Archival Report

Biological Psychiatry: GOS

Disentangling Independent and Mediated Causal Relationships Between Blood Metabolites, Cognitive Factors, and Alzheimer's Disease

Jodie Lord, Rebecca Green, Shing Wan Choi, Christopher Hübel, Dag Aarsland, Latha Velayudhan, Pak Sham, Cristina Legido-Quigley, Marcus Richards, Richard Dobson, and Petroula Proitsi, on behalf of the Alzheimer's Disease Neuroimaging Initiative, the GERAD1 Consortium, and AddNeuroMed

ABSTRACT

BACKGROUND: Education and cognition demonstrate consistent inverse associations with Alzheimer's disease (AD). The biological underpinnings, however, remain unclear. Blood metabolites reflect the end point of biological processes and are accessible and malleable. Identifying metabolites with etiological relevance to AD and disentangling how these relate to cognitive factors along the AD causal pathway could, therefore, offer unique insights into underlying causal mechanisms.

METHODS: Using data from the largest metabolomics genome-wide association study ($N \approx 24{,}925$) and three independent AD cohorts ($N = 4725$), cross-trait polygenic scores were generated and meta-analyzed. Metabolites genetically associated with AD were taken forward for causal analyses. Bidirectional two-sample Mendelian randomization interrogated univariable causal relationships between 1) metabolites and AD; 2) education and cognition; 3) metabolites, education, and cognition; and 4) education, cognition, and AD. Mediating relationships were computed using multivariable Mendelian randomization.

RESULTS: Thirty-four metabolites were genetically associated with AD at $p < .05$. Of these, glutamine and free cholesterol in extra-large high-density lipoproteins demonstrated a protective causal effect (glutamine: 95% confidence interval [CI], 0.70 to 0.92; free cholesterol in extra-large high-density lipoproteins: 95% CI, 0.75 to 0.92). An AD-protective effect was also observed for education (95% CI, 0.61 to 0.85) and cognition (95% CI, 0.60 to 0.89), with bidirectional mediation evident. Cognition as a mediator of the education-AD relationship was stronger than vice versa, however. No evidence of mediation via any metabolite was found.

CONCLUSIONS: Glutamine and free cholesterol in extra-large high-density lipoproteins show protective causal effects on AD. Education and cognition also demonstrate protection, though education's effect is almost entirely mediated by cognition. These insights provide key pieces of the AD causal puzzle, important for informing future multimodal work and progressing toward effective intervention strategies.

https://doi.org/10.1016/j.bpsgos.2021.07.010

Late-onset Alzheimer's disease (AD) impacts over 47 million individuals worldwide (1). Factors such as educational attainment (EA) and cognition demonstrate protective associations with AD (2–4), and these may indeed be causal (5–8). Etiological mechanisms underlying such relationships, however, remain unclear. Understanding the biological basis through which cognitive factors may exert their protective effect, as well as establishing direct markers of disease pathogenesis more generally, could therefore hold special value in advancing treatment and prevention strategies.

Blood metabolites—small molecular compounds such as lipids and amino acids—represent the crosstalk between genomic encoding and influences from the surrounding environment (9,10). These analytes could therefore provide mechanistic clues as to how cognitive exposures influence later AD risk. Further, as metabolites are quantifiable via a simple blood test and of plausible size to cross the blood-brain barrier (11), they represent promising candidates for direct treatment intervention. While research has indeed implicated several metabolites, particularly lipids, in AD and cognitive processing (12–15), the weight of evidence derives from observational studies. These remain problematic with respect to informing intervention strategies, as uncaptured confounding and reverse causation risk incorrect causal inferences. Moreover, associative studies allow little opportunity to understand specific pathways into disease end points, and as such, mediating relationships have been little explored.

ISSN: 2667-1743

Mendelian randomization (MR) presents a statistical methodology akin to a randomized controlled trial that allows researchers to investigate putative causal relationships through use of genetic variants as randomizing instruments (16). Multivariable (MV) extensions then allow for the interrogation of exposures that may exist along the same causal pathway, helping to disentangle independent from mediated causal effects (17,18). In a previous study, we utilized MR to investigate causal relationships between 19 candidate metabolites and AD, finding a protective effect of extra-large high-density lipoproteins (XL.HDLs) and a risk-increasing effect of glycoprotein acetyls (19). Candidate metabolites were, however, restricted to those previously associated with midlife cognition (20), rather than with AD specifically, neglecting other causal candidates that may be of relevance. Mediating relationships were also not explored, which, if demonstrated, could offer novel intervention sources and provide a richer understanding of the etiological drivers behind relationships observed.

This study sought to extend our previous findings, this time using cross-trait polygenic risk scoring (PRS) to screen for AD-specific candidate metabolites, selecting only those genetically predictive of AD diagnosis. Then, incorporating information from education and cognition together with that of the selected metabolites and AD, we computed independent versus mediated causal relationships using a combination of univariable MR (UVMR) and MVMR. In this way, we distinguished the extent to which causal relationships were mutually independent, or whether they reflected a chain of interdependent events along the same AD causal pathway.

METHODS AND MATERIALS

Data Sources

A summary of datasets utilized across both PRS and MR analyses can be found within Table 1, with further details available in Supplementary Information I2 in Supplement 1. Briefly, summary statistics from the largest metabolomics genome-wide association study (GWAS) (21) were utilized for both PRS and MR analyses. Nonoverlapping, individual-level genomic data across three AD cohorts—1) the Genetic and Environmental Risk in Alzheimer's Disease (GERAD1) consortium (https://gtr.ukri.org/project/B6C58A7C-3C3E-41CB-AF10-16DB59962C9E/), 2) the Alzheimer's Disease Neuroimaging Initiative (ADNI) (http://adni.loni.usc.edu/) (22), and 3) the AddNeuroMed and Dementia Case Register (ANM) (23,24)—were utilized as PRS target phenotypes, and summary level data from Kunkle *et al.* (25) represented AD in all MR analyses. Finally, summary data from Lee *et al.* (26) and Savage *et al.* (27) represented EA and cognition, respectively.

Polygenic Score Preparation

Quality control was conducted across AD cohorts separately, each following the same pipeline (Figure 1). Principal component analysis was performed using EIGENSOFT 6.1.4. (https://www.hsph.harvard.edu/alkes-price/software/), and genotyped data were imputed via the Sanger Imputation Service (https://imputation.sanger.ac.uk/). Individuals of non-European ancestry or whose most recent diagnosis was mild cognitive impairment or non-AD dementia were excluded. The minimum sample ages within the ADNI, ANM, and GERAD1 cohorts were 54, 53, and 43 years, respectively. As AD has a long symptom-free prodrome, samples were age-matched to a conservative minimum of 70 years. Single nucleotide polymorphisms (SNPs) within 750 kb of *APOE* were removed. Following imputation, separately typed platforms within the ADNI cohorts were merged.

To quantify signal within metabolite datasets, estimates of SNP heritability (h^2_{SNP}) were computed using linkage disequilibrium score regression (28). h^2Z scores ($h^2Z = h^2_{SNP}/se$) were subsequently computed (29), and datasets with $h^2Z < 2$ or with h^2_{SNP} 0 <> 1 were excluded.

MR Preparation

To maximize instrumental variable validity, SNPs were selected only if they 1) were associated with their exposure below genome-wide significance ($p < 5 \times 10^{-8}$), 2) demonstrated a computed $F > 10$ (30), and 3) were not within 750 kb of the *APOE* genomic region (owing to known pleiotropy) (31,32). For AD, EA, and intelligence, SNPs were clumped using an r^2 threshold of 0.001 (Tables S2–S4 in Supplement 2). For metabolites, instruments were selected from a set of precurated metabolite quantitative trait loci available within MR-Base (33); no additional clumping was required for these (Table S5 in Supplement 2). Finally, all datasets were harmonized, and any SNPs with noninferable palindromic SNPs or with minor allele frequency < 0.01 were excluded. Metabolite minor allele frequencies were used to infer AD allele

Table 1. Summary of Datasets Acquired for Use Across PRS and MR Analyses

Phenotype	Dataset	Phenotypic Measure	Sample *N*	Dataset *N*	Level	MR or PRS
Blood Metabolites	Kettunen *et al.* (21)	Nuclear magnetic resonance spectroscopy in plasma	~24,925	123	Summary	Both
Alzheimer's Disease	GERAD1[a]	Clinical diagnosis	4515	1	Individual	PRS
	ADNI (22)	Clinical diagnosis	1674	3[b]	Individual	PRS
	ANM (23,24)	Clinical diagnosis	1063	1	Individual	PRS
	Kunkle *et al.* (25)	Clinical diagnosis	63,926	1	Summary	MR
Educational Attainment	Lee *et al.* (26)	Highest education level obtained by 30 years of age	1.1 million	1	Summary	MR
Cognition	Savage *et al.* (27)	Multiple dimensions of cognitive functioning represented by a common latent "g" factor	269,867	1	Summary	MR

ADNI, Alzheimer's Disease Neuroimaging Initiative; ANM, AddNeuroMed and Dementia Case Register; GERAD1, Genetic and Environmental Risk in Alzheimer's Disease; MR, Mendelian randomization; PRS, polygenic risk scoring.

[a]For more information on the GERAD1 Consortium, see the Acknowledgments and Disclosures section.

[b]Available ADNI datasets were merged following data preprocessing.

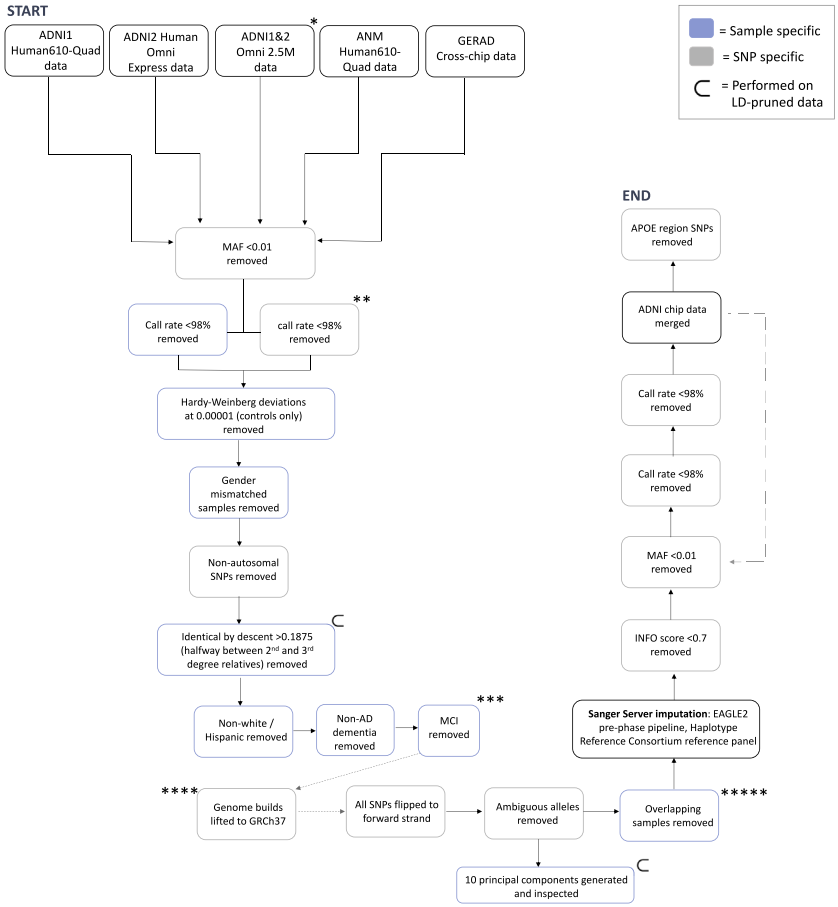

Figure 1. Flow chart overviewing quality control pipelines applied to all AD genotype datasets (separately) prior to polygenic risk scoring analyses. *The Illumina Omni2.5 microarray chip includes samples across ADNI1 and ADNI2, whole-genome sequenced at high coverage, and subsequently genotyped on the high-coverage Illumina chip. **Missingness for both SNPs and samples were inspected iteratively, from 90% to 98%, iterating between SNPs and samples in steps of 1%. ***Latest diagnosis was used to classify samples into cases and controls. Late-stage MCI, with MCI owing to probable AD, and clinician confidence score of 3–4 (indicating high confidence) remained in analyses as cases. ****Required for the Human610-Quad platform only. *****For overlaps between ADN1 chip data (Human610-Quad) and Omni 2.5M, duplicates in Omni 2.5M were removed and duplicates in Human610-Quad were retained. For overlaps between ADNI1 chip data (HumanOmniExpress) and Omni 2.5M, duplicates in HumanOmniExpress were removed and duplicates in Omni 2.5M were retained. No overlaps were observed between Human610-Quad and HumanOmniExpress. AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; ANM, AddNeuroMed and Dementia Case Register; GERAD, Genetic and Environmental Risk in Alzheimer's Disease; LD, linkage disequilibrium; MAF, minor allele frequency; MCI, mild cognitive impairment; SNP, single nucleotide polymorphism.

frequencies, as these were unavailable. All data extraction, preprocessing, and analyses were performed using the MR-Base package (v.0.4.25) (33). The detailed scope is provided in Supplementary Information I3 in Supplement 1.

Primary Statistical Analyses

Cross-Trait Genetic Associations Using Polygenic Scoring. PRS models were generated using PRSice-2 (34). Each metabolite was set as the model-generating base dataset that was then used to predict status within each AD dataset separately. For each metabolite, models were generated across 10 predefined p-value thresholds (P_Ts): 5×10^{-8}, 1×10^{-5}, 1×10^{-4}, .001, .01, .05, .1, .2, .5, 1. SNPs with $r^2 > 0.1$ were clumped, and those with p value $< P_T$ were weighted by their effect size, aggregated within the PRS of the corresponding model, and regressed on AD genotypes separately. Models were standardized and adjusted for sex, age, and seven principal components (Supplementary Information I4 in Supplement 1).

PRS Meta-analyses. PRS for metabolites at every P_T were meta-analyzed across the three AD datasets (35). For each metabolite, the 10 meta-analyzed results (one per P_T) were ranked by their p value to obtain the metabolite most significantly associated with AD status (Figure S1 in Supplement 1). Pseudo-R^2s were back-computed using the meta-analyzed regression coefficients (Supplementary Information I5 in Supplement 1), and I^2 and Cochran's Q assessed cross-study heterogeneity ($I^2 > 0.5$ and Q-$p < .05$). An adjusted alpha of $\alpha = 0.0002$ was computed to account for multiple testing (Supplementary Information I6 in Supplement 1).

Bidirectional Causal Analyses. Inverse variance–weighted (IVW) two-sample UVMR computed nonmediated causal relationships (32). To compute the total causal effect of each metabolite on AD, metabolite$_{1...j}$ was set as the exposure (x) in turn, and AD was set as the outcome (y). To interrogate reverse causation, x and y were then reversed. This bidirectional procedure was repeated for metabolites and education, metabolites and cognition, education and AD, cognition and AD, and education and cognition. While the temporal order of education (highest grade at 30 years of age) and AD make reverse causation implausible, bidirectional analyses were undertaken as a negative control. All results were computed in SD units.

Causal Mediation. MVMR was employed to interrogate mediation. This followed the same framework as UVMR

but with the addition of a second "mediating" variable (*m*), allowing for a causal estimate of *x* on *y* while holding *m* constant. A difference between the total estimate (*c*) derived from UVMR and the direct estimate (*c*′) derived from MVMR was used to signify a mediating effect of *m* on the *x*-*y* relationship (Figure 2) (18). To ensure that direct estimates were not biased by confounding of *m*, datasets were cross-checked to ensure that no instrumental variable overlap existed between *x* and *m* (Tables S2–S5 in Supplement 2). In this way, any *m*-to-*x* confounding would not enter into *c*′ (17).

For variables to be taken forward to MVMR, evidence of an association from *x* to *m* (when *m* = *y*) and from *m* to *y* (when *m* = *x*) was first required within UVMR. A total causal effect of *x* to *y* was not, however, necessary owing to potential suppression through *m* (36). UV results were therefore used to inform variable selection for MVMR. Any exposure with a UV effect on *y* but no UV effect on a potential *m* (when *m* = y) was automatically deemed to be acting independently of *m* and not selected for MVMR.

Sensitivity Analyses

For both UVMR and MVMR analyses, causal effects were re-estimated using MR-Egger (37). For UVMR, several additional sensitivities were undertaken, including 1) weighted median MR (32), 2) leave-one-out analysis, 3) instrument heterogeneity analysis (Cochran's *Q*) (38), and 4) MR–Bayesian model averaging (metabolite-AD analyses only) (39). Further information can be found in Supplementary Information I7 in Supplement 1.

Post Hoc Analyses

Following primary analyses, an instrumental variable outlier for one metabolite with a statistically significant causal association with AD—glutamine—was observed. To further explore the validity of glutamine as a causal analyte, two post hoc analyses were undertaken.

Single-SNP MR. The instrumental variable outlier (rs2657879) identified for glutamine demonstrated a notably significant genome-wide association ($p = 3 \times 10^{-70}$) relative to other glutamine SNPs. rs2657879 is also located within the genomic region of *GLS2*, a protein-encoding gene involved in the conversion of glutamine to glutamate as part of the glutamine-glutamate cycle (Figures S2 and S3 in Supplement 1) (40–42), indicating strong biological relevance. We therefore carried out a post hoc single-SNP MR using an independent dataset (43), setting only rs2657879 as the instrumental variable. The Wald ratio was used to assess causality (44).

Subthreshold MR. At the opposite extreme, we investigated how increasing SNP instrumental variable numbers might improve power in the absence of rs2657879. Relaxing the P_T risks introducing pleiotropic instruments. However, we used knowledge from our PRS analyses to inform the subthreshold selected, relaxing this to independent SNPs significant at $p < .0001$.

APOE-Related Polygenic Association. For all primary analyses, SNPs within the *APOE* genomic region were removed. For PRS, *APOE* was removed owing to its unusually

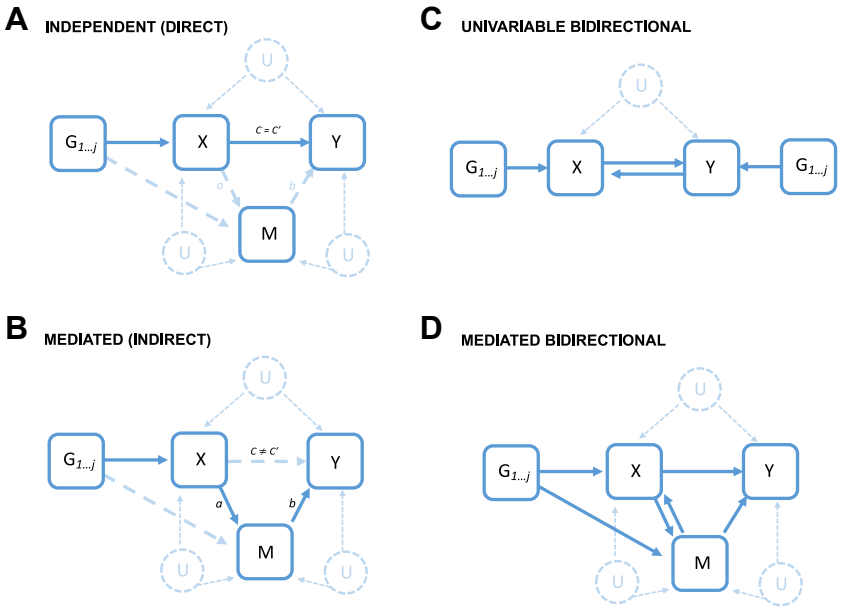

Figure 2. Diagrammatic illustration of causal paths identifiable using univariable and multivariable Mendelian randomization, assuming that Mendelian randomization assumptions are satisfied. $G_{1...j}$ represents genetic instruments used as proxy for random treatment assignment. X represents exposure of interest. Y represents outcome of interest. M represents mediator of interest. U represents potential unmeasured confounding. C represents total causal estimates (univariable estimate when mediator unaccounted for). C′ represents direct causal estimate (multivariable estimate when holding the mediator constant). a-path (X to M) * b-path (M to Y) represents mediated path of X to Y via M. Solid arrows represent observed causal effects in the direction depicted by the arrowhead. Dashed arrows represent potential directed relationships not fully observed. For panel **(A)**, dashed arrows from X to M and M to Y represent the possibility of either an active a-path or an active b-path but no significant effect of both. This would be expected if there was no difference between univariable (total) or multivariable (direct) estimates of X. For panel **(B)**, a dashed arrow from X to M represents an inactivation of the C′ path when M is introduced via multivariable models. C′ may be partially inactivated (reduced magnitude of C′ relative to C but a significant relationship maintained) or fully inactivated (complete loss of causal signal in C′ estimate) by the presence of M. Panels **(C)** and **(D)** depict bidirectional relationships. For panel **(C)**, a significant causal effect is observed when switching X and Y as the exposure of interest. For panel **(D)**, a significant causal effect is observed when switching X and M as the mediator of interest. All models allow for unmeasured confounding between X, M, and Y. Core MR assumptions are assumed to hold in all models: 1) no $G_{1...j}$ to X confounding, 2) no direct $G_{1...j}$-to-Y relationship, and 3) a robust relationship between $G_{1...j}$ and X.

large effect size for both AD (25) and lipid-related metabolites (21), risking dominating and drowning out the wider polygenic signal. For lipid-related metabolites associated with AD at $p < .05$ in PRS analyses, however, results were in the unexpected direction (see Results; Figure 3). To investigate whether this reflected *APOE* removal, PRS meta-analyses were recomputed for all metabolites significantly associated with AD at $p < .05$ but with SNPs restricted to the *APOE* genomic region only (chr19, bp 45786555-45037368). The quality control pipeline followed that outlined for primary analyses in Figure 1, with the exception of "*APOE* region SNPs removed." PRS analyses were then recomputed, restricting P_Ts to those identified as each metabolite's best threshold within primary analyses (Table S7 in Supplement 2).

Information on further post hoc analyses can be found in Supplementary Information I9 in Supplement 1.

RESULTS

Polygenic Association Between 34 Metabolites and Alzheimer's Disease

Following data preparation, 106 metabolites (Table S1 in Supplement 2) and 4725 AD case-control samples (GERAD1 = 3191, ADNI = 886, ANM = 648) (Table S8 in Supplement 2) were taken forward for PRS. When meta-analyzed, 34 metabolites were genetically associated with AD at $p < .05$ (Figure 3; Table S7 in Supplement 2), and no significant heterogeneity was evident ($I^2 < 0.5$, $Q\text{-}p > .05$). No PRS reached adjusted significance ($\alpha = 0.0002$), but glutamine came close at $p = .0009$ ($se = 0.03$, $P_T = 0.0001$) (Table S7 in Supplement 2). In the absence of associations meeting multiple testing thresholds, metabolites associated with AD at the more liberal, conventional threshold of $p < .05$, indicative of suggestive association, were taken forward for causal analyses.

Protective Causal Effect of Glutamine and Free Cholesterol in XL.HDLs on Alzheimer's Disease

An adjusted $\alpha = 0.004$ was computed for MR analyses (Supplementary Information I6 in Supplement 1). Two metabolites—glutamine and free cholesterol in XL.HDLs (XL.HDL.FC)—demonstrated evidence of a small protective causal effect on AD (glutamine: IVW odds ratio [OR], 0.80; $p = .002$; XL.HDL.FC: IVW-OR, 0.83; $p = .001$) (Figure 4). Consistent directionality was seen for both Egger and weighted median estimates, though lower precision of Egger resulted in confidence intervals (CIs) crossing the null (Figure 4A). No pleiotropy was evidenced by the Egger intercept, nor was heterogeneity apparent (Table S9 in Supplement 2). MR–Bayesian model averaging also corroborated, ranking XL.HDL.FC and glutamine with the highest marginal inclusion probability, indicative of being the strongest "true causal" candidates of those analyzed (Table S19 in Supplement 2). Both also represented the most frequent metabolites within MR–Bayesian model averaging group models (Supplementary Information I10 in Supplement 1, Table S20 in Supplement 2). Leave-one-out analysis did, however, identify one outlier for glutamine (rs2657879), the removal of which resulted in non-significance (IVW-OR, 0.82; 95% CI, 0.62 to 1.10) (Figure S12 in Supplement 1). No other metabolite demonstrated evidence of a causal association with AD, in either direction (Table S9 and S10 in Supplement 2, Figures S4 and S5 in Supplement 1).

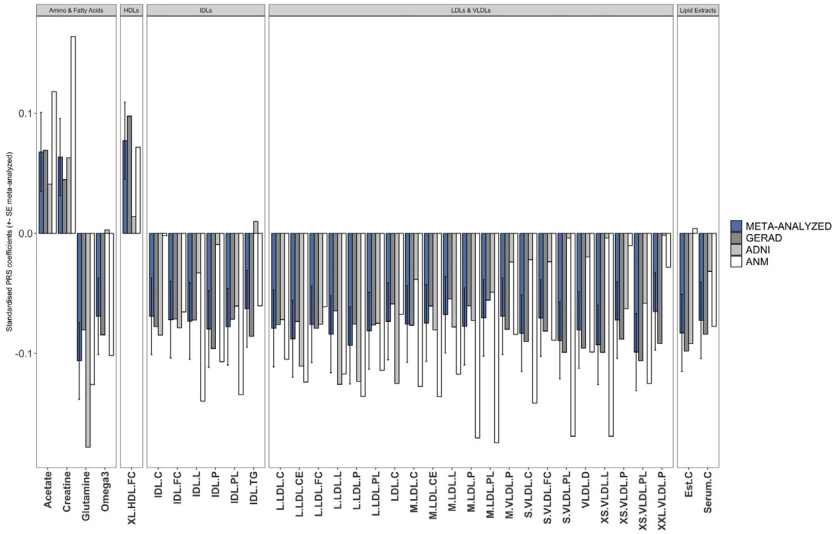

Figure 3. Bar chart demonstrating the magnitude of effect and meta-analyzed standard errors for metabolite–Alzheimer's disease cross-trait polygenic score associations. Results are displayed for meta-analyzed associations significant at $p < .05$. Meta-analyzed coefficients are represented by blue bars. Non–meta-analyzed coefficients are displayed in gray hues for comparison. Metabolites are grouped into metabolite families. ADNI, Alzheimer's Disease Neuroimaging Initiative; ANM, AddNeuroMed and Dementia Case Register; Est.C, free-cholesterol-to-esterified-cholesterol ratio; GERAD1, Genetic and Environmental Risk in Alzheimer's Disease; HDL, high-density lipoprotein; IDL, intermediate-density lipoprotein; IDL.C, IDL total cholesterol; IDL.FC, IDL free cholesterol; IDL.L, IDL lipids; IDL.P, IDL concentration of particles; IDL.PL, IDL phospholipids; IDL.TG, IDL triglycerides; LDL, low-density lipoprotein; L.LDL.C, large LDL total cholesterol; L.LDL.CE, large LDL cholesterol esters; L.LDL.FC, large LDL free cholesterol; L.LDL.L, large LDL lipids; L.LDL.P, large LDL concentration of particles; L.LDL.PL, large LDL phospholipids; LDL.C, LDL total cholesterol; M.LDL.C, medium LDL total cholesterol; M.LDL.CE, medium LDL cholesterol esters; M.LDL.L, medium LDL lipids; M.LDL.P, medium LDL concentration of particles; M.LDL.PL, medium LDL phospholipids; M.VLDL.P, medium VLDL concentration of particles; S.VLDL.C, small VLDL total cholesterol; S.VLDL.FC, small VLDL free cholesterol; S.VLDL.PL, small VLDL phospholipids; Serum.C, serum cholesterol; VLDL, very low-density lipoprotein; VLDL.D, VLDL mean diameter; XL.HDL.FC, extra-large HDL free cholesterol; XS.VLDL.L, extra-small VLDL lipids; XS.VLDL.P, extra-small VLDL concentration of particles; XS.VLDL.PL, extra-small VLDL phospholipids; XXL.VLDL.P, extra-extra-large VLDL concentration of particles.

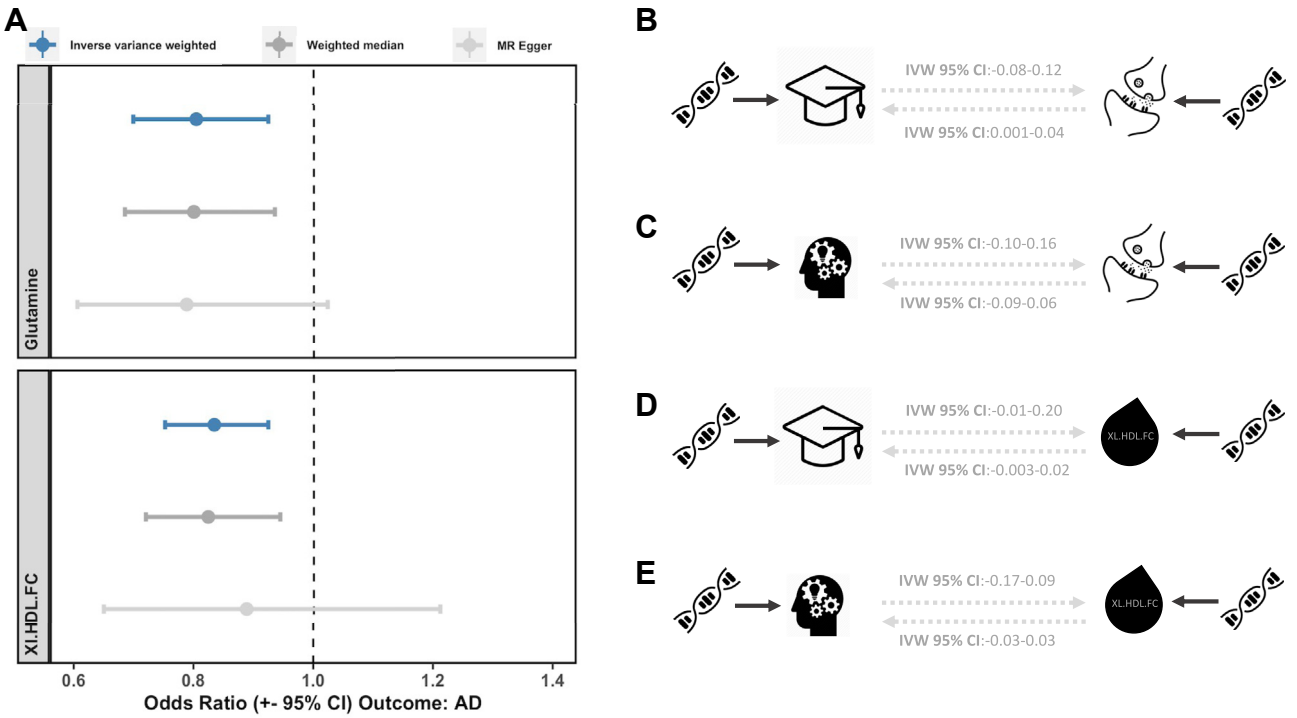

Figure 4. Causal effects of glutamine and XL.HDL.FC. **(A)** Forest plot illustrating odds ratio point estimates and 95% CIs for the effect of glutamine (top panel) and XL.HDL.FC (bottom panel) on AD in primary univariable analyses (blue bars) and secondary analyses (weighted median = dark gray hue, Egger = light gray hue). **(B)** Causal diagram and 95% CIs from primary univariable MR of the causal effect of education on glutamine (top arrows) and of glutamine on education (bottom arrow). **(C)** Causal diagram and 95% CIs from primary univariable MR of the causal effect of cognition on glutamine (top arrows) and of glutamine on cognition (bottom arrow). **(D)** Causal diagram and 95% CIs from primary univariable MR of the causal effect of education on XL.HDL.FC (top arrows) and of XL.HDL.FC on education (bottom arrow). **(E)** Causal diagram and 95% CIs from primary univariable MR of the causal effect of cognition on XL.HDL.FC (top arrows) and of XL.HDL.FC on cognition (bottom arrow). Double helix icons represent genetic instrumental variables used in MR analyses. Causal arrows and CIs are displayed in a grayed hue to represent the nonsignificant causal relationship observed across these variables, indicating no evidence of shared causal pathways. AD, Alzheimer's disease; CI, confidence interval; IVW, inverse variance weighted; MR, Mendelian randomization; XL.HDL.FC, free cholesterol in extra-large high-density lipoprotein.

Negative Causal Effect of Cognition on Lipid-Related Metabolites

A nominally significant causal effect of cognition on 18 lipid-related metabolites was observed. Metabolites belonged primarily to the low-density lipoprotein (LDL) or very LDL (VLDL) family (n = 11). All associations were in the negative direction (Figure S6 in Supplement 1), indicating higher cognition may result in lowered levels of these metabolites. However, no estimate survived multiple testing and inconsistent directionality was seen for 17 of the 18 metabolites in Egger analyses (Table S11 in Supplement 2). No evidence of a causal effect in the opposite direction (metabolite to cognition) was found (Table S12 in Supplement 2).

Causal Association Between EA and Lipid-Related Metabolites

A nominally significant causal effect of education on nine lipid-related metabolites (Figure S7 in Supplement 1, Table S13 in Supplement 2) was found. One metabolite, triglycerides in intermediate-density lipoproteins (IDL.TGs), remained significant at the adjusted level (p = .002). This effect was negative (IVW-β = −0.18; 95% CI, −0.29 to −0.07), indicating that higher education results in lower IDL.TG. Directionality was consistent for both Egger and weighted median, and no pleiotropy was evidenced by the Egger intercept (Table S13 in Supplement 2).

In the opposite direction, omega-3 fatty acids (FAω3) was the only metabolite to demonstrate evidence of a causal effect on education. This was in the positive direction, though the magnitude of effect was small, and significance was at the nominal level only (IVW-β = 0.02; 95% CI, 0.01 to 0.46; p = .04) (Table S14 in Supplement 2). FAω3 was also one of nine metabolites associated with education in the opposite direction (IVW-β = 0.16; 95% CI, 0.02 to 0.30; p = .03). However, instrument heterogeneity was evident (Q-p = .01), and inconsistent directionality was observed for Egger (Table S13 in Supplement 2).

Protective Causal Effect of Cognition and EA on AD, With Bidirectional Mediation

Both education and cognition demonstrated evidence of a negative causal association with AD (Figures S8 and S9 in Supplement 1). These were both significant at the adjusted level, indicating a protective effect (education: IVW-OR, 0.72; 95% CI, 0.61 to 0.84, $p = 7.34 \times 10^{-5}$; cognition: IVW-OR, 0.73; 95% CI, 0.60 to 0.90, p = .002) (Table S15 in Supplement 2). Sensitivity analyses demonstrated consistent directionality for both Egger and weighted median. Wider CIs

were observed for Egger estimates, but no pleiotropy was indicated by the Egger intercept (Table S15 in Supplement 2). Some instrument heterogeneity was evident for cognition (Q-p = .01), though leave-one-out analysis identified no significant change following per-instrument removal (Figures S14 and S15 in Supplement 1). No evidence of a causal effect in the opposite direction was found (Table S16 in Supplement 2, Figures S10 and S11 in Supplement 1).

A bidirectional causal relationship between cognition and education was also observed (Figure 5A). This was larger in the direction of education to cognition (education to cognition: IVW-β = 0.67; 95% CI, 0.63 to 0.71; cognition to education: IVW-β = 0.30; 95% CI, 0.27 to 0.33) (Table S17 in Supplement 2).

MVMR also demonstrated evidence of bidirectional mediation between cognition and education with respect to their effect on AD. Mediation via cognition was, however, stronger than via education. More specifically, when cognition was introduced into education-AD models, the causal estimate of education on AD became smaller in magnitude (MV-IVW-OR, 0.84) and was no longer significant (p = .17), indicating total mediation via cognition (Figure 5B; Table S18 in Supplement 2). While magnitude of the cognition-AD relationship also dropped when introducing education (MV-IVW-OR, 0.81), the direct effect of cognition retained nominal significance, indicating only partial mediation via education (p = .049).

No Evidence of Mediation Between Metabolites and Cognitive Factors on AD

Neither glutamine nor XL.HDL.FC demonstrated a causal association with education or cognition (Figure 4B–E; Tables S12 and S14 in Supplement 2). Thus, the effect of both these metabolites on AD was deemed independent of education and cognition, with no MV model necessary. Similarly, while a number of suggestive associations between cognition and metabolites were observed (see Causal Effect of Cognition on Lipid-Related Metabolites), these were in the direction of cognition to metabolite. As none of these metabolites demonstrated a causal association with AD (Table S9 in Supplement 2), a significant b-path required for mediation was absent, and the effect of cognition on AD deemed independent of these metabolites. There was one metabolite—FAω3—that demonstrated suggestive evidence of contributing to increased EA (Table S14 in Supplement 2). Investigating EA as a mediator on the causal pathway from FAω3 to AD, however, demonstrated no evidence of a mediating effect (Table S18 in Supplement 2).

Post Hoc Analyses: Glutamine as a Protective Analyte for AD

The Wald ratio was used to re-estimate the causal association between glutamine and AD using only influential SNP rs2657879 in an independent dataset (Table S21 in Supplement 2) (43). Results corroborated primary analyses with notably greater effect magnitude, though lower precision (OR, 0.035; 95% CI, 0.003 to 0.381; p = .006) (Table S22 in Supplement 2; Figure 6).

A PRS-informed subthreshold of 100 glutamine-associated instruments (r^2 = 0) significant at $p < .0001$ (rs2657879 excluded) (Table S23 in Supplement 2) also corroborated a protective causal effect of glutamine, with comparable magnitude to primary results (OR, 0.89; 95% CI, 0.81 to 0.98; p = .0009) (Figure 6; Table S22 in Supplement 2).

Post Hoc Analyses: *APOE*

Of the 34 metabolites associated with AD at $p < .05$ in primary PRS analyses, 30 were *APOE*-related lipid-subfractions

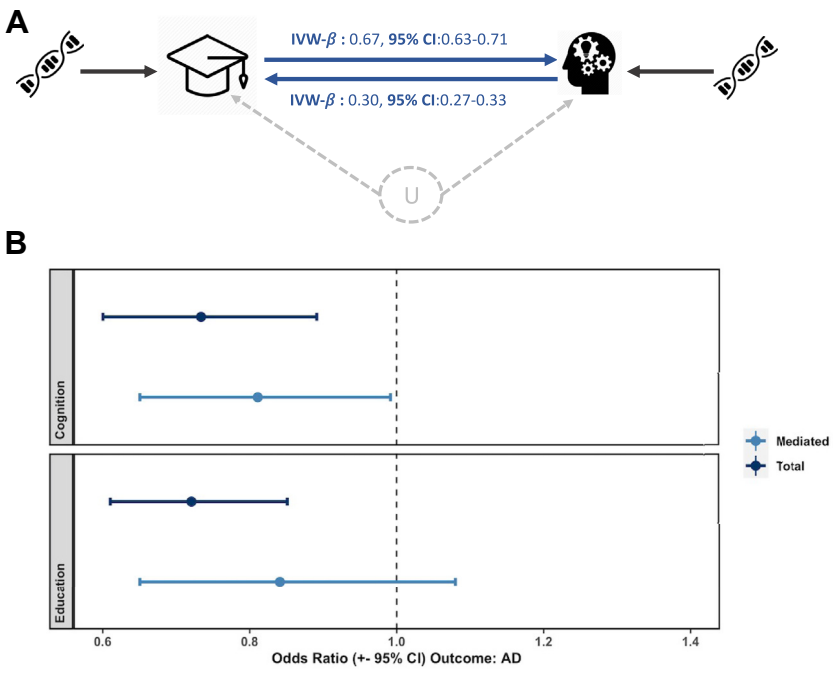

Figure 5. Bidirectional effects of intelligence and educational attainment. **(A)** Causal diagram illustrating the univariable bidirectional relationship between education (left) and cognition (right). Double helix icons on the far left and far right represent genetic instrumental variables ($_{1...j}$). U represents unobserved confounding (allowable provided no instrumental-variable-to-U relationship). The darker-hued arrow confirms the causal estimate when x = education and y = cognition. The lighter-hued arrow confirms the causal estimate when x = cognition and y = education. **(B)** Forest plot confirming the total (univariable) causal estimate of education and cognition on Alzheimer's disease (AD) (darker hue) and the mediated causal estimate of education on AD when controlling for cognition and of cognition on AD when controlling for education (lighter hue). CI, confidence interval; IVW, inverse variance weighted.

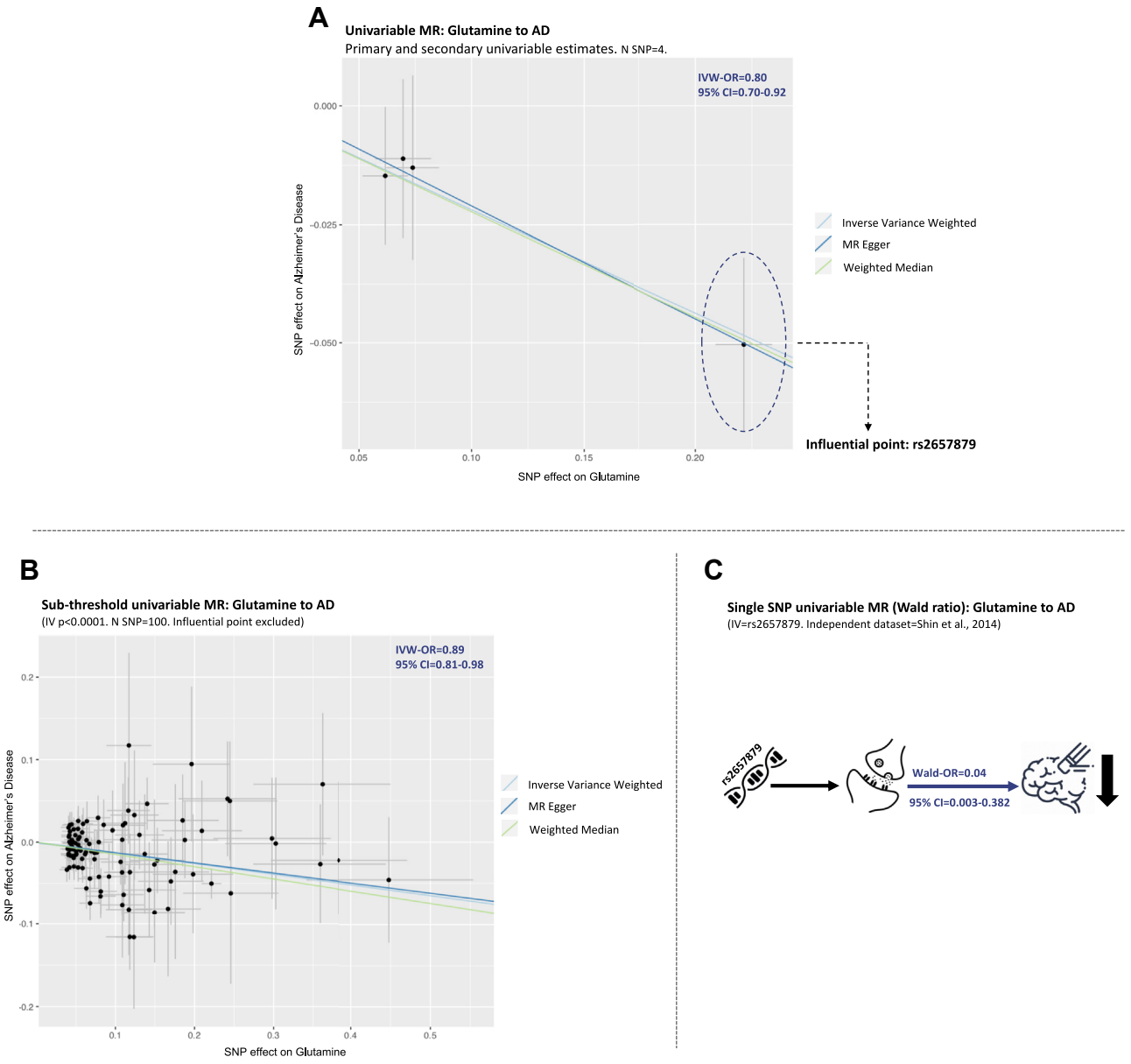

Figure 6. Corroboration between primary, secondary, and post hoc estimates of glutamine on AD. **(A)** Scatterplot illustrating pooled per-instrument estimates of the effect of glutamine on AD within primary IVW analyses (light blue slope) and secondary MR-Egger (dark blue slope) and weighted median (green slope) analyses. For each instrument, 95% CIs are displayed for 1) SNP outcome estimates (vertical bars) and 2) SNP exposure estimates (horizontal bars). Slopes across the three MR estimates indicate comparable protective causal estimates. Influential point rs2657879 is highlighted (far right). **(B)** Scatterplot illustrating pooled per-instrument estimates for the effect of glutamine on AD within post hoc subthreshold analyses, with independent instruments (R^2 = 0) associated with glutamine at $p < .0001$ included in analyses (most predictive cross-trait polygenic risk scoring threshold when glutamine scoring models used to predict AD). Slopes for IVW, Egger, and weighted median subthreshold estimates exclude the influential SNP rs2657879. **(C)** Directed acyclic graph and corresponding causal estimate and 95% CIs for the effect of glutamine on AD within post hoc analyses using influential point rs2657879 as a single SNP instrumental variable estimated using the Wald ratio, and in an independent dataset from Shin *et al.* (43). The far-left icon represents the genetic instrumental variable, the middle icon represents glutamine as X, and the far-right icon represents AD as Y, with the downward arrow signifying a reduced effect of AD given higher levels of glutamine. AD, Alzheimer's disease; CI, confidence interval; IV, instrumental variable; IVW, inverse variance–weighted; MR, Mendelian randomization; OR, odds ratio; SNP, single nucleotide polymorphism.

(VLDLs = 9, LDLs = 11, IDLs = 7, HDLs = 1, cholesterols = 2) (Table S7 in Supplement 2), and each of these demonstrated associations in the unexpected direction (Figure 3). Reperforming PRS for these 30 metabolites, but restricting SNPs to the *APOE* region, rectified directionality, with associations matching that which would be expected based on

wider literature. Statistical significance also greatly improved, ranging from $p = .001$ (XL.HDL.FC) to $p = 2 \times 10^{-59}$ (FC in small VLDLs) (Table S24 in Supplement 2).

DISCUSSION

To our knowledge, this is the first study to triangulate knowledge across PRS, UVMR, and MVMR to disentangle causal relationships between blood metabolites, cognitive factors, and AD. Polygenic scores allowed us to identify, from a wider set of metabolites, those demonstrating plausible genetic overlap with AD. MR then allowed us to interrogate both direct and indirect causality, identifying two metabolites, glutamine and XL.HDL.FC, as having direct AD-protective effects, independent of EA or cognition. An AD-protective effect was also confirmed for both education and cognition, though no evidence of mediation via any of our metabolites was evident. There was, however, bidirectional mediation confirmed between education and cognition, with the mediating effect of cognition being strongest. In other words, education's protective effect appears to work almost entirely through its positive effect on cognition, and this, in turn, reduces AD risk.

Glutamine and AD

Glutamine demonstrated the strongest genetic overlap with AD in cross-trait polygenic scoring. It was also one of two metabolites demonstrating a causal effect on AD. Though sensitivity analyses indicated glutamine's causal estimate was primarily driven by influential SNP rs2657879, post hoc analyses added weight to initial conclusions, corroborating primary findings. Glutamine also has biological relevance, as it is critically implicated in neuronal transmission as part of the glutamine-glutamate cycle (40,42) (Figure S3 in Supplement 1). The leading view in literature is that glutamine is indeed AD protective (45,46), though some positive associations, both with AD and with lower cognition, have also been reported (45,47). The small but statistically significant protective effect found here adds weight to the former. Given glutamine's intrinsic link to glutamate, findings also implicitly implicate glutamate in AD etiology. This assertion would align with wider literature, which has indeed found a link between this metabolite and AD pathology (45). Unfortunately, glutamate was unable to be interrogated directly within our study owing to its nonavailability within Kettunen *et al.* (21), and while it was available within a smaller GWAS (43), no genome-wide significant SNPs were identified. It will be an important endeavor to incorporate this metabolite into future analyses once instruments of adequate power become available.

XL.HDL.FC and AD

As reported in our previous study (19), XL.HDL.FC demonstrated a protective causal effect on AD. This aligns with wider literature that consistently regards HDLs as health promoting, while regarding LDLs and IDLs as their risk-increasing counterparts (48–50). HDLs have been implicated in reduced cognitive decline and AD more specifically (45,51), and hold biological relevance, having shown evidence of protection against neuroinflammation and cerebral amyloid angiopathy (51). Interestingly, however, our PRS results painted a somewhat opposing picture to that of MR. Here, XL.HDL.FC was found to have a positive AD-association, while all LDLs and IDLs demonstrated negative associations. Expected directionality was recovered, however, in post hoc PRS that restricted SNPs to only those within the *APOE* genomic region, indicating that primary PRS results may, in part, reflect the removal of *APOE*-related SNPs. A priori removal was necessary owing to the risk of *APOE*'s unusually large effect size (21,25) drowning out the wider polygenic signal. However, evidence from post hoc results suggest that because *APOE* largely dominates the genetic relationship between lipid subfractions and AD, removal may bring to bear counterintuitive associations that could bias conclusions if not properly interrogated. This phenomenon may be particularly pertinent for polygenic scores as opposed to MR, as PRS explicitly assumes polygenicity, resulting in bias when such assumptions do not hold, such as in the presence of unconventionally large effects (52). While it is outside the scope of the current study, it will be interesting to expand on the genetic associations observed here and further investigate the polygenic relationship between these metabolites and AD on the basis of, and interactions with, *APOE* status.

Three additional HDL subfractions, as well as one marker of inflammation—glycoprotein acetyls—demonstrated AD causal associations within our previous study (19). These were not estimated here owing to their lack of AD polygenic overlap. This highlights the importance of adopting an expanded repertoire of screening methods for identifying causal candidates, with previous selection based on phenotypic associations with midlife cognition (19) as opposed to genetic overlap with AD-specific diagnosis, as was done here.

Cognitive Factors and AD

A strong bidirectional causal relationship was observed between education and cognition, indicating a causal feedback loop. In line with this, both education and cognition demonstrated a total (nonmediated) protective effect on AD, with similar magnitudes of effect. Additionally, when the direct effect for each was measured with consideration of the other, evidence of bidirectional mediation was present. However, while cognition was only partially mediated by education, the independent effect of education was entirely attenuated by education, suggesting that education's protective effect on AD is working via its positive effect on cognition. These results mirror findings from a recent study (5) using smaller-scaled data from 1) Lambert *et al.* (53) for AD, 2) Okbay *et al.* (54) for education, and 3) Hill *et al.* (55) for cognition. Here, too, a bidirectional relationship between education and cognition, and a mediating effect of cognition on the education-AD relationship, was found (5). To our knowledge, ours is the first study to successfully replicate these 2020 findings on larger, independent data.

Cognitive Factors and Metabolites

Evidence was found for education causing lower levels of IDL.TG. This causal association was the only to survive multiple testing when investigating cognitive factors and metabolites. IDLs lie between LDLs and VLDLs, and therefore most closely resemble a family of metabolites consistently associated with adverse vascular outcomes (56–59). Similarly, higher

triglyceride levels have been implicated in poor neurocognitive outcomes (60–62) and have previously shown associations with dementia (63). Interestingly, while our results indicated that education may reduce levels of this potentially harmful metabolite, there was no evidence that this translated through to either cognition or later AD. Considering the magnitude of education's effect on IDL.TGs, it may be that its small effect is not of large enough magnitude to then translate through to a detectable indirect effect on subsequent cognition. As education's effect on AD is totally mediated by cognition, this would mean that any indirect effect of IDL.TGs on AD would not be observed owing to the absent IDL.TG-cognition path. Alternatively, such results may reflect power, as many more instruments were available for both education (n = 277) and cognition (n = 133) relative to metabolites (maximum: n = 15; IDL.TG: n = 12). For lipid-related metabolites such as IDL.TG, a large proportion of genetic signal was also removed through exclusion of *APOE*-related instruments. This was necessary owing to known pleiotropy (32). However, coupled with low instrumental variable numbers to begin with, this likely attenuated detectable causal signals. Inferences regarding the extent of IDL.TG's impact on cognition and AD should therefore remain conservative until larger-powered samples become available.

Several additional causal associations were observed but failed to survive multiple testing. We advise caution in overinterpreting these and thus refrain from discussing them here. A brief interpretation can, however, be found in Supplementary Information I11 in Supplement 1.

Limitations

Metabolites studied here were limited to those available within the GWAS literature. As a result, only 123 were available for investigation. The human metabolome is estimated to contain over 250,000 metabolites (64). The extent to which we have captured all relevant metabolic mechanisms within our study therefore remains doubtful. Moreover, of those that were studied, quantification relied on nuclear magnetic resonance spectroscopy, a method with limited specificity relative to alternative methods such as mass spectrometry. It is worth noting that GWAS data for a larger number of mass spectrometry–quantified metabolites ($N \approx 400$) were available at the time of study (43). However, sample power ($N \approx 7824$) lagged considerable behind data used within our study ($N \approx 24{,}925$), resulting in fewer available instrumental variables. This alternative dataset did, however, provide an accessible independent dataset for post hoc interrogation of an influential point identified for glutamine. Moreover, while breadth and specificity were suboptimal in our selected study data (21), key metabolites previously implicated in cognition, cognitive decline, and dementia were indeed present.

Throughout analyses, AD was also quantified using a binary diagnostic measure. However, AD clinical manifestations become apparent only after a long symptom-free prodromal period (65). It remains plausible, therefore, that our clinical phenotype contains noise, with prodromal AD cases incorrectly classified among controls. It is worth noting that an attempt to avoid such noise was made throughout. For PRS, a conservative age-matched cutoff of 70 years across all samples was implemented to avoid contamination where possible. For MR, our smaller GWAS of clinically diagnosed individuals was selected over a larger alternative dataset of AD-by-proxy samples, in which diagnosis was derived from self-reported parental dementia, a phenotype of lower specificity (7). Nonetheless, quantification of the AD phenotype, rather than relying on potentially erroneous diagnostic boundaries, should be sought in future studies to improve signal to noise. This could be achieved using an endophenotype approach or through use of existing imaging or cerebrospinal fluid biomarkers (66–68) as proxies for AD status.

Summary

Combining knowledge across polygenic scores, MVMR, and UVMR, our results identified two blood metabolites—glutamine and XL.HDL.FC—with evidence of protective effects on AD. The biological mechanisms underpinning the relationship between education, cognition, and AD remain elusive, with no evidence of mediation via any of our metabolites. However, the effect of education on AD was shown to be almost entirely driven by cognition, implying that methods aimed at increasing cognition either indirectly through education or directly via brain training could hold protective utility against AD risk. Disentangling wider, multimodal risk factors and understanding how these connect along the AD causal pathway will be an important future endeavor if we hope to appropriately inform treatment strategies. This study provides some important, initial pieces to this causal puzzle, offering biological and nonbiological sources of insight to feed into wider multimodal work.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was made possible only through generous funding from key funding bodies. JL is funded by the van Geest endowment fund and PP is funded by Alzheimer's Research UK (ARUK). MR is supported by the Medical Research Council (MRC) (Grant No. MC_UU_12019/3). This study represents independent research additionally funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. A proportion of data collection and sharing for this project was also funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant No. U01 AG024904). The ADNI is funded by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from Abbott; the Alzheimer's Association; the Alzheimer's Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California San Diego. ADNI data are disseminated by the Laboratory of Neuro Imaging at the University of California, Los Angeles. For GERAD1 data, Cardiff University was

supported by the Wellcome Trust, MRC, ARUK, and the Welsh Assembly Government. Cambridge University and King's College London acknowledge support from the MRC. ARUK supported sample collections at the South West Dementia Bank and the Universities of Nottingham, Manchester, and Belfast. The Belfast group acknowledges support from the Alzheimer's Society, Ulster Garden Villages, Northern Ireland R&D Office, and the Royal College of Physicians/Dunhill Medical Trust. The MRC and Mercer's Institute for Research on Ageing supported the Trinity College group. The South West Dementia Brain Bank acknowledges support from Bristol Research into Alzheimer's and Care of the Elderly. The Charles Wolfson Charitable Trust supported the OPTIMA group. Washington University was funded by National Institutes of Health (NIH) grants, the Barnes Jewish Foundation, and the Charles and Joanne Knight Alzheimer's Research Initiative. Patient recruitment for the MRC Prion Unit/University College London (UCL) Department of Neurodegenerative Disease collection was supported by the UCL Hospitals/UCL Biomedical Centre and NIHR Queen Square Dementia Biomedical Research Unit. LASER-AD was funded by Lundbeck SA. The Bonn group was supported by the German Federal Ministry of Education and Research, Competence Network Dementia and Competence Network Degenerative Dementia, and Alfried Krupp von Bohlen und Halbach-Stiftung. The Genetic and Environmental Risk for Alzheimer's Disease (GERAD1) Consortium also used samples ascertained by the National Institute of Mental Health Alzheimer's Disease Genetics Initiative. The i-Select chip was funded by the French National Foundation on Alzheimer's disease and related disorders. The European Alzheimer's Disease Initiative was supported by a LABEX (Laboratory of Excellence Program Investment for the Future) DISTALZ grant, Institut National de la Santé et de la Recherche Médicale, Institut Pasteur de Lille, Université de Lille 2, and the Lille University Hospital. The GERAD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium was supported by the MRC (Grant No. 503480), ARUK (Grant No. 503176), the Wellcome Trust (Grant No. 082604/2/07/Z), and German Federal Ministry of Education and Research: Competence Network Dementia Grant Nos. 01GI0102, 01GI0711, and 01GI0420. The Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium was partly supported by NIH / NIA Grant No. R01 AG033193; NIA Grant No. AG081220; AGES Contract No. N01-AG-12100; National Heart, Lung, and Blood Institute Grant No. R01 HL105756; the Icelandic Heart Association; and the Erasmus Medical Center and Erasmus University. The Alzheimer's Disease Genetics Consortium was supported by NIH / NIA Grant Nos. U01 AG032984, U24 AG021886, and U01 AG016976; and Alzheimer's Association Grant No. ADGC-10-196728.

Data used in preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/ wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. The members of ADNI are listed in Supplement 1. Data used in the preparation of this article were obtained from the GERAD1 Consortium. As such, the investigators within the GERAD1 consortia contributed to the design and implementation of GERAD1 and/or provided data but did not participate in analysis or writing of this report. A full list of GERAD1 investigators can be found in Supplement 1.

JL performed all data preprocessing, statistical analyses, and manuscript writing. PP supervised the study, provided statistical input, and reviewed the manuscript. RG provided statistical input and reviewed the manuscript. SWC aided polygenic risk scoring statistical insight and back-conversion of R^2 statistics and reviewed the manuscript. CH aided in establishing an appropriate tool for correlated multiple test corrections and reviewed the manuscript. DA and LV provided ANM data and reviewed the manuscript. MR, PC, CL-Q, and RD provided background support and reviewed the manuscript.

We thank members of the Statistical Genetics Unit at the Social, Genetic, and Developmental Psychiatry Centre, King's College London, for their continued support and discussion on methodological aspects of this study. We also thank all providers of summary genomic data for use within this study. For the provision of the International Genomics of Alzheimer's Project (IGAP), IGAP investigators within contributed to the design and implementation of IGAP and provided data but did not participate in analysis or writing of this report. The IGAP was made possible by the generous participation of the control subjects, the patients, and their families.

A previous version of this article was published as a preprint on medRxiv: https://www.medrxiv.org/content/10.1101/2021.02.12.21251409v1.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Institute of Psychology, Psychiatry and Neuroscience (JL, RG, CH, DA, LV, RD, PP) and the Institute of Pharmaceutical Science (CL-Q), King's College London; NIHR Maudsley Biomedical Research Centre (RG, CH, RD), South London and Maudsley NHS Foundation Trust; the MRC Unit for Lifelong Health and Ageing at UCL (MR), University College London; and the Farr Institute of Health Informatics Research (RD), UCL Institute of Health Informatics, London, United Kingdom; Department of Genetics and Genomic Sciences (SWC), Icahn School of Medicine at Mount Sinai, New York, New York; Steno Diabetes Center (CLQ), Copenhagen, and the National Centre for Register-based Research (CH), Aarhus University, Aarhus, Denmark; the Center for Age-Related Medicine (DA), Stavanger University Hospital, Stavanger, Norway; and the Department of Psychiatry (PS), University of Hong Kong, Hong Kong, China.

Address correspondence to Petroula Proitsi, Ph.D., at petroula.proitsi@kcl.ac.uk.

Received May 17, 2021; revised Jun 18, 2021; accepted Jul 7, 2021.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsgos.2021.07.010.

REFERENCES

1. Prince M, Comas-Herrera A, Knapp M, Guerchet M, Karagiannidou M (2016): World Alzheimer Report 2016: Improving Healthcare for People Living With Dementia: Coverage, Quality and Costs Now and in the Future. London: Alzheimer's Disease International.
2. Sharp ES, Gatz M (2011): Relationship between education and dementia: An updated systematic review. Alzheimer Dis Assoc Disord 25:289–304.
3. Yeo RA, Arden R, Jung RE (2011): Alzheimer's disease and intelligence. Curr Alzheimer Res 8:345–353.
4. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, *et al.* (2020): Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 396:413–446.
5. Anderson EL, Howe LD, Wade KH, Ben-Shlomo Y, Hill WD, Deary IJ, *et al.* (2020): Education, intelligence and Alzheimer's disease: Evidence from a multivariable two-sample Mendelian randomization study. Int J Epidemiol 49:1163–1172.
6. Larsson SC, Traylor M, Malik R, Dichgans M, Burgess S, Markus HS (2017): Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. BMJ 359:j5375.
7. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, *et al.* (2019): Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet 51:404–413.
8. Raghavan NS, Vardarajan B, Mayeux R (2019): Genomic variation in educational attainment modifies Alzheimer disease risk. Neurol Genet 5:e310.
9. Enche Ady CNA, Lim SM, Teh LK, Salleh MZ, Chin A-V, Tan MP, *et al.* (2017): Metabolomic-guided discovery of Alzheimer's disease biomarkers from body fluid. J Neurosci Res 95:2005–2024.
10. Brink-Jensen K, Bak S, Jørgensen K, Ekstrøm CT (2013): Integrative analysis of metabolomics and transcriptomics data: A unified model framework to identify underlying system pathways. PLoS One 8: e72116.
11. Snowden S, Dahlén S-E, Wheelock CE (2012): Application of metabolomics approaches to the study of respiratory diseases. Bioanalysis 4:2265–2290.

12. Whiley L, Sen A, Heaton J, Proitsi P, García-Gómez D, Leung R, *et al.* (2014): Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. Neurobiol Aging 35:271–278.
13. Proitsi P, Kim M, Whiley L, Simmons A, Sattlecker M, Velayudhan L, *et al.* (2017): Association of blood lipids with Alzheimer's disease: A comprehensive lipidomics analysis. Alzheimers Dement 13:140–151.
14. Wang G, Zhou Y, Huang F-J, Tang H-D, Xu X-H, Liu J-J, *et al.* (2014): Plasma metabolite profiles of Alzheimer's disease and mild cognitive impairment. J Proteome Res 13:2649–2658.
15. González-Domínguez R, García-Barrera T, Gómez-Ariza JL (2015): Metabolite profiling for the identification of altered metabolic pathways in Alzheimer's disease. J Pharm Biomed Anal 107:75–81.
16. Smith GD, Ebrahim S (2003): 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 32:1–22.
17. Sanderson E, Davey Smith G, Windmeijer F, Bowden J (2019): An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol 48:713–727.
18. Burgess S, Thompson DJ, Rees JMB, Day FR, Perry JR, Ong KK (2017): Dissecting causal pathways using mendelian randomization with summarized genetic data: Application to age at menarche and risk of breast cancer. Genetics 207:481–487.
19. Lord J, Jermy B, Green R, Wong A, Xu J, Legido-Quigley C, *et al.* (2021): Mendelian randomization identifies blood metabolites previously linked to midlife cognition as causal candidates in Alzheimer's disease. Proc Natl Acad Sci U S A 118:e2009808118.
20. Proitsi P, Kuh D, Wong A, Maddock J, Bendayan R, Wulaningsih W, *et al.* (2018): Lifetime cognition and late midlife blood metabolites: Findings from a British birth cohort. Transl Psychiatry 8:203.
21. Kettunen J, Demirkan A, Würtz P, Draisma HHM, Haller T, Rawal R, *et al.* (2016): Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun 7:11122.
22. Weiner MW, Aisen PS, Jack CR, Jagust WJ, Trojanowski JQ, Shaw L, *et al.* (2010): The Alzheimer's disease neuroimaging initiative: Progress report and future plans. Alzheimers Dement 6:202–211.e7.
23. Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H, *et al.* (2009): AddNeuroMed–The European collaboration for the discovery of novel biomarkers for Alzheimer's disease. Ann N Y Acad Sci 1180:36–46.
24. Sattlecker M, Kiddle SJ, Newhouse S, Proitsi P, Nelson S, Williams S, *et al.* (2014): Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. Alzheimers Dement 10:724–734.
25. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, *et al.* (2019): Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nat Genet 51:414–430.
26. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, *et al.* (2018): Gene discovery and polygenic prediction from a 1.1-million-person GWAS of educational attainment. Nat Genet 50:1112–1121.
27. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, *et al.* (2018): Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet 50:912–919.
28. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, *et al.* (2015): LD score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47:291–295.
29. Hill WD, Davies G, Harris SE, Hagenaars SP, Liewald DC, Penke L, *et al.* (2016): Molecular genetic aetiology of general cognitive function is enriched in evolutionarily conserved regions. Transl Psychiatry 6:e980.
30. Burgess S, Butterworth A, Thompson SG (2013): Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 37:658–665.
31. van der Harst P, Verweij N (2018): Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. Circ Res 122:433–443.
32. Bowden J, Davey Smith G, Haycock PC, Burgess S (2016): Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol 40:304–314.
33. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, *et al.* (2018): The MR-Base platform supports systematic causal inference across the human phenome. Elife 7:e34408.
34. Choi SW, O'Reilly PF (2019): PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience 8:giz082.
35. Viechtbauer W (2010): Conducting meta-analyses in R with the metafor package. J Stat Softw 36:1–48.
36. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V (2002): A comparison of methods to test mediation and other intervening variable effects. Psychol Methods 7:83–104.
37. Bowden J, Davey Smith G, Burgess S (2015): Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. Int J Epidemiol 44:512–525.
38. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, *et al.* (2019): Guidelines for performing Mendelian randomization investigations. Wellcome Open Res 4:186.
39. Zuber V, Colijn JM, Klaver C, Burgess S (2020): Selecting likely causal risk factors from high-throughput experiments using multivariable Mendelian randomization. Nat Commun 11:29.
40. Stobart JL, Anderson CM (2013): Multifunctional role of astrocytes as gatekeepers of neuronal energy supply. Front Cell Neurosci 7:38.
41. Tsepilov YA, Shin S-Y, Soranzo N, Spector TD, Prehn C, Adamski J, *et al.* (2015): Nonadditive effects of genes in human metabolomics. Genetics 200:707–718.
42. Scott WK, Medie FM, Ruffin F, Sharma-Kuinkel BK, Cyr DD, Guo S, *et al.* (2018): Human genetic variation in GLS2 is associated with development of complicated Staphylococcus aureus bacteremia. PLoS Genet 14:e1007667.
43. Shin S-Y, Fauman EB, Petersen A-K, Krumsiek J, Santos R, Huang J, *et al.* (2014): An atlas of genetic influences on human blood metabolites. Nat Genet 46:543–550.
44. Teumer A (2018): Common methods for performing Mendelian randomization. Front Cardiovasc Med 5:51.
45. van der Lee SJ, Teunissen CE, Pool R, Shipley MJ, Teumer A, Chouraki V, *et al.* (2018): Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies. Alzheimers Dement 14:707–722.
46. Chen J, Herrup K (2015): Chapter 70 - Glutamine as a potential neuroprotectant in Alzheimer's disease. In: Martin CR, Preedy VR, editors. Diet and Nutrition in Dementia and Cognitive Decline. San Diego: Academic Press, 761–771.
47. Madeira C, Vargas-Lopes C, Brandão CO, Reis T, Laks J, Panizzutti R, *et al.* (2018): Elevated glutamate and glutamine levels in the cerebrospinal fluid of patients with probable Alzheimer's disease and depression. Front Psychiatry 9:561.
48. Wilson PW, Garrison RJ, Castelli WP, Feinleib M, McNamara PM, Kannel WB (1980): Prevalence of coronary heart disease in the Framingham Offspring Study: Role of lipoprotein cholesterols. Am J Cardiol 46:649–654.
49. Ouimet M, Barrett TJ, Fisher EA (2019): HDL and reverse cholesterol transport. Circ Res 124:1505–1518.
50. Bardagjy AS, Steinberg FM (2019): Relationship between HDL functional characteristics and cardiovascular health and potential impact of dietary patterns: A narrative review. Nutrients 11:1231.
51. Button EB, Robert J, Caffrey TM, Fan J, Zhao W, Wellington CL (2019): HDL from an Alzheimer's disease perspective. Curr Opin Lipidol 30:224–234.
52. Zhang Q, Sidorenko J, Couvy-Duchesne B, Marioni RE, Wright MJ, Goate AM, *et al.* (2020): Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture. Nat Commun 11:4799.
53. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, *et al.* (2013): Meta-analysis of 74,046 individuals

identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 45:1452–1458.
54. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, *et al.* (2016): Genome-wide association study identifies 74 loci associated with educational attainment. Nature 533:539–542.
55. Hill WD, Marioni RE, Maghzian O, Ritchie SJ, Hagenaars SP, McIntosh AM, *et al.* (2019): A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Mol Psychiatry 24:169–181.
56. Zeljkovic A, Vekic J, Spasojevic-Kalimanovska V, Jelic-Ivanovic Z, Bogavac-Stanojevic N, Gulan B, *et al.* (2010): LDL and HDL subclasses in acute ischemic stroke: Prediction of risk and short-term mortality. Atherosclerosis 210:548–554.
57. de Oliveira J, Hort MA, Moreira ELG, Glaser V, Ribeiro-do-Valle RM, Prediger RD, *et al.* (2011): Positive correlation between elevated plasma cholesterol levels and cognitive impairments in LDL receptor knockout mice: Relevance of cortico-cerebral mitochondrial dysfunction and oxidative stress. Neuroscience 197:99–106.
58. Kovács KR, Bajkó Z, Szekeres CC, Csapó K, Oláh L, Magyar MT, *et al.* (2014): Elevated LDL-C combined with hypertension worsens subclinical vascular impairment and cognitive function. J Am Soc Hypertens 8:550–560.
59. Imamura T, Doi Y, Arima H, Yonemoto K, Hata J, Kubo M, *et al.* (2009): LDL cholesterol and the development of stroke subtypes and coronary heart disease in a general Japanese population: The Hisayama study. Stroke 40:382–388.
60. Parthasarathy V, Frazier DT, Bettcher BM, Jastrzab L, Chao L, Reed B, *et al.* (2017): Triglycerides are negatively correlated with cognitive function in nondemented aging adults. Neuropsychology 31:682–688.
61. Farr SA, Yamada KA, Butterfield DA, Abdul HM, Xu L, Miller NE, *et al.* (2008): Obesity and hypertriglyceridemia produce cognitive impairment. Endocrinology 149:2628–2636.
62. Morley JE, Banks WA (2010): Lipids and cognition. J Alzheimers Dis 20:737–747.
63. Bernath MM, Bhattacharyya S, Nho K, Barupal DK, Fiehn O, Baillie R, *et al.* (2020): Serum triglycerides in Alzheimer disease: Relation to neuroimaging and CSF biomarkers. Neurology 94:e2088–e2098.
64. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, *et al.* (2018): HMDB 4.0: The human metabolome database for 2018. Nucleic Acids Res 46:D608–D617.
65. Bature F, Guinn B-A, Pang D, Pappas Y (2017): Signs and symptoms preceding the diagnosis of Alzheimer's disease: A systematic scoping review of literature from 1937 to 2016. BMJ Open 7: e015746.
66. Dhiman K, Blennow K, Zetterberg H, Martins RN, Gupta VB (2019): Cerebrospinal fluid biomarkers for understanding multiple aspects of Alzheimer's disease pathogenesis. Cell Mol Life Sci 76:1833–1863.
67. Zetterberg H, Skillbäck T, Mattsson N, Trojanowski JQ, Portelius E, Shaw LM, *et al.* (2016): Association of cerebrospinal fluid neurofilament light concentration with Alzheimer disease progression. JAMA Neurol 73:60–77.
68. Bloudek LM, Spackman DE, Blankenburg M, Sullivan SD (2011): Review and meta-analysis of biomarkers and diagnostic imaging in Alzheimer's disease. J Alzheimers Dis 26:627–645.