DOI: 10.1002/alz70861_108661

BIOMARKERS

PODIUM PRESENTATION

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

BIOMARKERS (NON-NEUROIMAGING)

Identification of an Alzheimer's Disease Biomarker Signature Using Reproducible Proteomics Data Mining

María Fernanda Zambrano-Astorga | Aldo Moreno-Ulloa

CICESE, Ensenada, BJ, Mexico

Correspondence
María Fernanda Zambrano-Astorga, CICESE, Ensenada, BJ, Mexico.
Email: mzambrano@cicese.edu.mx

Abstract

Background: Traditional diagnosis of Alzheimer's disease (AD) has focused on detecting beta-amyloid and phosphorylated tau, which only capture a portion of the disease's complexity. Proteomic studies have revealed broader molecular and cellular alterations in AD, highlighting the need for novel biomarkers that provide a more comprehensive understanding of the disease. In this study, we performed multi-cohort data mining of published proteomics datasets to identify a reproducible cerebrospinal fluid (CSF) biomarker signature for AD diagnosis.

Method: We systematically selected seven CSF proteomics studies (n = 1,350) comparing AD patients to non-AD subjects, using predefined inclusion and exclusion criteria. Dysregulated proteins were identified based on statistical significance ($p < 0.05$) and a fold change threshold ($|\log2\ FC| \geq 0.6$). To ensure reproducibility, a co-occurrence analysis was conducted to retain proteins dysregulated in at least two independent studies. The diagnostic performance of the resulting protein signature was evaluated using support vector machine-based ROC curve analysis across three independent proteomics-based cohorts (n = 813). We also compared the diagnostic performance of our signature with that of previously reported AD biomarker panels.

Result: Across the seven studies included (Table 1), we identified 61 dysregulated proteins that overlapped between at least two datasets. After extensive filtering, 11 proteins demonstrating consistent directionality of change were consolidated into a reproducible diagnostic signature, referred to as PPAV11 (Table 2). This signature demonstrated outstanding diagnostic performance, with ROC-AUC values ranging from 0.96 to 0.98 and sensitivity/specificity values > 0.92. Compared to nine previously published signatures, our panel consistently showed superior diagnostic accuracy across three independent proteomics datasets (Figure 1).

Conclusion: This study identified a reproducible and robust 11-protein CSF signature capable of accurately distinguishing individuals with AD from those without AD. Our biomarker panel outperformed several previously published signatures in diagnostic

metrics. These findings highlight the potential of proteomics-based approaches for improving disease diagnosis.

Data collection and sharing for ADNI cohort dataset was funded by the National Institutes of Health Grant U01 AG024904 and Department of Defense award number W81XWH-12-2-0012.

Table 1. Overview of Study Demographics for PPAV11 Protein Signature Discovery and Validation

Cohorts	Total subjects, n	Non-AD (%)	AD (%)	Age, years	Females (%)	Males (%)	Black/African American (%)	Caucasian/ White (%)	Another group (%)
DISCOVERY									
Bader et al., 2020	98	43 (43.88)	55 (56.12)	72.42±7.43	48 (48.98)	50 (51.02)	NA	NA	NA
Dammer et al., 2024	300	140 (46.66)	160 (53.34)	NA	189 (63.0)	111 (37.0)	29 (9.67)	268 (89.33)	3 (1.00)
Higginbotham et al., 2020	437	217 (49.67)	220 (50.33)	66.51±12.40	252 (57.67)	185 (42.33)	74 (16.93)	356 (81.47)	7 (1.60)
Liu et al., 2023	60	30 (50.0)	30 (50.0)	65.4±11.34	27 (45.0)	33 (55.0)	NA	NA	NA
Sathe et al., 2019	10	5 (50.0)	5 (50.0)	NA	NA	NA	NA	NA	NA
van der Ende et al., 2023	425	195 (45.88)	230 (54.12)	62.0±10.0	174 (40.94)	251 (59.06)	NA	NA	NA
Wang et al., 2020	20	9 (45.0)	11 (55.0)	NA	NA	NA	NA	NA	NA
VALIDATION									
ADNI	317	145 (45.74)	172 (54.26)	74.78±7.19	141 (44.48)	176 (55.52)	10 (3.15)	302 (95.27)	5 (1.58)
Bangs et al. 2025	431	213 (49.42)	218 (50.58)	66.37±8.41	267 (61.95)	164 (38.05)	111 (25.75)	318 (73.78)	2 (0.46)
Johnson et al., 2020	65	32 (49.23)	33 (50.77)	67.64 ±8.34	32 (49.23)	26 (40.00)	NA	NA	NA

Abbreviations: AD, Alzheimer's Disease; Non-AD, non-Alzheimer's Disease; n, number; NA, Not available; ADNI, The Alzheimer's Disease Neuroimaging Initiative. Continuous data is expressed as mean ± standard deviation.

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

Table 2. Comparative Overview of Proteomics Datasets and Published Protein Signatures Utilized for PPAV11 Validation

Cohorts		ADNI: Non-AD n = 145, AD n = 172		Bangs et al., 2025: Non-AD n = 213, AD n = 218		Johnson et al., 2020: Non-AD n = 32, AD n = 33	
Comparative signatures	Original protein signature (n)	Proteins found and tested in ADNI cohort (n)	Genes ID	Proteins found and tested in Bangs cohort (n)	Genes ID	Proteins found and tested in Johnson cohort (n)	Genes ID
This study-PPAV11	11	11	CHI3L1, CYCS, DDAH1, GDA, LRRC4B, NPTX2, PKM, SMOC1, SPON1, YWHAG, YWHAZ	11	CHI3L1, CYCS, DDAH1, GDA, LRRC4B, NPTX2, PKM, SMOC1, SPON1, YWHAG, YWHAZ	11	CHI3L1, CYCS, DDAH1, GDA, LRRC4B, NPTX2, PKM, SMOC1, SPON1, YWHAG, YWHAZ
Ali et al., 2025	10	10	CD248, HHIP, LRRN1, MIF, PPP1R1A, PPP3R1, SMOC1, TMED2, TMOD2, YWHAG	8	CD248, HHIP, LRRN1, MIF, PPP3R1, SMOC1, TMED2, YWHAG	4	CD248, LRRN1, SMOC1, YWHAG
Bader et al., 2020	5	5	ALDOC, IMPA1, MAPT, PKM, YWHAZ	5	ALDOC, IMPA1, MAPT, PKM, YWHAZ	3	ALDOC, PKM, YWHAZ
Del Campo et al., 2022	8	5	ABL1, ITGB2, MMP10, SPON2, TREM1	2	ITGB2, SPON2	0	
Liu et al., 2023	3	2	PRDX3, SRGN	2	PRDX3, SRGN	0	
Sathe et al., 2019	3	3	NPTX2, PKM, YWHAG	3	NPTX2, PKM, YWHAG	3	NPTX2, PKM, YWHAG
Shen et al., 2024	6	6	GFAP, NPTX2, PEA15, SMOC1, SMOC2, TNFRSF1B	5	GFAP, NPTX2, PEA15, SMOC1, SMOC2	2	NPTX2, SMOC1
Tao et al., 2024	19	14	ART3, CHRD, COL18A1, CXCL16, FRZB, GM2A, IGHM, MAN1C1, MGAT2, MGP, PCDHGC5, PCSK1N, PLTP, SHBG	19	ART3, B3GALNT1,C16orf89, CHRD, COL18A1, CXCL16, FRZB, GALNT7, GM2A, IGHM, MAN1C1, MAN2A2, MGAT2, MGP, PCDHGC5, PCSK1N PLTP, SCRG1, SHBG	16	ART3, B3GALNT1,C16orf89, CHRD, COL18A1, CXCL16, FRZB, GALNT7, GM2A, IGHM, MAN2A2,MGP, PCSK1N, PLTP, SHBG, SCRG1
van Zalm et al., 2023	3	3	ALDOA, PKM, LDHB	3	ALDOA, PKM, LDHB	3	ALDOA, PKM, LDHB
Wang et al., 2020	6	5	GFAP, MAPT, NTN1, PRDX3, SMOC1	4	GFAP, MAPT, PRDX3, SMOC1	1	SMOC1

Abbreviations: AD, Alzheimer's Disease; Non-AD, non-Alzheimer's Disease; n, number of individuals; ADNI, The Alzheimer's Disease Neuroimaging Initiative.

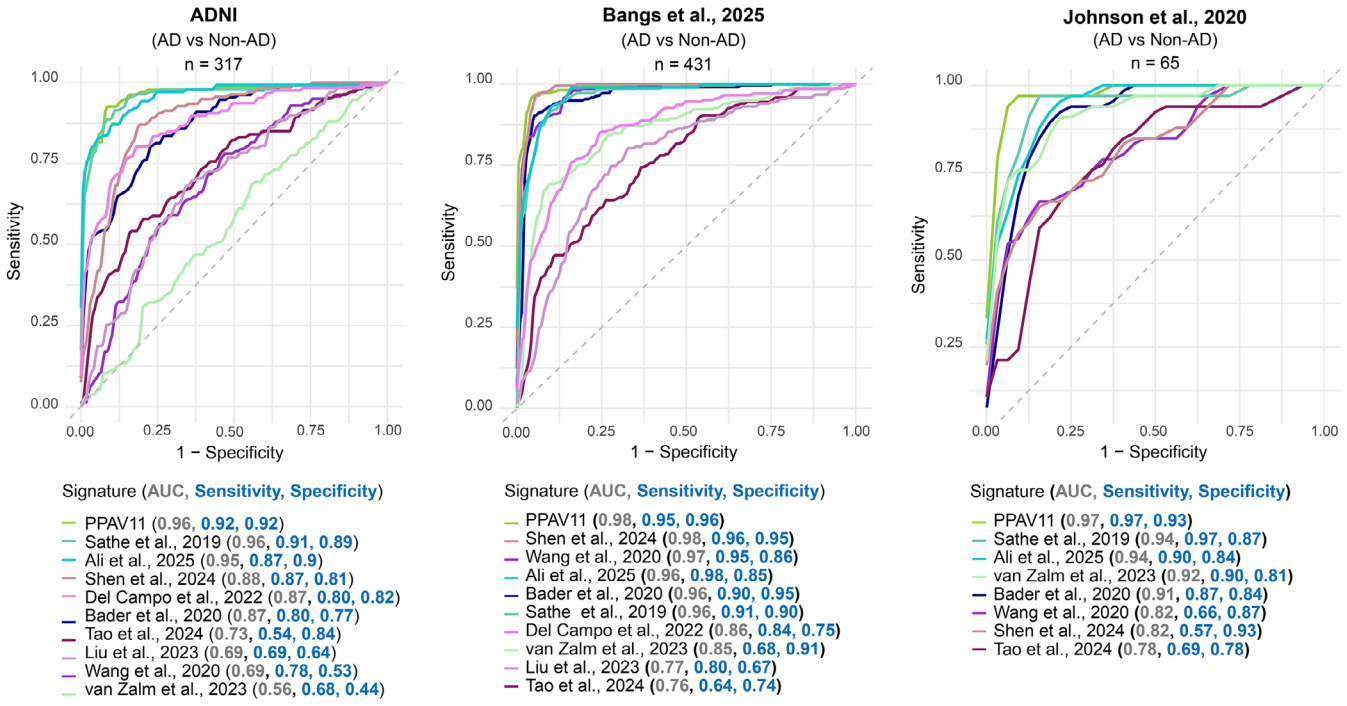

Figure 1. Receiver Operating Characteristic (ROC) Curve Analysis of PPAV11 and Published Protein Signatures for Alzheimer's Disease Diagnosis. Evaluation was conducted using pROC package in R Studio software (SVM algorithm).