	N-(2-furoyl)glycine	935	0.2502	1.80E-11	1.36E-08	2.72E-09	Xenobiotics
	tartronate (hydroxymalonate)	1118	-0.1517	4.15E-10	3.14E-07	5.23E-08	Xenobiotics
	oxalate (ethanedioate)	1120	-0.1005	2.35E-08	1.78E-05	2.54E-06	Cofactors and Vitamins
	3-carboxy-4-methyl-5-pentyl-2-furanpropionate (3-Cmpfp)**	1120	-0.1138	2.85E-08	2.16E-05	2.70E-06	Lipid
	1-palmitoyl-2-oleoyl-GPC (16:0/18:1)	1120	0.0454	1.17E-07	8.82E-05	9.80E-06	Lipid
	glycerate	1120	-0.0744	1.63E-07	1.23E-04	1.23E-05	Carbohydrate
	4-vinylphenol sulfate	1119	0.2438	8.88E-07	6.72E-04	6.11E-05	Xenobiotics

	piperine	1117	-0.2718	1.16E-06	8.79E-04	7.32E-05	Xenobiotics
	1-palmitoyl-2-palmitoleoyl-GPC (16:0/16:1)*	1120	0.0960	2.41E-06	1.82E-03	1.40E-04	Lipid
	cysteine sulfinic acid	836	0.0816	6.05E-06	4.58E-03	3.27E-04	Amino Acid
	threonate	1120	-0.0990	7.97E-06	6.04E-03	4.02E-04	Cofactors and Vitamins
	N-acetylcarnosine	1120	-0.0797	1.59E-05	1.20E-02	7.51E-04	Amino Acid
	4-hydroxychlorothalonil	1120	-0.0974	1.90E-05	1.44E-02	8.46E-04	Xenobiotics
	quinolinate	1099	-0.0818	2.36E-05	1.78E-02	9.90E-04	Cofactors and Vitamins
	indole-3-carboxylic acid	1023	-0.0857	3.19E-05	2.41E-02	1.27E-03	Amino Acid
	indolepropionate	1118	-0.1528	3.76E-05	2.84E-02	1.42E-03	Amino Acid
	beta-cryptoxanthin	1120	-0.1361	3.97E-05	3.00E-02	1.43E-03	Cofactors and Vitamins
	caffeine	1094	-0.2541	4.40E-05	3.33E-02	1.51E-03	Xenobiotics
	1-arachidonoyl-GPE (20:4n6)*	1120	0.0509	4.61E-05	3.49E-02	1.52E-03	Lipid
Caffeine intake	5-acetylamino-6-amino-3-methyluracil	1106	0.2675	4.28E-91	3.24E-88	3.24E-88	Xenobiotics
	1,7-dimethylurate	1096	0.2178	7.33E-65	5.55E-62	2.77E-62	Xenobiotics
	theophylline	1086	0.2178	3.57E-64	2.71E-61	9.02E-62	Xenobiotics
	1,3-dimethylurate	1013	0.1752	5.86E-57	4.44E-54	1.11E-54	Xenobiotics
	1-methylxanthine	1036	0.1887	2.05E-55	1.55E-52	3.10E-53	Xenobiotics
	paraxanthine	1020	0.1577	5.41E-41	4.10E-38	6.01E-39	Xenobiotics
	3-hydroxypyridine sulfate	1113	0.1828	5.56E-41	4.21E-38	6.01E-39	Xenobiotics
	caffeine	1094	0.2189	3.08E-38	2.33E-35	2.91E-36	Xenobiotics
	trigonelline (N'-methylnicotinate)	1120	0.1464	8.90E-38	6.74E-35	7.49E-36	Cofactors and Vitamins
	1-methylurate	898	0.1584	2.47E-37	1.87E-34	1.87E-35	Xenobiotics
	quinate	1114	0.2234	7.71E-32	5.84E-29	5.30E-30	Xenobiotics
	5-acetylamino-6-formylamino-3-methyluracil	1005	0.1333	3.32E-23	2.51E-20	2.09E-21	Xenobiotics
	3-methyl catechol sulfate (1)	1110	0.1211	2.28E-21	1.73E-18	1.33E-19	Xenobiotics
	1,3,7-trimethylurate	952	0.1154	1.16E-15	8.80E-13	6.28E-14	Xenobiotics
	hippurate	1120	0.0672	3.24E-10	2.45E-07	1.64E-08	Xenobiotics
	cinnamoylglycine	1073	0.0929	1.49E-08	1.13E-05	7.03E-07	Xenobiotics
	N-(2-furoyl)glycine	935	0.0590	1.75E-08	1.33E-05	7.81E-07	Xenobiotics
	catechol sulfate	1120	0.0441	9.65E-08	7.31E-05	4.06E-06	Xenobiotics
	3-phenylpropionate (hydrocinnamate)	1019	0.0666	1.36E-07	1.03E-04	5.43E-06	Xenobiotics
	2-hydroxyglutarate	1115	0.0222	1.28E-06	9.66E-04	4.83E-05	Lipid
	guaiacol sulfate	1120	0.0406	3.12E-06	2.36E-03	1.12E-04	Xenobiotics
	taurine	1120	0.0151	7.56E-06	5.72E-03	2.60E-04	Amino Acid
	N-acetylasparagine	1120	0.0200	2.14E-05	1.62E-02	7.05E-04	Amino Acid
	5alpha-androstan-3beta,17alpha-diol disulfate	935	0.0565	6.04E-05	4.57E-02	1.90E-03	Lipid

The covariates adjusted for in the MIND diet model were age, sex, years of education, BMI, and physical activity.

The covariates adjusted for in the physical activity model were age, sex, years of education, BMI, MIND diet score, and smoking.

The covariates adjusted for in the smoking model were age, sex, years of education, BMI, physical activity, and caffeine intake.

The covariates adjusted for in the caffeine model were age, sex, years of education, BMI, and smoking.

Table 3. Significant results of the associations between modifiable risk factor-associated plasma metabolites and AD endophenotypes.

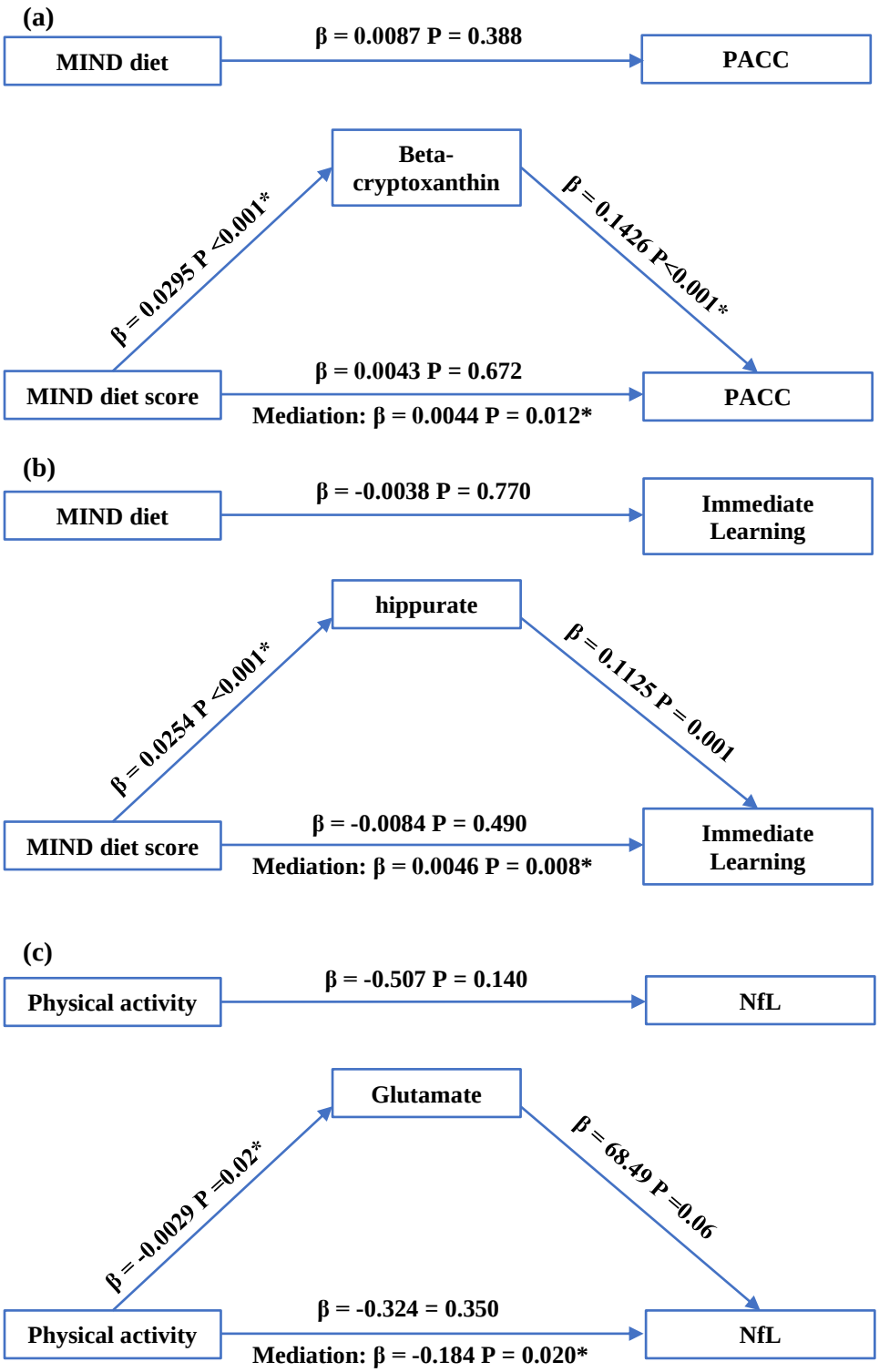
Modifiable risk factor	Associated metabolite	Estimate	p	Adjusted p	FDR q	Associated endophenotypes
MIND diet	heptenedioate (C7:1-DC)	3.08	2.98E-02	2.98E-02	2.98E-02	**P-tau**
	heptenedioate (C7:1-DC)	30.90	1.67E-03	4.51E-02	4.51E-02	**T-tau**
	beta-cryptoxanthin	0.14	6.06E-04	1.64E-02	1.64E-02	**PACC**
	hippurate	0.11	1.74E-03	4.69E-02	3.74E-02	**Immediate Learning**
Physical activity	glutamate	65.70	8.89E-03	2.67E-02	2.67E-02	**NfL**
Smoking	beta-cryptoxanthin	0.14	6.15E-04	1.41E-02	1.41E-02	**PACC**

Age, sex, years of education, BMI, and physical activity, were adjusted for in the model of MIND-diet associated metabolites.

Age, sex, years of education, BMI, MIND diet score, and smoking were adjusted for in the model of physical activity associated metabolites.

Age, sex, years of education, BMI, and caffeine intake were adjusted for in the model of smoking associated metabolites.

Figure


(a)
MIND diet
β = 0.0087 P = 0.388
PACC
Beta-cryptoxanthin
β = 0.0295 P <0.001*
β = 0.1426 P<0.001*
MIND diet score
β = 0.0043 P = 0.672
PACC
Mediation: β = 0.0044 P = 0.012*
(b)
MIND diet
β = -0.0038 P = 0.770
Immediate Learning
hippurate
β = 0.0254 P <0.001*
β = 0.1125 P = 0.001
MIND diet score
β = -0.0084 P = 0.490
Immediate Learning
Mediation: β = 0.0046 P = 0.008*
(c)
Physical activity
β = -0.507 P = 0.140
NfL
Glutamate
β = -0.0029 P =0.02*
β = 68.49 P =0.06
Physical activity
β = -0.324 = 0.350
NfL
Mediation: β = -0.184 P = 0.020*

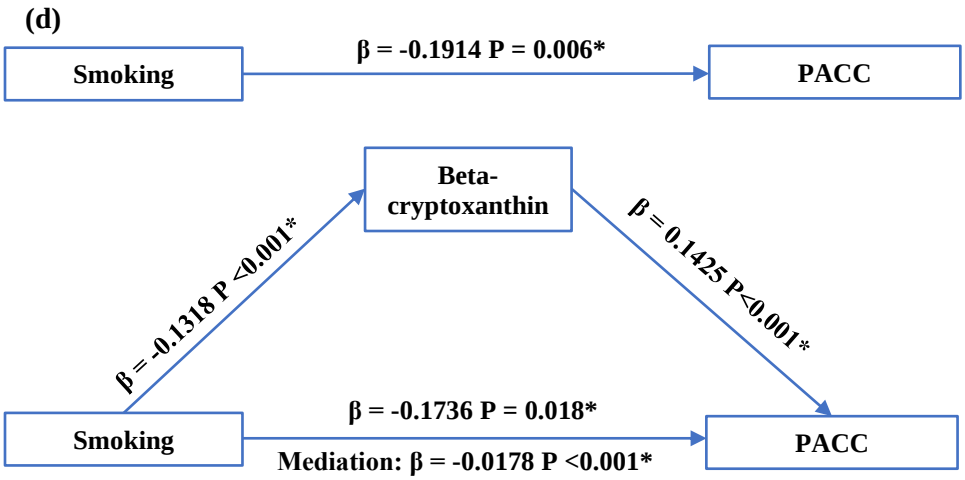


Figure 1. Estimated total effects, direct effects, and mediation effects of metabolites between modifiable risk factors and AD endophenotypes. (a) MIND diet→beta-cryptoxanthin→PACC (b) MIND diet→hippurate→Immediate Learning (c) physical activity→glutamate→NfL (d) smoking → beta-cryptoxanthin→PACC. The upper part of the figure represents the total effect estimated from the modifiable risk factor to the AD endophenotype. The lower triangle represents the corresponding estimated effect between the modifiable risk factor and the metabolite (upper left of the triangle), between the metabolite and AD endophenotype (upper right of the triangle), the direct effect of the modifiable risk factor on the AD endophenotype (base of triangle above the arrow), and the mediation (indirect) effect (base of triangle below the arrow). * indicates the effect was significant statistically.