

CROSS-COHORT VALIDATION OF METABOLOMIC PREDICTORS OF GLOBAL COGNITION IN NON-DEMENTED ADULTS USING ELASTIC NET REGRESSION

Vasiliki-Amaryllis Skyfa, Brandi Vollmer, Christopher Papandreou, Christian Habeck, Michal Beerli, Yaakov Stern, Nikolaos Scarmeas, Yian Gu

INTRODUCTION

- Data on metabolomic profiles associated with cognition across the adult life course remain limited and lack external validation.
- Identifying metabolite signatures of global cognition in non-demented populations may reveal early metabolic alterations relevant to cognitive aging trajectories.

METHODS

Discovery dataset

Hellenic Longitudinal Investigation of Aging and Diet (HELIAD)

- The participants in this cohort were selected through random sampling from Greek community-dwelling individuals over the age of 65 years.

Validation Datasets

Aiginition Longitudinal Biomarker

Investigation Of Neurodegeneration (ALBION)

- Recruited participants aged over 40 who are cognitively normal (CN) or have mild cognitive impairment (MCI) in Athens, Greece.

Reference Ability Neural Network study and the Cognitive Reserve study (CR/RANN)

- Included cognitively normal adults aged 20-80 from the greater New York area.

Figure 1. Flowchart of HELIAD participants.

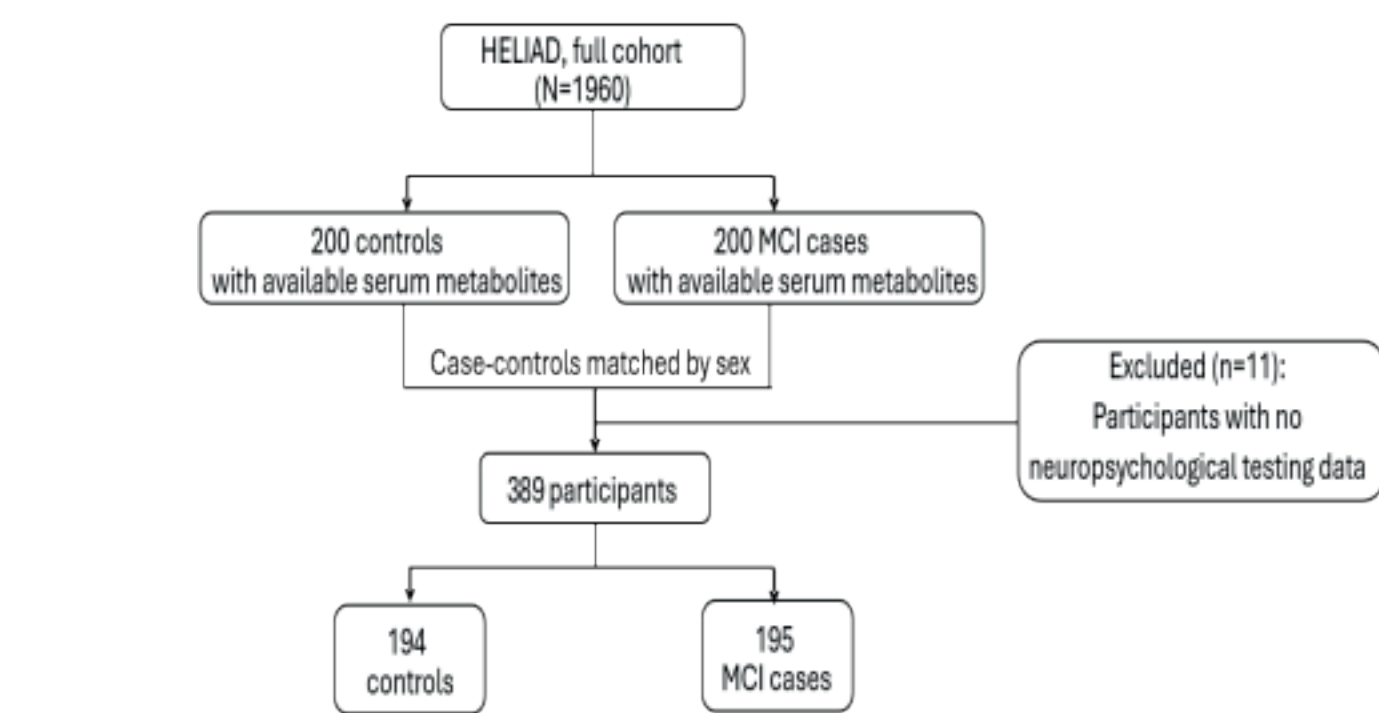


Figure 2. Flowchart of ALBION participants.

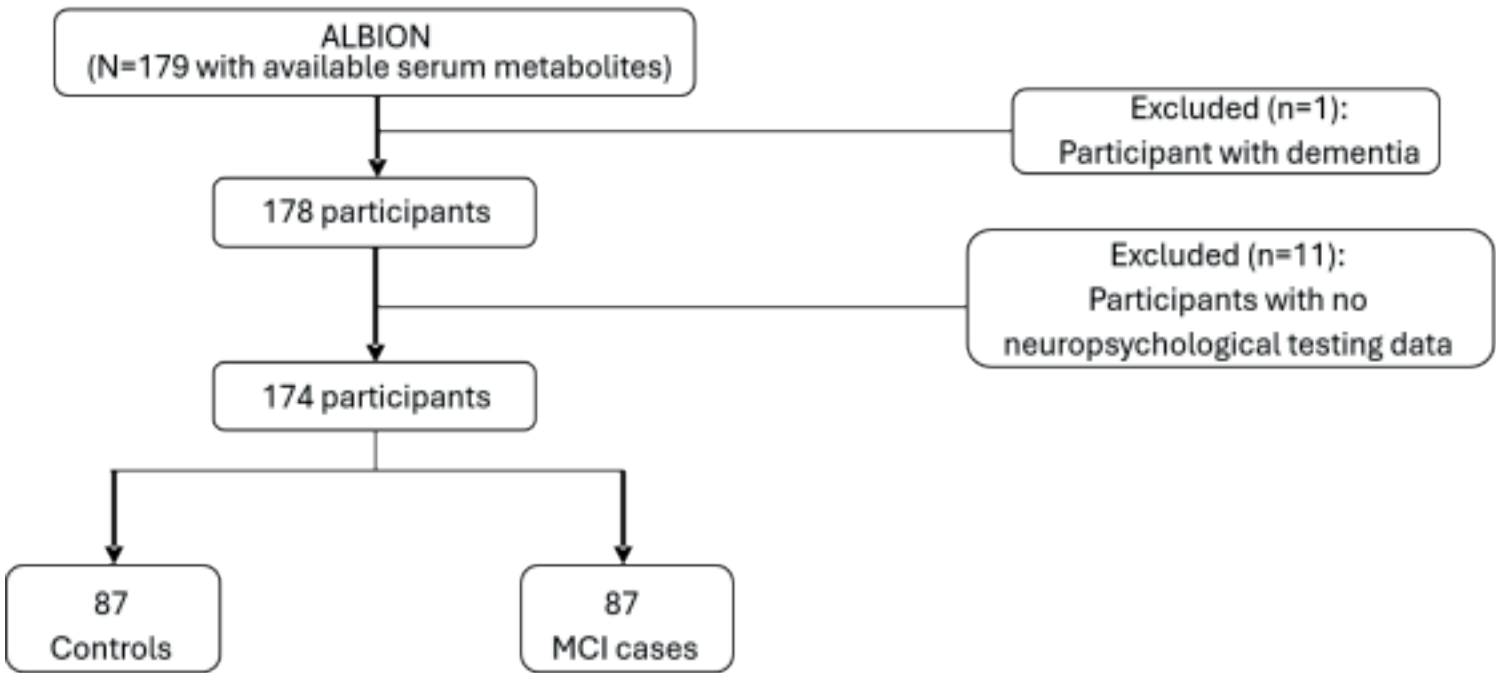
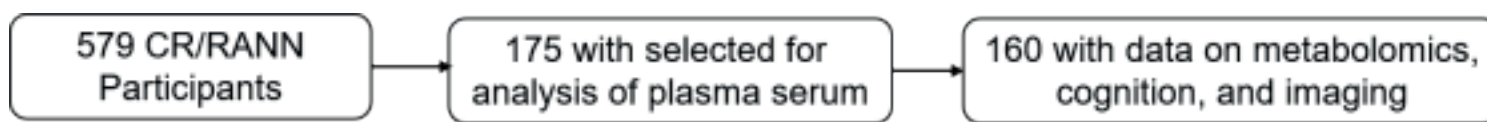


Figure 3. Flowchart of CR/RANN participants.



Data Collection

Metabolomics HELIAD & ALBION

- Utilized a targeted approach.
- Serum collected at baseline underwent liquid chromatography coupled to high-resolution mass spectrometry, liquid chromatography coupled to tandem mass spectrometry, and gas chromatography coupled to tandem mass spectrometry.

Metabolomics CR/RANN

- Utilized an untargeted approach.
- Plasma collected at baseline underwent liquid chromatography and high-resolution mass spectrometry.
- 241 Metabolites overlapped all 3 cohorts and were considered for analysis.

Composite Cognition Score

Global cognitive composite score was calculated by averaging all domain-level z-scores.

Analysis

Identification of metabolites and construction of the signature

- Standardized Composite Cognitive Score as the outcome and 241 inverse-normal transformed metabolites as predictors.
- 5 independent elastic net regression models for feature selection, within a 10-fold cross-validation framework.

- Keep only the 20 metabolites appearing in all 5 models.
- 5 independent Ridge regression models with the 20 selected metabolites within a 10-fold cross-validation framework, for coefficients estimation.

Evaluation of predictive performance of the signature

- Due to limited transferability of the HELIAD-derived coefficients, five independent Ridge models were trained in each cohort (ALBION, CR/RANN), employing 10-fold cross-validation to determine the optimal λ that minimized the RMSE.

The metabolic pathways in which the metabolites were involved were identified using MetaboAnalyst.

RESULTS

Table 1. Study sample characteristics.

	Discovery Dataset HELIAD (N=389)		Testing Dataset ALBION (N=174)		Testing Dataset CR/RANN (N=160)	
Characteristic	Mean or N	SD or %	Mean or N	SD or %	Mean or N	SD or %
Age (Years)	74.79	5.14	64.14	9.39	57	16.2
Education (Years)	5.61	3.72	13.37	3.88	15.9	2.4
Female	227	58.4%	116	66.7%	84	52.1%
Male	162	41.6%	58	33.3%	77	47.9%
Composite Cognitive Score	-0.69	0.78	-0.44	0.91		

Figure 4. Regression coefficients of the metabolite signature for all cohorts.

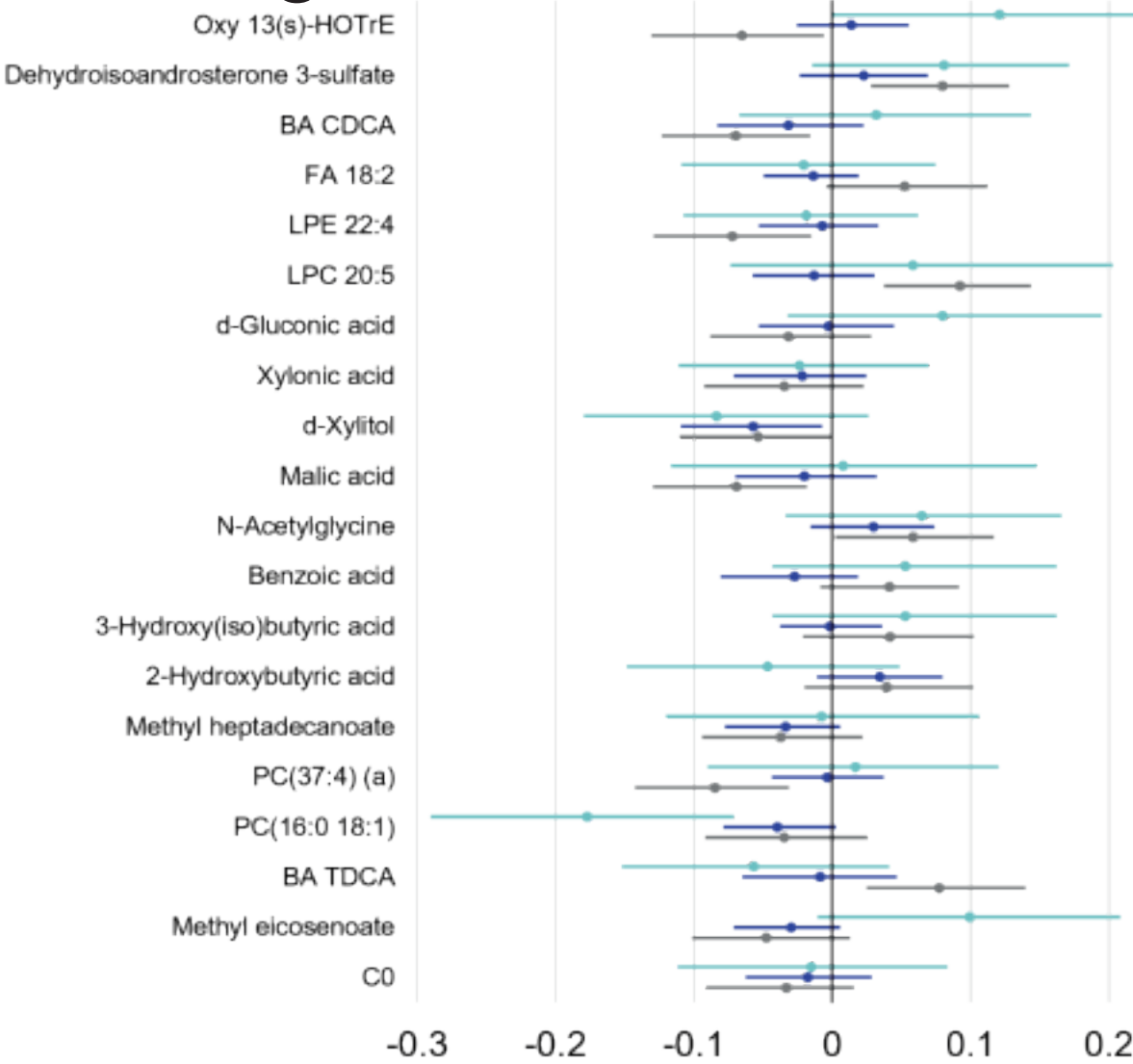


Table 2. Performance metrics for optimized models in each cohort.

	HELIAD	ALBION	CR/RANN
RMSE	0.686	0.866	1.066
R-Squared	0.234	0.109	0.165
Pearson	0.484	0.330	0.407
Lambda	0.182	1.546	0.576

Pathway analysis

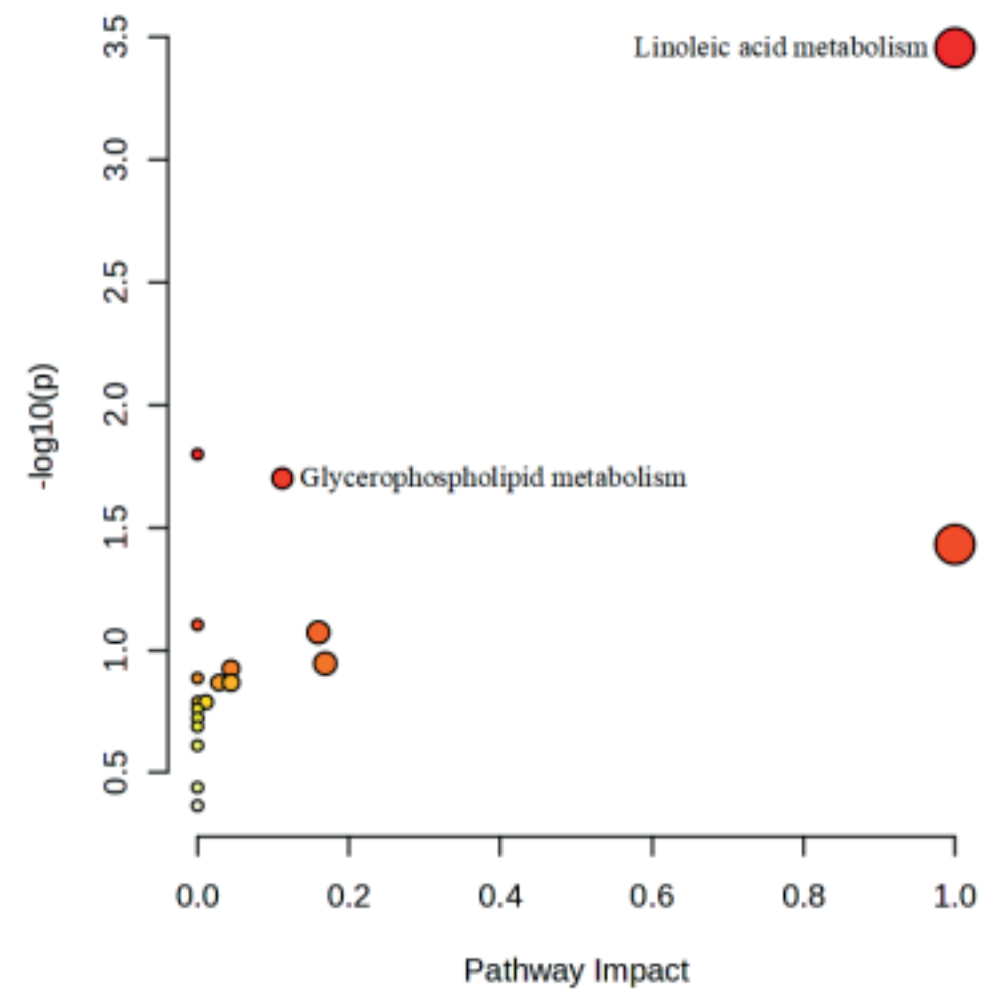


Figure 5. Metabolic pathway enrichment analysis.

- Linoleic acid metabolism showed the strongest association ($p = 0.00043$, FDR = 0.0344), driven in part by phosphatidylcholine and linoleate (FA 18:2).
- Glycerophospholipid metabolism was also enriched ($p = 0.0241$, FDR = 0.962), also driven by phosphatidylcholine though the false discovery rate suggests a weaker overall significance.

CONCLUSIONS

- A reproducible 20-metabolite signature associated with global cognition was identified.
- External validation suggested that the underlying metabolite pattern was preserved across cohorts despite differences in population characteristics and metabolomic platforms.
- These findings support the potential of metabolomics to identify early markers of cognitive vulnerability and may contribute to precision medicine approaches in cognitive aging and neurodegeneration.

CONTACT

ns257@cumc.columbia.edu

papchris@hmu.gr

amaryllis.skyfa@hotmail.gr