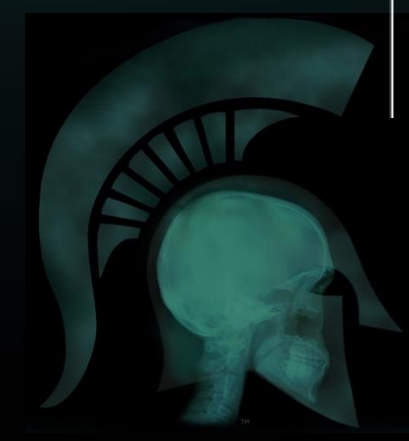




Dementia Etiology Classification Using NULISA Plasma Biomarkers and Machine Learning

Kelly N. DuBois^{1,2}, Subhamoy Pal², Amanada Cook Maher^{2,3}, Judith Heidebrink^{2,4}, Carol Persad³, Bruno Giordani^{2,3}, Benjamin M. Hampstead^{2,5,6}, Kelly M. Bakulski², David G. Morgan^{1,2,7}, Nicholas M. Kanaan^{1,2,7}

¹Department of Translational Neuroscience, College of Human Medicine, Michigan State University; ²Michigan Alzheimer's Disease Research Center, University of Michigan; ³Department of Psychiatry, University of Michigan; ⁴Department of Neurology, University of Michigan; ⁵Research Program on Cognition and Neuromodulation Based Interventions (RP-CNBI), Department of Psychiatry & Neurology, University of Michigan; ⁶VA Ann Arbor Healthcare System, Neuropsychology Section, Mental Health Service; ⁷Neuroscience Program, Michigan State University, East Lansing, MI, USA

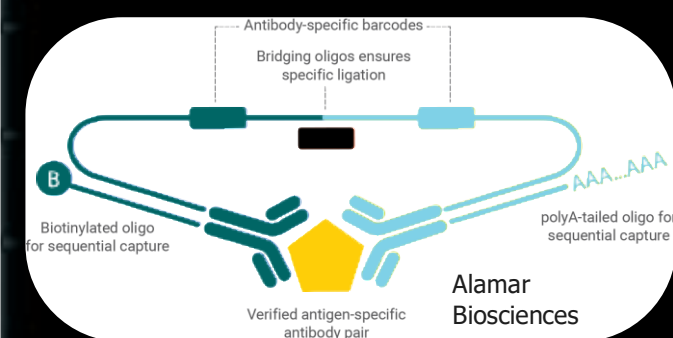


Link to Publication



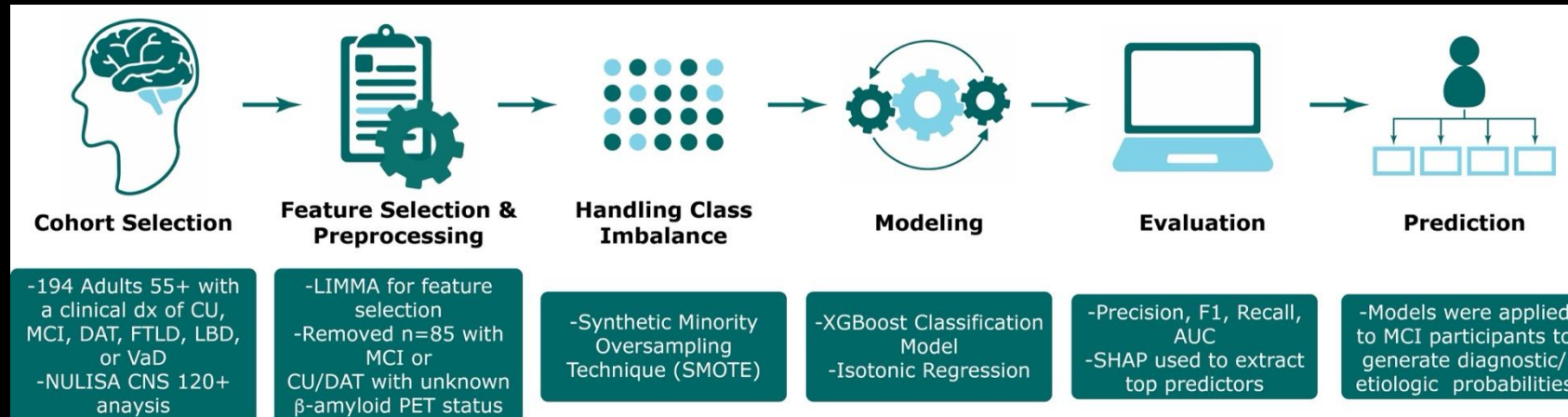
Background

The clinical and molecular overlap among dementia syndromes complicates accurate diagnosis, particularly for individuals with atypical or mixed presentations. Integrating multiplexed plasma proteomics with supervised machine-learning approaches offers a promising, minimally invasive method for differentiating underlying disease etiologies.



Methods

Plasma from 194 participants from the MADRC University of Michigan Memory and Aging Project was analyzed using the Nucleic acid Linked Immuno-Sandwich Assay (NULISA) Disease CNS 120+ panel. Differentially expressed proteins were identified among participants with etiologies of Alzheimer's disease, frontotemporal lobar degeneration, Lewy body dementia, and vascular dementia. These proteins were used to train supervised XGBoost models distinguishing each dementia etiology from all others. The resulting models were applied to participants with mild cognitive impairment to generate data-driven probabilities of disease etiology.

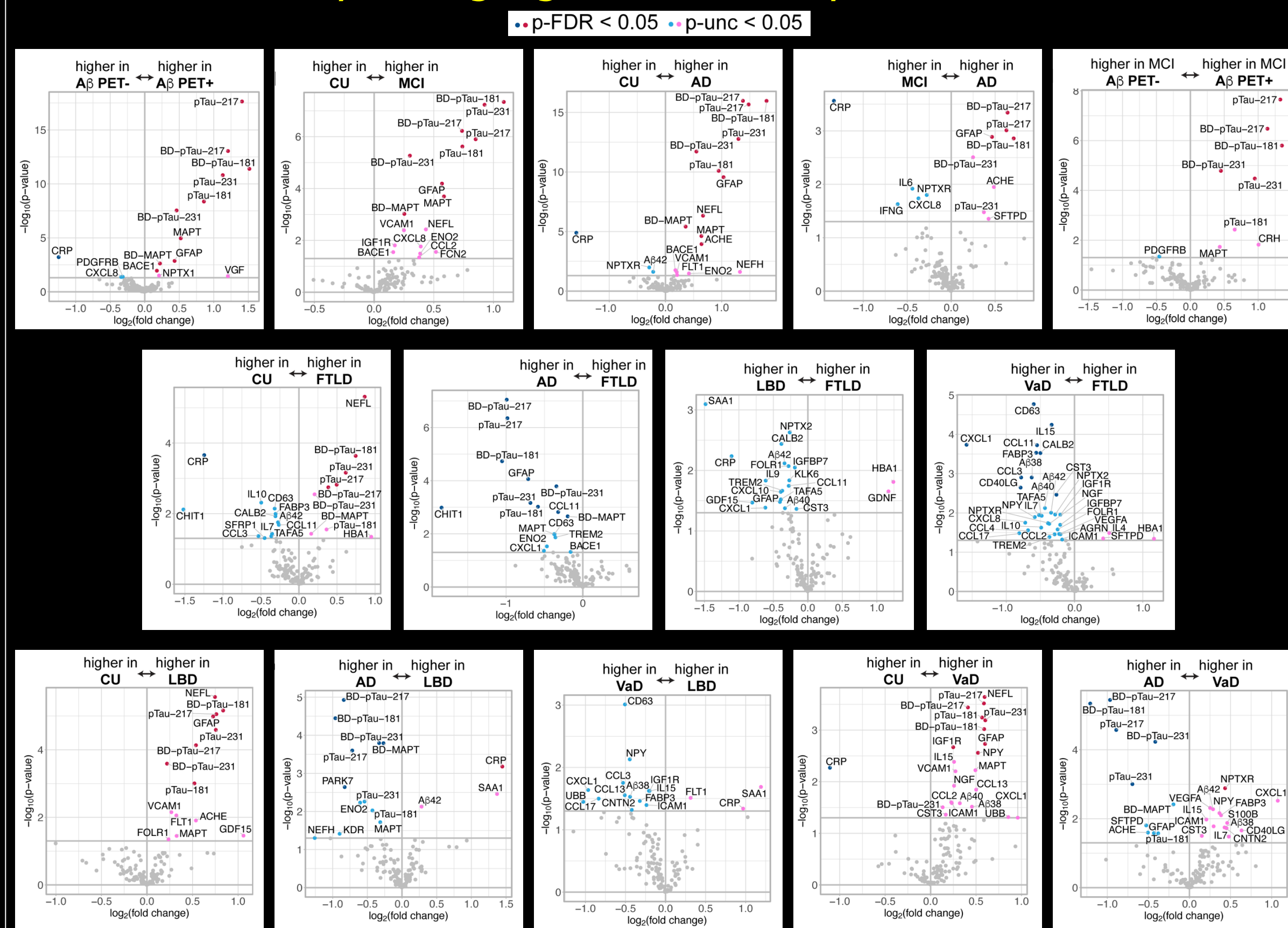


Participant Characteristics

Characteristics	Overall (N=194) Median (IQR/N%)	CU (N=35) Median (IQR/N%)	MCI (N=53) Median (IQR/N%)	AD (N=33) Median (IQR/N%)	FTLD (N=30) Median (IQR/N%)	LBD (N=26) Median (IQR/N%)	VaD (N=17) Median (IQR/N%)	p-value	Adjusted p-value
Age (years)	72.0 (67.3, 77.0)	73.0 (70.0, 76.0)	72.0 (69.0, 77.0)	70.0 (66.0, 76.0)	64.0 (61.3, 68.8)	77.0 (71.5, 80.0)	76.0 (71.0, 82.0)	<0.001***	<0.001***
Sex									
Male	100 (51.5%)	15 (42.9%)	29 (54.7%)	16 (48.5%)	16 (53.3%)	19 (73.1%)	5 (29.4%)	0.088	0.876
Years Education	16 (14, 18)	18 (15, 18)	16 (14, 18)	16 (23, 18)	16 (14, 17.3)	16 (14, 18)	14 (14, 16)	0.090	0.904
Missing	6	0	4	0	2	0	0		
APOE-ε4 allele (any)	88 (45.4%)	8 (22.9%)	23 (43.4%)	22 (66.7%)	10 (33.3%)	7 (26.9%)	8 (47.1%)	0.052	0.524
Missing	13	0	7	1	3	1	1		
Race									
White	138 (71.1%)	23 (65.7%)	34 (64.2%)	23 (69.7%)	30 (100%)	22 (84.6%)	6 (35.3%)	<0.001***	0.004**
Black/African American	55 (28.4%)	12 (34.3%)	19 (35.8%)	9 (28.1%)	0 (0%)	4 (15.4%)	11 (64.7%)		
CDR-Sum	1.0 (0.6, 3.9)	0 (0.0, 0.0)	0.5 (0.5, 1.0)	4.5 (2.5, 5.0)	4 (1.6, 6.0)	5.5 (2.0-8.9)	0.5 (0.5, 1.0)	<0.001***	<0.001***
MoCA Total	23 (18, 26)	28 (26, 29)	24 (22, 25)	17 (12, 19)	19 (11, 22)	21 (15, 25)	24 (21, 26)	<0.001***	<0.001***
Missing	7	0	4	1	1	1	0		
Aβ PET									
Positive	39	0	18	20	0	1	0	<0.001***	<0.001***
Negative	28	15	11	0	1	0	1		
Missing	127	20	24	13	29	25	16		

Numeric variables = median (IQR). Categorical variables = number (percent). P-values for continuous variables are from Dunn's post-hoc tests; p-values for categorical variables are from Pearson's Chi-squared test and adjusted p-value with Bonferroni corrections. CDR-Sum, Clinical Dementia Rating-Sum of Boxes; MoCA, Montreal Cognitive Assessment; CU, cognitively unimpaired; MCI, mild cognitive impairment; AD, Alzheimer's disease; FTLD, frontotemporal lobar degeneration; LBD, Lewy body dementia; VaD, vascular dementia.

LIMMA Analyses Highlight Differentially Abundant Proteins



Volcano plots demonstrate differentially abundant proteins. Dark dots (red and blue) represent comparisons with FDR-corrected $p < 0.05$, while light dots (pink and light blue) represent comparisons with uncorrected (unc) $p < 0.05$. Gray vertical line signifies 0 fold change, gray horizontal line signifies $p = 0.05$ (uncorrected).

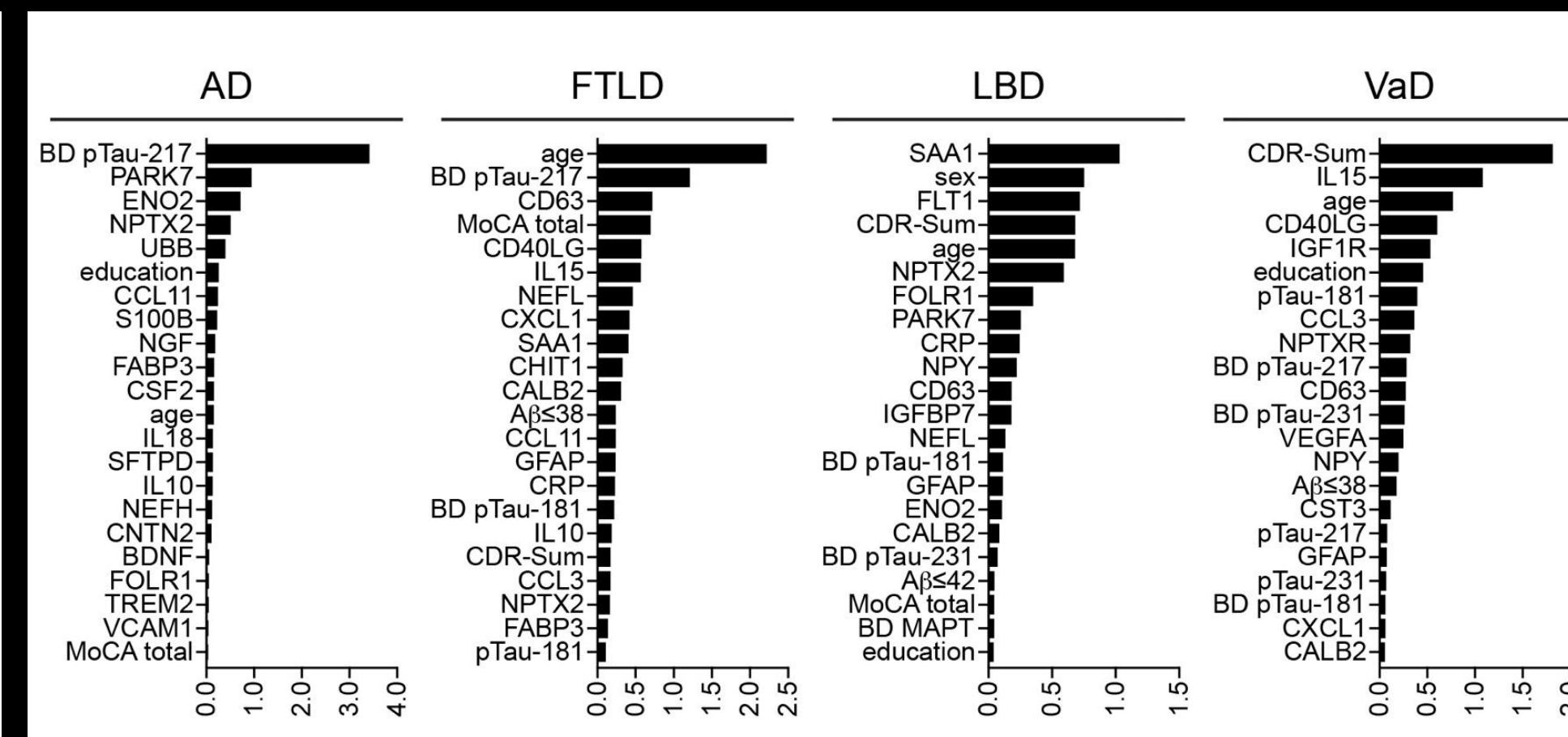
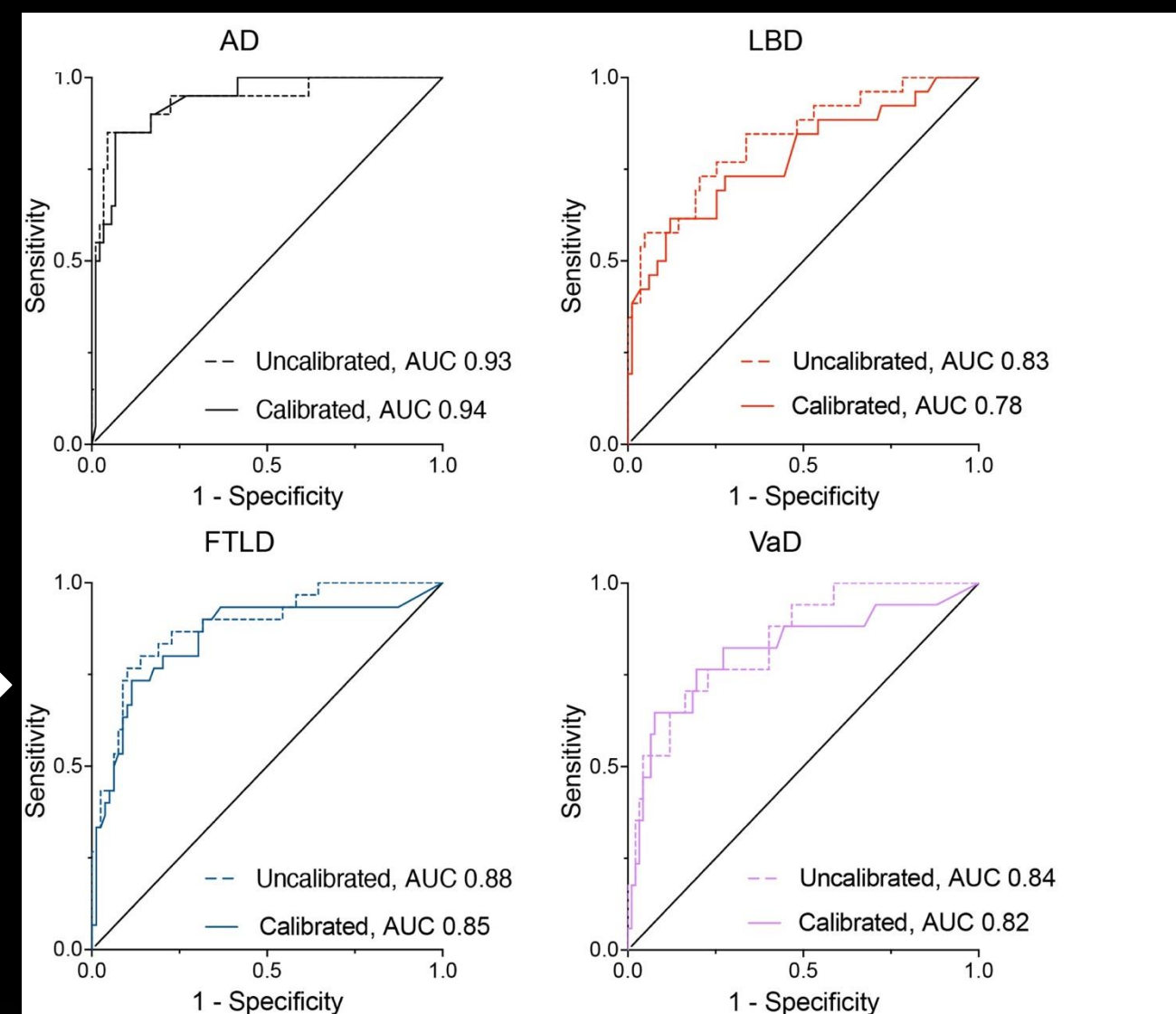
Abbreviations: Aβ, Amyloid beta; CU, cognitively unimpaired; MCI, mild cognitive impairment; AD, Alzheimer's disease; FTLD, frontotemporal lobar degeneration; LBD, Lewy body dementia; VaD, vascular dementia.

XGBoost ML Models Discriminate Among Dementia Etiologies

Models were adjusted for Age, Sex, Education (years), CDR-Sum, and MoCA total

Receiver operator characteristic curves for each one-versus-rest XGBoost model

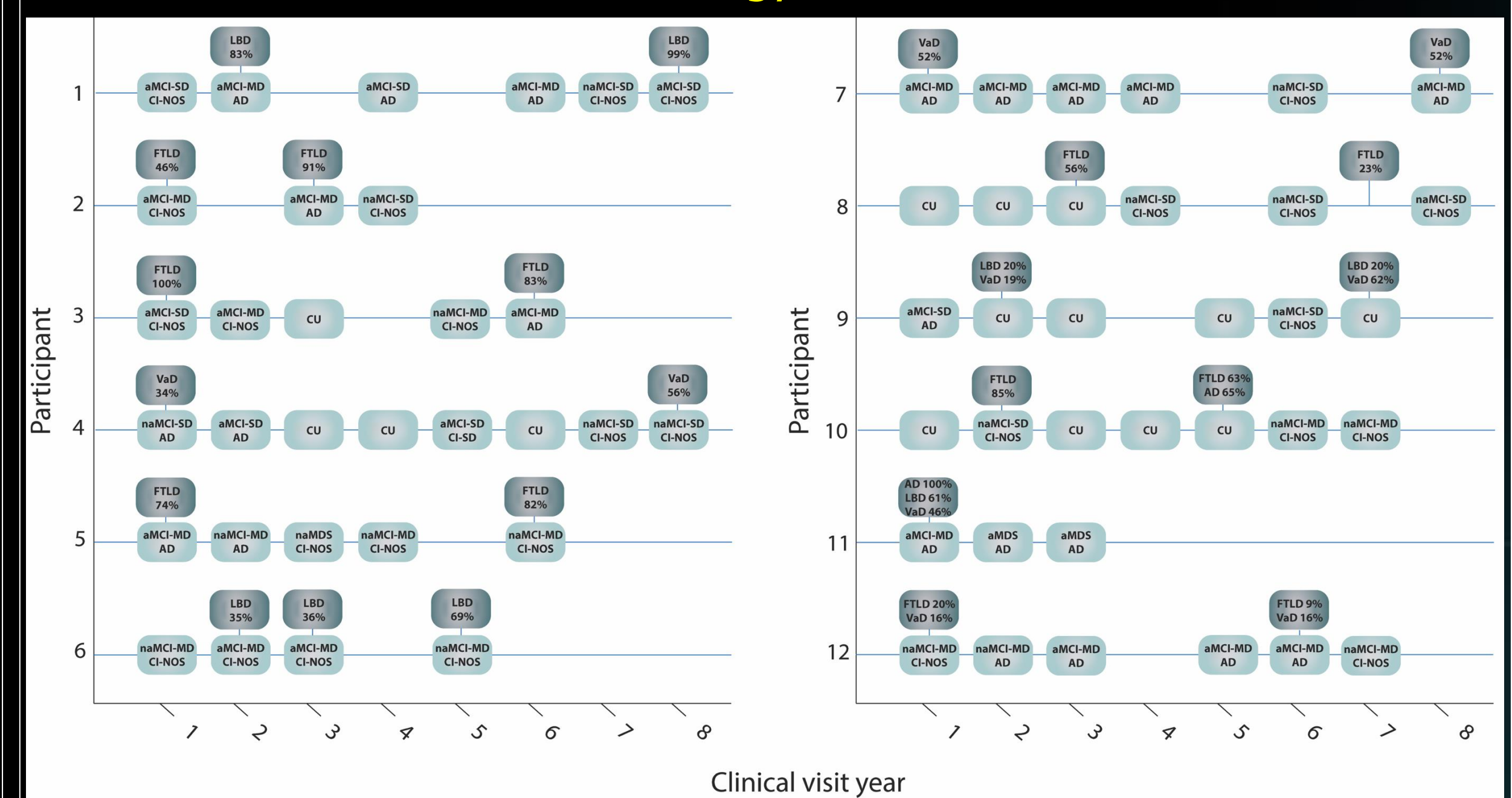
AUC: Area Under the Curve



SHAP Analysis Feature Importance

Abbreviations: AD, Alzheimer's disease; FTLD, frontotemporal lobar degeneration; LBD, Lewy body dementia; VaD, vascular dementia.

Model Etiology Predictions



Light boxes = clinical diagnosis and etiology. Dark boxes = etiology prediction derived from the model and probability. Abbreviations: aMCI, amnesic MCI; naMCI, non-amnesic MCI; SD, single domain; MD, multiple domains; CI, cognitive impairment; NOS, not otherwise specified; aMDS, amnesic multidomain dementia syndrome; naMDS, non-amnesic multidomain dementia syndrome.

Conclusions

- Etiology-specific protein patterns can be detected by NULISA plasma proteome analysis.
- XGBoost classifier machine-learning models can differentiate between etiology groups with high specificity (0.87-0.94) using differentially expressed NULISA proteins.
- Models identified the top discriminatory features for each etiology.
- Applications of models to participants with mild cognitive impairment yields plausible etiological predictions, highlighting potential diagnostic applications of this approach.

Future Directions & Support

This framework has the potential to augment clinical decision-making, improve diagnostic confidence, and support enrollment in disease-modifying therapy trials. Future studies will utilize a larger, more well-balanced cohort with neuropathology-confirmed etiologies to enable the inclusion of more demographic and clinical features in models and increase confidence.

This project was supported by NIA P30 AG072931, R01 AG068338, R01AG058724, and the Malbach-Smiley Endowment (MSU)