

Journal Club - Paper

- **Title:** AI-based Differential Diagnosis of Dementia Etiologies on Multimodal Data
- **Authors:** Chonghua Xue, et al.
- **Institutions:** Various, including Boston University
- **Journal:** Nature Medicine
- **Publishing Date:** June 2024
- **DOI:** <https://doi.org/10.1038/s41591-024-03118-z>
- **Contribution:** AI model to aid differential dementia diagnosis



Introduction: Presented Problems and Solutions

In this section, we will explore the:

Key problems highlighted
Solutions proposed

Problem 1: Overlapping Symptoms

Overlapping presentations of different dementias.

- Memory loss: Alzheimer's disease, vascular dementia, depression.
- Behavioral changes: Alzheimer's disease, frontotemporal dementia.
- Problems with executive functioning or language: Lewy body dementia or Alzheimer's disease.
- Similar idea from Stephen Wood's talk on Schizophrenia.

More difficult to diagnose.

Problem 2: Variability in Dementia Presentations

Heterogeneity in dementia symptoms and imaging complicates diagnosis.

- Different symptoms for the same dementia type (e.g., memory loss vs. language issues in Alzheimer's).
- Imaging variability confounded by comorbidities.
- Leads to inconsistent diagnostic criteria, misdiagnosis, or delayed diagnosis.

More difficult to diagnose.

Problem 3: Resource and Specialist Shortages

Limited access to specialists and tools in dementia diagnosis.

- Lancet Neurology: shortage of neurologists and geriatricians.
- Remote practices lack access to advanced imaging resources.

More difficult to diagnose.

Challenges

Misdiagnosis:

- Suitable for early-stage AD: anti-amyloid therapies.
- Increased risk with VD coexistence: leads to complications.
- Result: negative treatment outcomes.

Delayed diagnosis:

- Impact: critical loss of time for effective treatment.
- Consequence: worsens prognosis, accelerates patient decline.

Proposed AI Model

A multimodal AI model:

- Inputs: Demographics, medical history, neuropsychological assessments, multimodal neuroimaging.
- Outputs: Probabilities for each potential dementia etiology.
- Aim: A tool to enhance clinicians' abilities, reducing misdiagnosis and speeding up diagnosis.

Demo

For a demo, click here:

<https://huggingface.co/spaces/vkola-lab/nmed2024>

Methods: AI Model and Datasets

In this section, we will explore the:

- Model Architecture
- Datasets
- Training
- Handling of Missing Data

Inputs

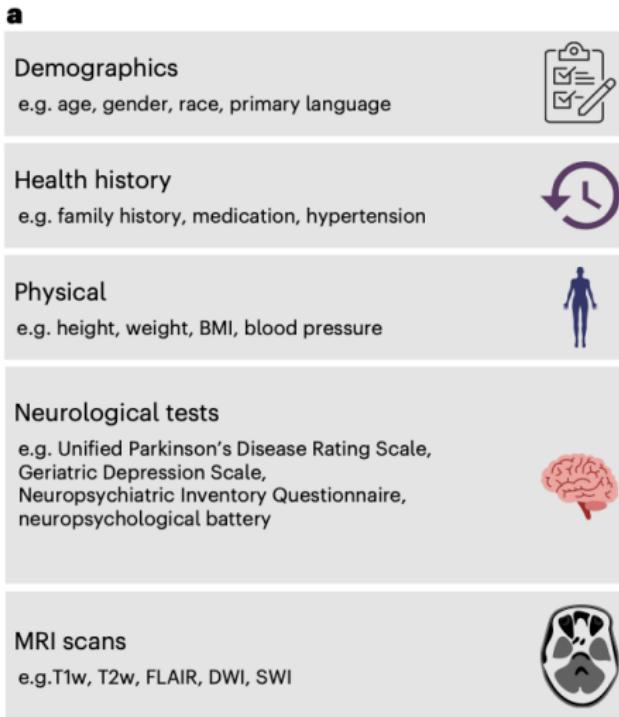


Figure: Diverse set of inputs

Model Outputs: Diagnostic Categories

Glossary 1	
Acronym	Description
NC	Normal cognition
MCI	Mild cognitive impairment
DE	Dementia
AD	Alzheimer's disease
LBD	Lewy body dementia, including dementia with Lewy bodies and Parkinson's disease dementia
VD	Vascular dementia, vascular brain injury and vascular dementia, including stroke
PRD	Prion disease including Creutzfeldt-Jakob disease
FTD	Frontotemporal lobar degeneration and its variants, including primary progressive aphasia, corticobasal degeneration and progressive supranuclear palsy, and with or without amyotrophic lateral sclerosis
NPH	Normal pressure hydrocephalus
SEF	Systemic and environmental factors including infectious diseases (HIV included), metabolic, substance abuse / alcohol, medications, systemic disease and delirium
PSY	Psychiatric conditions including schizophrenia, depression, bipolar disorder, anxiety and posttraumatic stress disorder
TBI	Moderate/severe traumatic brain injury, repetitive head injury and chronic traumatic encephalopathy
ODE	Other dementia conditions, including neoplasms, Down syndrome, multiple systems atrophy, Huntington's disease and seizures

- **Neurologist-Defined Categories:** Aligned with clinical management strategies.
- **Example 1:** LBD & Parkinson's disease: Movement disorder specialists.
- **Example 2:** VD & Stroke: Specialists in stroke and vascular diseases.

Figure: Outputs

Model Architecture

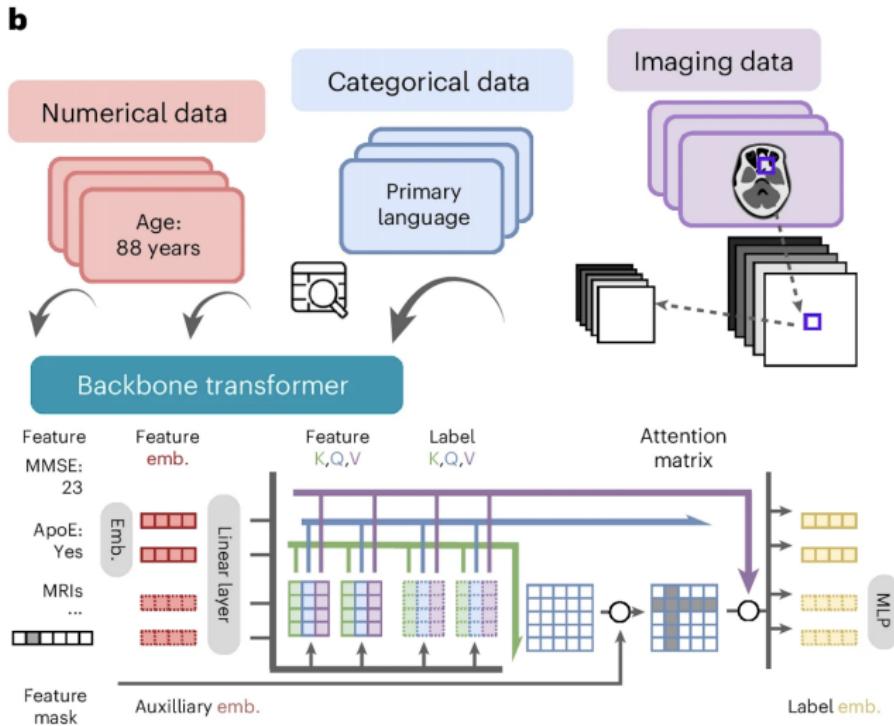


Figure: Transformer Backbone for multi-label classification problem

Datasets Overview

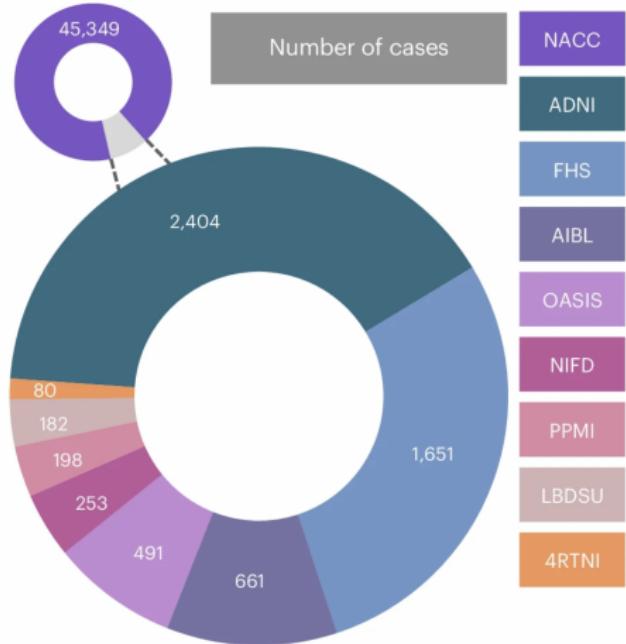


Figure: Datasets

- National Alzheimer's Coordinating Center (NACC) (N=45,349)
- Alzheimer's Disease Neuroimaging Initiative (ADNI) (N=2,404)
- FTD Neuroimaging Initiative (NIFD) (N=253)
- Parkinson's Progression Marker Initiative (PPMI) (N=198)
- Australian Imaging, Biomarker and Lifestyle Flagship Study of Ageing (AIBL) (N=661)
- Open Access Series of Imaging Studies-3 (OASIS) (N=491)
- 4 Repeat Tauopathy Neuroimaging Initiative (4RTNI) (N=80)
- Lewy Body Dementia Center for Excellence at Stanford University (LBDSU) (N=182)
- Framingham Heart Study (FHS) (N=1,651)

Data Set Info

Dataset (group)	Age (y), mean±s.d.	Male, n (%)	Education (y), mean±s.d.	Race (White, Black, Asian, American Indian, Pacific, multirace), n	CDR, mean±s.d.
NACC					
NC [n = 17,242]	71.25±11.16	6,009, 34.85%	15.83±2.98^	(13,266, 2541, 528, 109, 10, 575)^	0.05±0.15
MCI [n = 7,582]	73.72±9.81	3,615, 47.68%	15.16±3.45^	(5,708, 1185, 231, 53, 5, 276)^	0.45±0.18
AD [n = 16,131]	76.0±10.31	7,234, 44.85%	14.52±3.74^	(13,161, 1702, 354, 92, 10, 458)^	1.2±0.73
LBD [n = 1,913]	75.01±8.55	1,365, 71.35%	15.12±3.63^	(1,659, 128, 39, 17, 0, 37)^	1.29±0.78
VD [n = 1,919]	80.32±8.76	947, 49.35%	14.15±4.22^	(1,394, 332, 67, 2, 1, 68)^	1.22±0.74
PRD [n = 114]	60.07±10.36	62, 54.39%	14.8±3.33^	(93, 5, 5, 0, 1, 1)^	1.95±0.95
FTD [n = 2,898]	65.86±9.36	1,603, 55.31%	15.45±3.09^	(2,664, 69, 73, 4, 5, 39)^	1.2±0.83
NPH [n = 138]	79.1±9.24	69, 50.0%	15.0±3.28^	(119, 10, 4, 0, 0, 4)^	1.18±0.71
SEF [n = 808]	76.3±11.15	413, 51.11%	14.6±3.77^	(646, 95, 15, 5, 2, 31)^	1.11±0.7
PSY [n = 2,700]	73.74±10.78	1,102, 40.81%	14.13±4.12^	(2,163, 238, 59, 14, 5, 87)^	1.1±0.64
TBI [n = 265]	72.87±11.23	192, 72.45%	14.42±4.13^	(212, 27, 3, 2, 1, 11)^	1.11±0.69
ODE [n = 1,234]	72.94±12.14	654, 53.0%	14.5±3.78^	(1,046, 93, 28, 5, 4, 36)^	1.2±0.76
P value	<1.0 × 10 ⁻²⁰⁰	<1.0 × 10 ⁻²⁰⁰	<1.0 × 10 ⁻²⁰⁰	8.341 × 10 ⁻¹⁴⁵	<1.0 × 10 ⁻²⁰⁰
NIFD					
NC [n = 124]	63.21±7.27	56, 45.16%	17.48±1.87^	(89, 0, 0, 0, 0, 3)^	0.03±0.12^
FTD [n = 129]	63.66±7.33	75, 58.14%	16.18±3.29^	(109, 1, 1, 0, 0, 4)^	0.82±0.54^
P value	6.266 × 10 ⁻¹	5.246 × 10 ⁻²	2.606 × 10 ⁻⁴	6.531 × 10 ⁻¹	4.333 × 10 ⁻²⁸

Figure: Datasets: NACC, NIFD

Data Set Info (contd.)

Dataset (group)	Age (y), mean±s.d.	Male, n (%)	Education (y), mean±s.d.	Race (White, Black, Asian, American Indian, Pacific, multirace), n	CDR, mean±s.d.
PPMI					
NC [n = 171]	62.74±10.12	109, 63.74%	15.82±2.93	(163, 3, 2, 0, 0, 1)^\wedge	NA
MCI [n = 27]	68.04±7.32	22, 81.48%	15.52±3.08	(24, 1, 1, 0, 0, 1)	NA
P value	1.006×10^{-2}	1.115×10^{-1}	6.194×10^{-1}	2.910×10^{-1}	NA
AIBL					
NC [n = 480]	72.45±6.22	203, 42.29%	NA	NA	0.03±0.12
MCI [n = 102]	74.73±7.11	53, 51.96%	NA	NA	0.47±0.14
AD [n = 79]	73.34±7.77	33, 41.77%	NA	NA	0.93±0.54
P value	5.521×10^{-3}	1.887×10^{-1}	NA	NA	4.542×10^{-158}
OASIS					
NC [n = 424]	71.34±9.43	164, 38.68%	15.79±2.62^\wedge	(53, 18, 1, 0, 0, 0)^\wedge	0.0±0.02
MCI [n = 27]	75.04±7.25	14, 51.85%	15.19±2.76	(4, 1, 0, 0, 0, 0)^\wedge	0.52±0.09
AD [n = 32]	77.44±7.42	20, 62.5%	15.19±2.8	(8, 1, 0, 0, 0, 0)^\wedge	0.86±0.44
LBD [n = 4]	74.75±5.67	4, 100.0%	16.0±2.83	NA	1.0±0.0
FTD [n = 4]	64.25±8.61	3, 75.0%	16.5±2.96	(4, 0, 0, 0, 0, 0)	1.25±0.75
P value	7.789×10^{-4}	3.239×10^{-3}	5.507×10^{-1}	8.735×10^{-1}	2.855×10^{-169}

Figure: Datasets: PPMI, AIBL, OASIS

Data Set Info (contd.)

Dataset (group)	Age (y), mean±s.d.	Male, n (%)	Education (y), mean±s.d.	Race (White, Black, Asian, American Indian, Pacific, multirace), n	CDR, mean±s.d.
LBDSU					
NC [n = 134]	68.77±7.62	61, 45.52%	17.27±2.47^	NA	NA
MCI [n = 35]	70.16±8.41	26, 74.29%	16.6±2.58	NA	NA
LBD [n = 13]	73.42±7.81	8, 61.54%	16.77±2.15	NA	NA
P value	1.033×10^{-1}	7.863×10^{-3}	3.243×10^{-1}	NA	NA
4RTNI					
NC [n = 12]	68.08±4.92	5, 41.67%	15.45±2.57^	(12, 0, 0, 0, 0)	0.0±0.0
MCI [n = 31]	67.61±7.0	11, 35.48%	16.68±4.02	(25, 1, 2, 0, 1, 1)^	0.55±0.15
FTD [n = 37]	69.14±7.43	20, 54.05%	16.46±4.21	(31, 1, 0, 0, 1, 2)^	1.27±0.55
P value	6.691×10^{-1}	2.992×10^{-1}	6.843×10^{-1}	7.620×10^{-1}	5.700×10^{-16}
ADNI					
NC [n = 868]	72.7±6.57	383, 44.12%	16.51±2.52	(730, 92, 28, 2, 0, 12)^	0.0±0.04^
MCI [n = 1119]	72.77±7.65	648, 57.91%	15.97±2.75	(1,023, 56, 17, 2, 2, 13)^	0.5±0.06
AD [n = 417]	74.99±7.78	232, 55.64%	15.25±2.92	(383, 20, 10, 0, 0, 4)	0.77±0.27
P value	8.911×10^{-8}	3.090×10^{-09}	2.869×10^{-14}	2.828×10^{-5}	$<1.0 \times 10^{-200}$

Figure: Datasets: LBDSU, 4RTNI, ADNI

Data Set Info (contd.)

Dataset (group)	Age (y), mean±s.d.	Male, n (%)	Education (y), mean±s.d.	Race (White, Black, Asian, American Indian, Pacific, multirace), n	CDR, mean±s.d.
FHS					
NC [n = 394]	74.9±10.22^	206, 52.28%	NA	(394, 0, 0, 0, 0, 0)	0.0±0.0
MCI [n = 434]	79.92±8.8^	203, 46.77%	NA	(434, 0, 0, 0, 0, 0)	0.49±0.07
AD [n = 687]	82.99±7.87^	211, 30.71%	NA	(687, 0, 0, 0, 0, 0)	2.04±0.88
LBD [n = 73]	79.34±9.37^	46, 63.01%	NA	(73, 0, 0, 0, 0, 0)	1.84±0.84
VD [n = 113]	81.74±7.3^	48, 42.48%	NA	(113, 0, 0, 0, 0, 0)	1.85±0.8
FTD [n = 8]	85.67±5.91^	4, 50.0%	NA	(8, 0, 0, 0, 0, 0)	2.0±0.87
P value	1.316×10^{-31}	7.905×10^{-14}	NA	1.0	$<1.0 \times 10^{-200}$

High level of missing features, for e.g.: MRI scans unavailable for most of NACC subjects, **ADNI**: 69% missing data, **FHS**: 94% fewer features

Dataset (group)	T1	T2	FLAIR	SWI
NACC	1970	352	318	32
NIFD	633	414	537	3
PPMI	241	N.A.	N.A.	N.A.
AIBL	681	N.A.	334	N.A.
OASIS	662	N.A.	N.A.	N.A.
LBDSU	181	N.A.	N.A.	N.A.
4RTNI	165	119	120	N.A.
ADNI	1055	N.A.	N.A.	N.A.
FHS	115	109	114	N.A.

Overview of Dementia Types in the Study

Dementia Category	No.
Normal Cognition (NC)	19,849
Mild Cognitive Impairment (MCI)	9,357
Alzheimer's Disease (AD)	17,346
Lewy Body Dementia (LBD)	2,003
Vascular Dementia (VD) including Stroke	2,032
Prion Disease (PRD)	114
Frontotemporal Lobar Degeneration (FTD)	3,076
Normal Pressure Hydrocephalus (NPH)	138
Systemic and External Factors (SEF)	808
Psychiatric Diseases (PSY)	2,700
Traumatic Brain Injury (TBI)	265
Other Causes (ODE)	1,234

Table: Categories of dementia and the number of participants in each category.

Data Splitting and Preprocessing

Training dataset

$n = 38,319$

NACC(36,454), AIBL, PPMI, NIFD, LBDSU, OASIS, 4RTNI

Testing dataset

$n = 12,950$

NACC*(8,895), ADNI, FHS

- Non-imaging (**$n = 391$**) preprocessing: Data standardization across cohorts (Uniform Data Set (UDS) 3.0 dictionary).
- Imaging preprocessing:
Skull strip → Reorient → flirt → Intensity normalisation
- Handling missing data: Random feature masking, masking of missing labels to enhance model robustness.

Methods: Embedding Data for Model Input

- **Numerical Data:**

- Example: Age
- Embedding: Linear projection to preserve order.

- **Categorical Data:**

- Example: Gender ("Male", "Female")
- Embedding: Lookup table.

- **Imaging Data:**

- Example: MRI Scans
- Embedding: Swin UNETR, downsampled for transformer input.

Training and Loss Function Overview

- **Combined Loss Function:**
 - Focal Loss (FL) for class imbalance
 - Ranking Loss (RL) for inter-class relationships
 - Includes L2 regularization
- **Optimizer:** AdamW, learning rate = 0.001
- **Scheduler:** Cosine learning rate scheduler with warm restarts
- **Training:** 256 epochs, model selected based on validation performance
- **Handling Class Imbalance:** Focal loss improves performance on imbalanced data based on model without focal loss across 13 diagnostic categories

Results and Validation

In this section, we will explore the following results:

NC, MCI, Dementia Classification

Prodromal AD Detection

Alignment with CDR Scores

Single and Co-occurring Dementias Detection

Validation with Biomarkers/ Neuropathological Evidence

Clinical Assessments

NC, MCI and Dementia Classification

Test Cases: Evaluated on NC, MCI, and dementia (NACC, ADNI, FHS datasets)

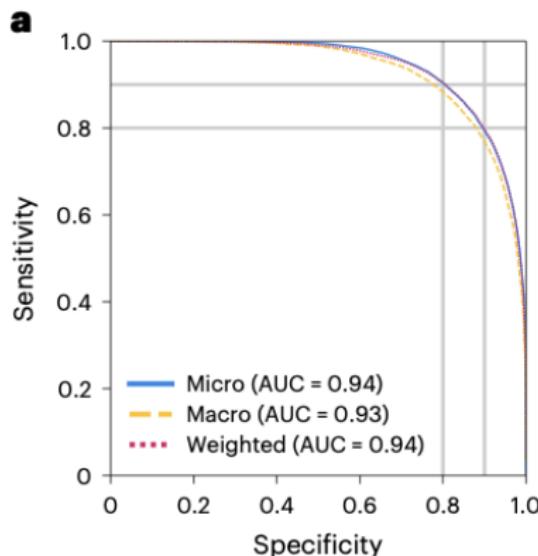


Figure: ROC

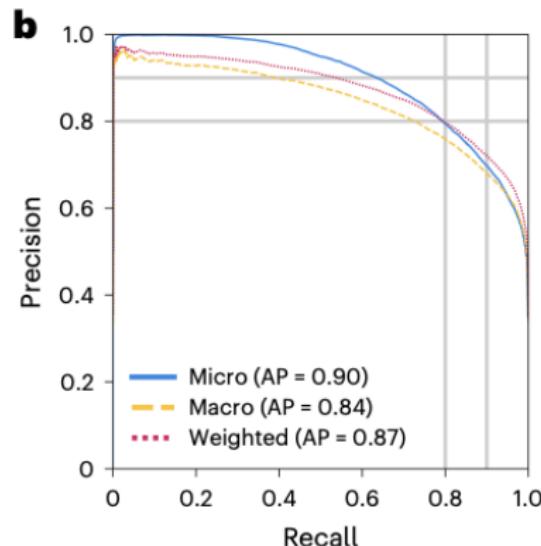
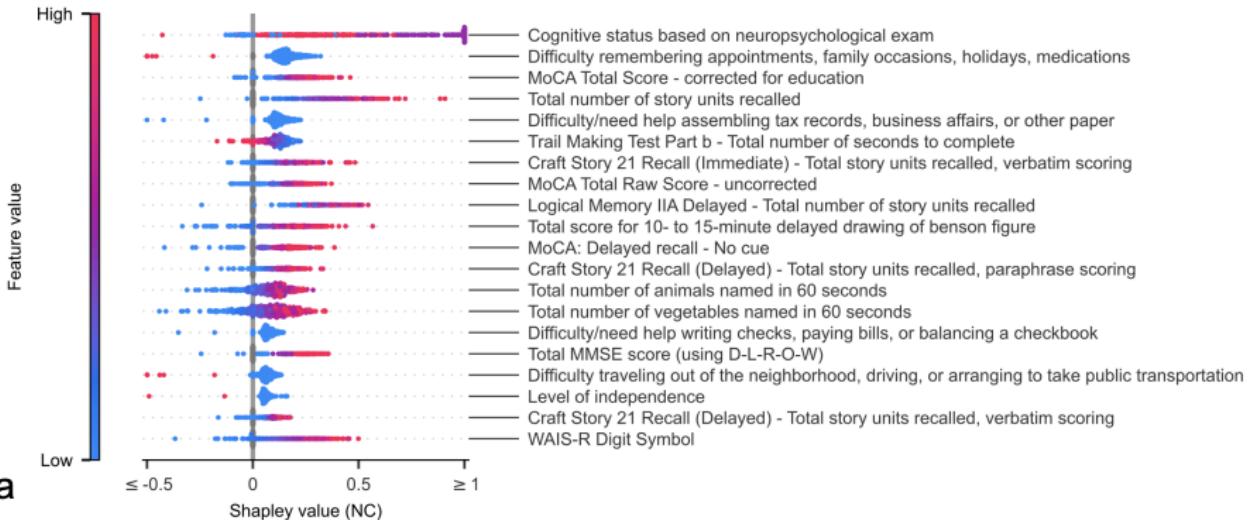


Figure: PR

NC, MCI and Dementia Classification (contd.)

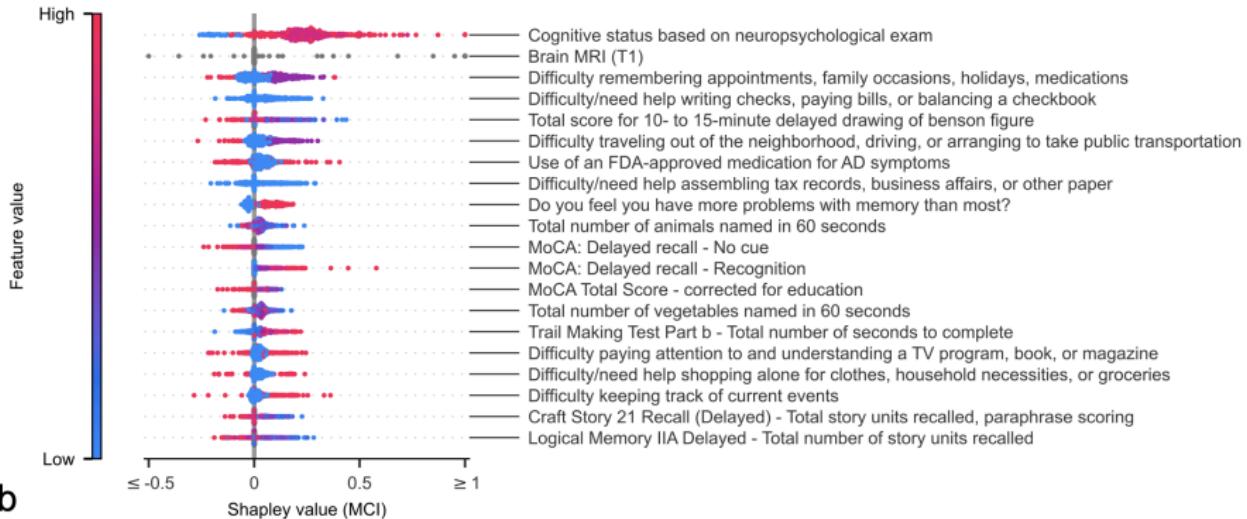
- **Consistent Across Subgroups:** Age, gender, race
 - Microaveraged AUC ≥ 0.88 , AUPR ≥ 0.82
- **Benchmarking:**
 - Compared against CatBoost
 - Similar performance on NACC dataset
 - Outperformed CatBoost on ADNI and FHS
 - AUROC improvements: 0.02 to 0.21, AUPR improvements: 0.03 to 0.17

Interpretability - NC (Normal Cognition)



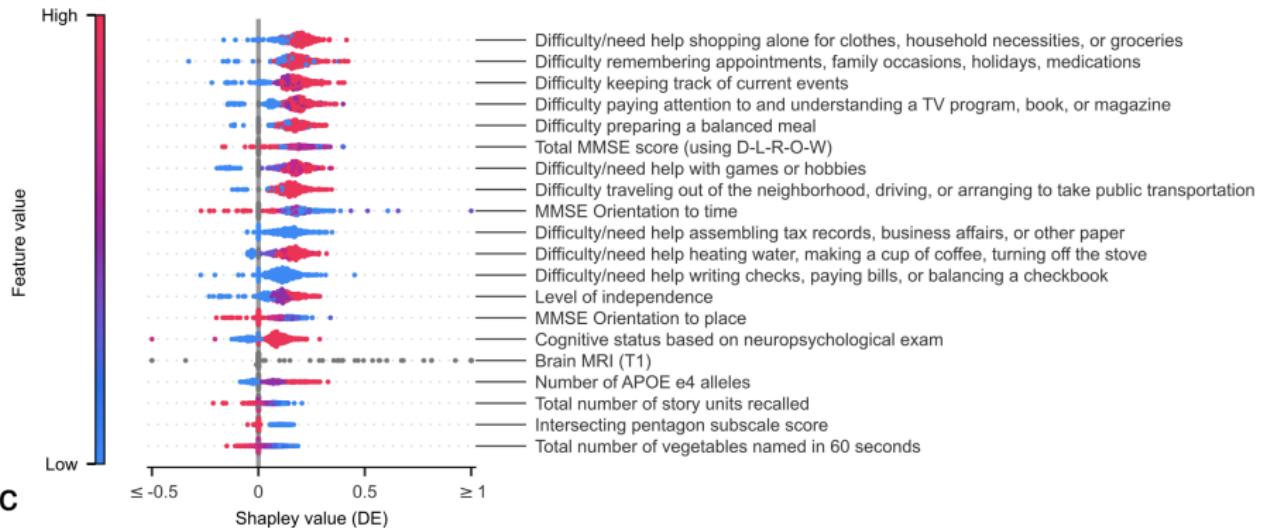
Shapley Key Features: Cognitive Status, MoCA Scores, Memory Tasks

Interpretability - MCI (Mild Cognitive Impairment)



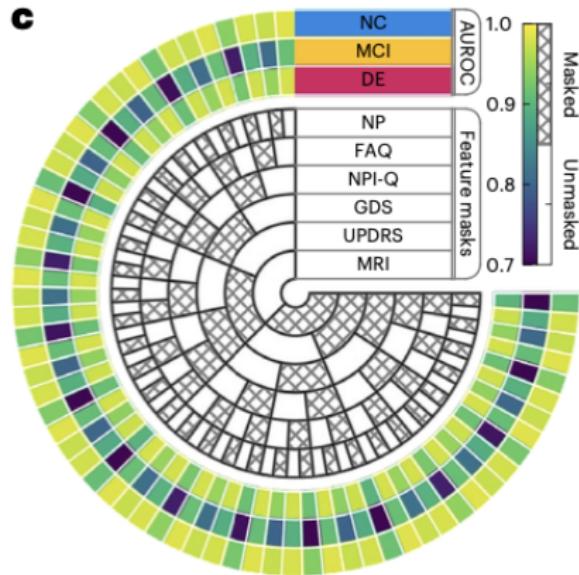
Shapley Key Features: Memory Recall, Functional Decline, MRI Changes

Interpretability - DE (Dementia)



Shapley Key Features: Functional Decline, MMSE Scores, APOE4 Alleles

Incomplete Data



Testing on NACC:

Missing data simulated yet consistently reliable scores despite missing features (e.g., MRI, FAQ, NP tests).

External Testing:

ADNI: 69% missing data, AUROC 0.91, AUPR 0.86

FHS: 94% fewer features, AUROC 0.68, AUPR 0.53

Prodromal AD Detection

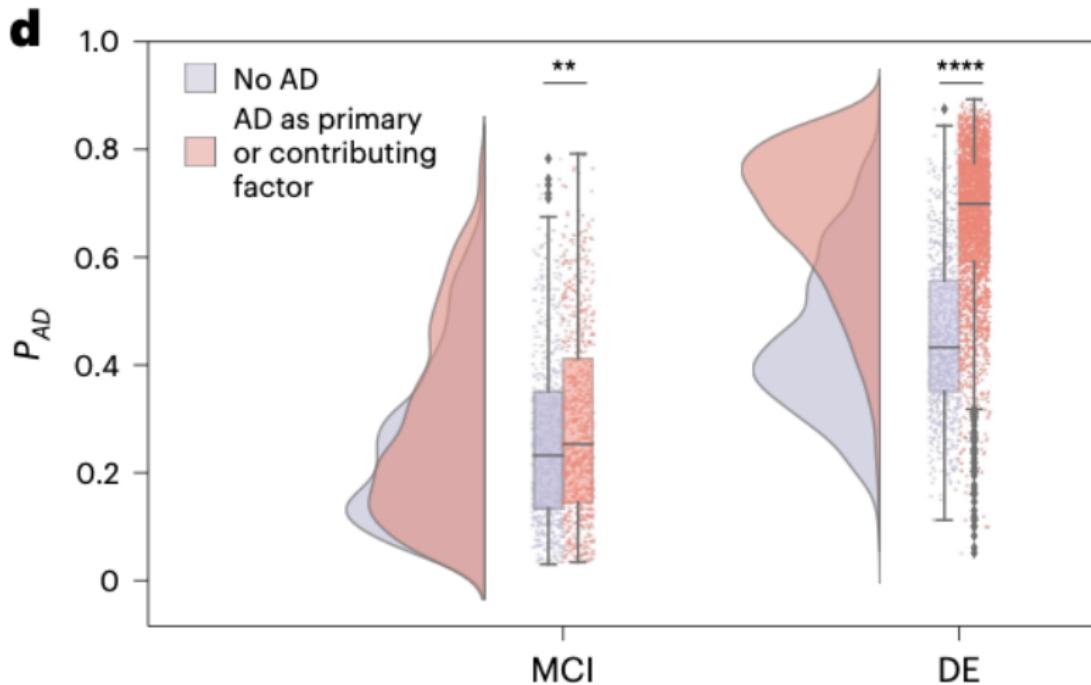
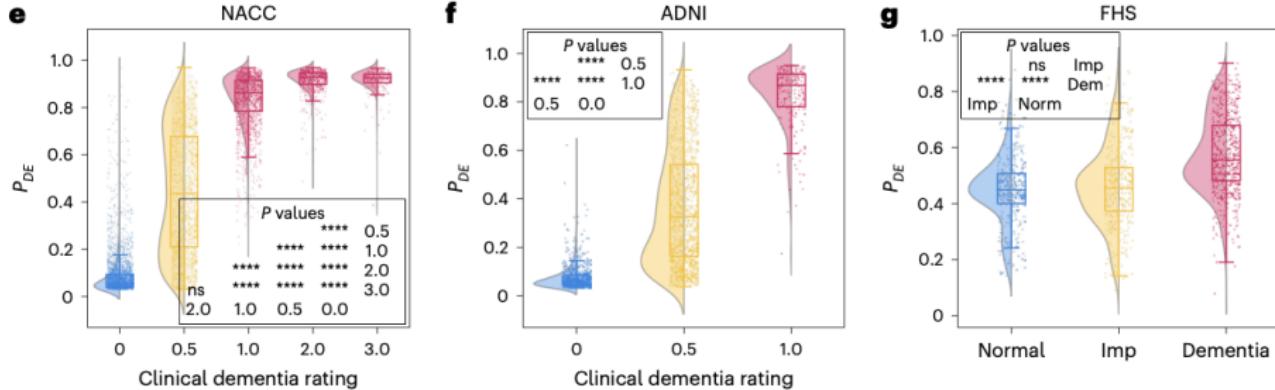


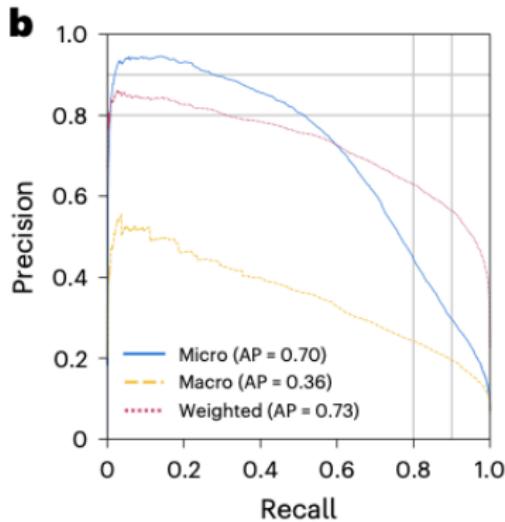
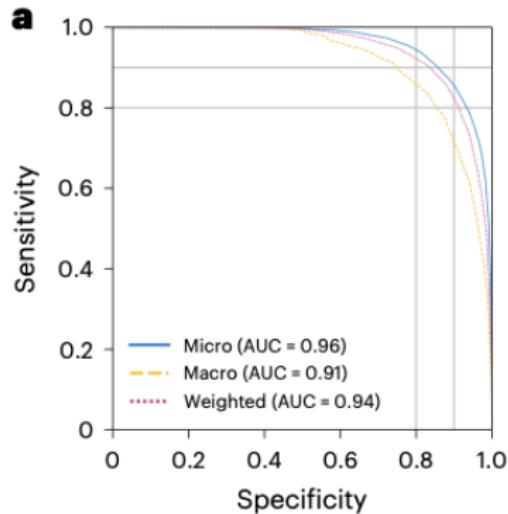
Figure: Probability of AD, $P(\text{AD})$, for MCI and DE. Model generally assigns higher $P(\text{AD})$ for MCI with AD etiology. Helpful for early diagnosis?

Alignment of P(DE) with CDR Scores



- Strong correlation between model's $P(\text{DE})$ and CDR scores.
- $P(\text{DE})$ increases with higher CDR scores (NACC and ADNI datasets).
- Challenge in distinguishing early cognitive decline (FHS dataset). Ascribed to lack of CDR in FHS.

Single Dementia Detection



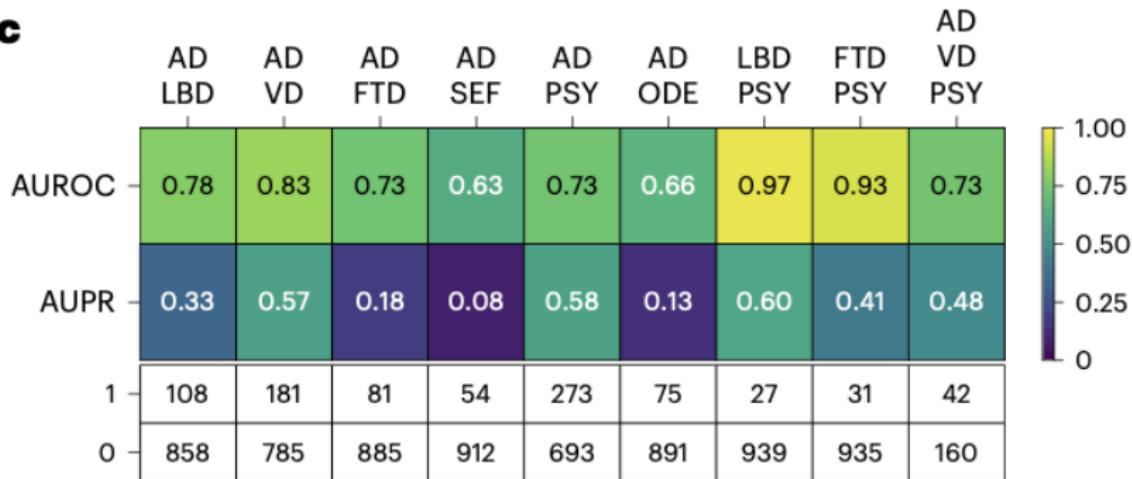
- **Macroaverage AUPR:** Lower scores suggest better performance on some diagnoses than others.
- **Weighted AUPR:** Weighted-average higher because poorer performance is on rare dementias.
- **Stable Across Demographics:** Consistent performance across age, gender, and race, with microaveraged AUC > 0.94 and AP > 0.66.

Single Dementia Detection (contd.)

Dataset (group)	Balanced Accuracy	Precision	Sensitivity	Specificity	F1 Score	MCC	AUROC	AUPR
NACC								
NC	0.93	0.9	0.92	0.94	0.91	0.85	0.98	0.97
MCI	0.83	0.52	0.80	0.85	0.63	0.56	0.91	0.67
DE	0.94	0.92	0.94	0.93	0.93	0.87	0.99	0.98
AD	0.89	0.84	0.87	0.91	0.86	0.78	0.96	0.93
LBD	0.87	0.43	0.80	0.95	0.56	0.56	0.96	0.68
VD	0.83	0.28	0.74	0.92	0.41	0.42	0.93	0.47
PRD	0.67	0.09	0.35	0.99	0.14	0.17	0.96	0.12
FTD	0.89	0.36	0.90	0.89	0.51	0.53	0.96	0.67
NPH	0.55	0.12	0.11	1.00	0.12	0.12	0.91	0.077
SEF	0.66	0.069	0.42	0.90	0.12	0.13	0.82	0.064
PSY	0.79	0.24	0.71	0.86	0.36	0.36	0.90	0.36
TBI	0.62	0.07	0.26	0.98	0.11	0.12	0.90	0.098
ODE	0.68	0.11	0.46	0.89	0.17	0.18	0.84	0.11
ADNI								
NC	0.83	0.64	0.97	0.69	0.77	0.64	0.94	0.89
MCI	0.76	0.82	0.63	0.88	0.71	0.53	0.87	0.83
DE	0.90	0.64	0.91	0.89	0.75	0.71	0.97	0.88
AD	0.91	0.70	0.89	0.92	0.78	0.74	0.97	0.86
FHS								
NC	0.59	0.35	0.42	0.76	0.38	0.17	0.66	0.33
MCI	0.53	0.40	0.13	0.93	0.20	0.098	0.59	0.34
DE	0.68	0.68	0.68	0.68	0.68	0.36	0.73	0.71
AD	0.65	0.63	0.52	0.78	0.57	0.32	0.72	0.64
LBD	0.52	0.077	0.068	0.96	0.072	0.032	0.62	0.071
VD	0.65	0.18	0.44	0.85	0.26	0.20	0.74	0.30
FTD	0.59	0.016	0.25	0.92	0.03	0.045	0.71	0.028

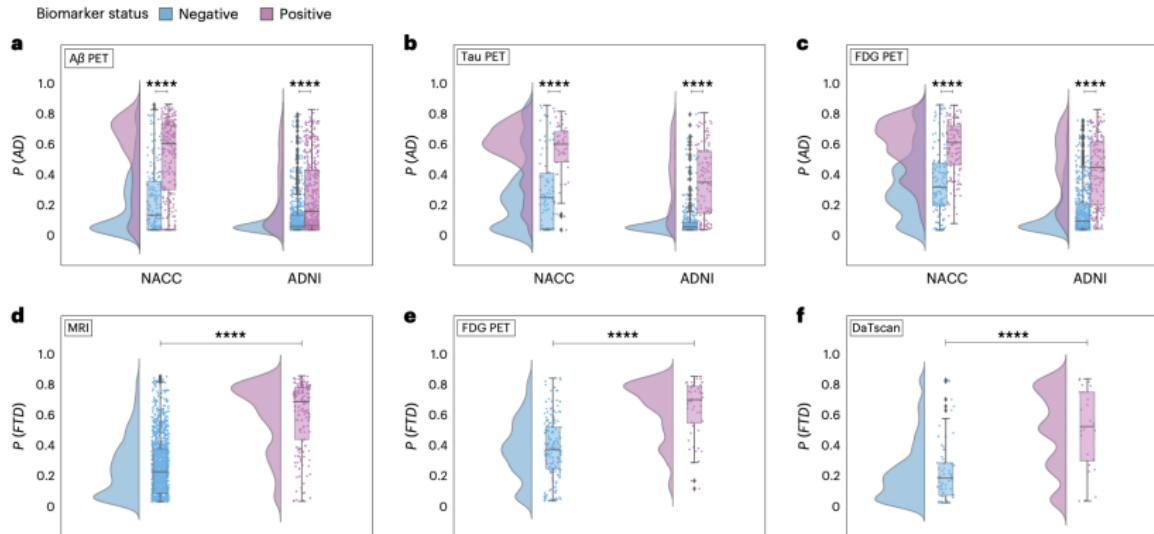
Co-occurring Dementias Detection

C



- AUROC for co-occurrences with more than 25 samples.
- Number of positive and negative samples available.
- Only NACC used.

Validation with Biomarkers



- **AD:** Higher $P(\text{AD})$ with $A\beta$, tau, FDG PET ($P < 0.0001$).
- **FTD:** Higher $P(\text{FTD})$ with MRI, FDG PET.
- **LBD:** Higher $P(\text{LBD})$ with DaTscan. [Figure f (typo)]
- Model aligns with biomarker profiles.
NACC and ADNI for AD and NACC for FTD, LBD

Validation with Neuropathological Evidence

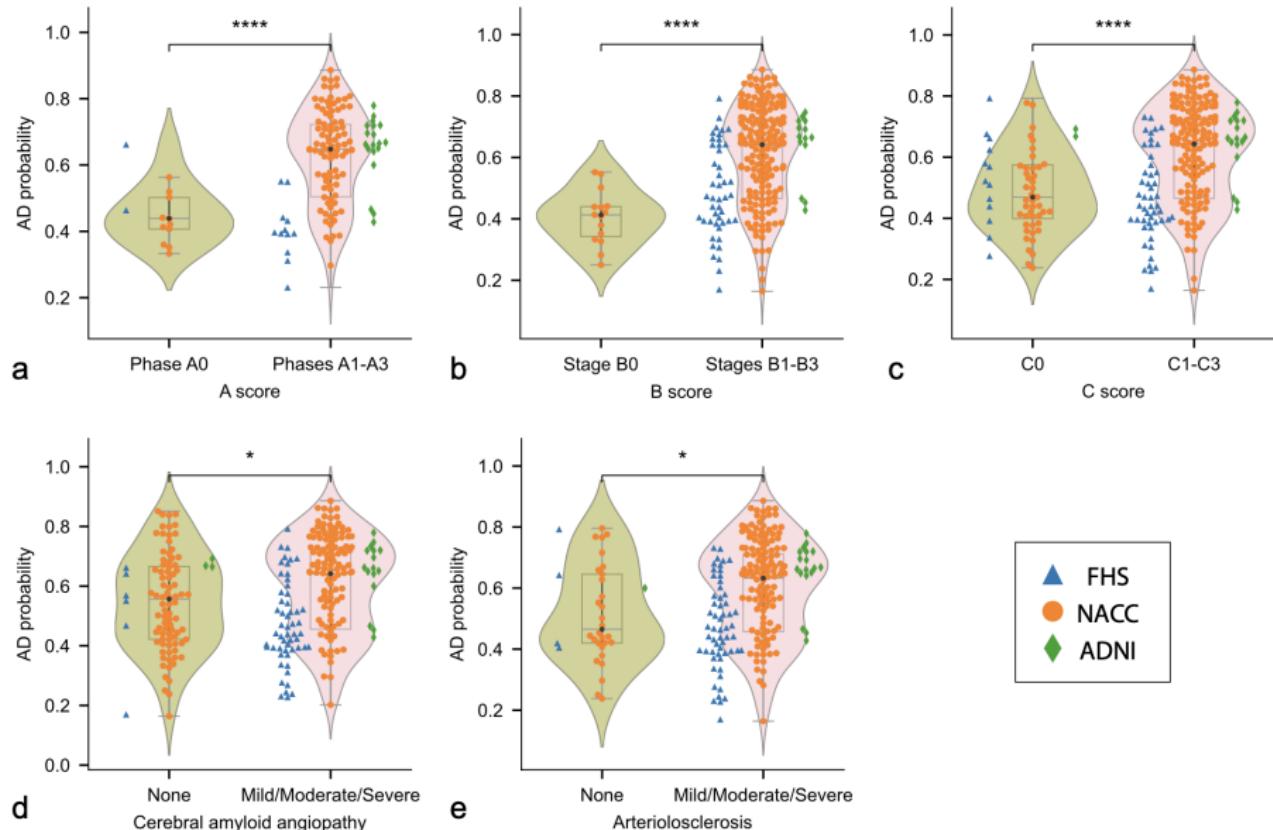
Validation: Model's etiology-specific scores validated against postmortem neuropathological markers (smaller sample size).

Cohort	Age at death (years) mean ± std	Male gender (percentage)	CDR mean ± std
NACC			
AD [n = 131]	79.28 ± 10.61	74, 56.4%	1.44 ± 0.75
LBD [n = 44]	76.64 ± 8.31	36, 81.82%	1.44 ± 0.66
VD [n = 15]	83.73 ± 10.39	7, 46.67%	1.37 ± 0.86
PRD [n = 5]	562.6 ± 4.96	5, 100%	0.7 ± 0.24
FTD [n = 20]	67.15 ± 8.61	13, 65%	1.82 ± 0.98
NPH [n = 1]	77 ± 0.0	1, 100%	0.5 ± 0.0
SEF [n = 3]	73 ± 24.04	2, 66.67%	0.83 ± 0.23
TBI [n = 1]	87 ± 0.0	1, 100%	2.0 ± 0.0
ODE [n = 11]	68.82 ± 14.99	7, 63.64%	1.45 ± 0.86
p-value	1.428e-08	5.956e-02	1.471e-01
ADNI			
AD [n = 19]	82.05 ± 9.03	13, 68.42%	0.86 ± 0.22
FHS			
AD [n = 55]	91 ± 7.0	18, 32.72%	1.01 ± 0.15
LBD [n = 4]	92.75 ± 2.16	3, 75%	1.0 ± 0.0
VD [n = 5]	93 ± 3.35	4, 80%	1.2 ± 0.4
FTD [n = 2]	79 ± 4.0	1, 50%	1.0 ± 0.0
p-value	1.584e-01	8.227e-02	2.972e-01

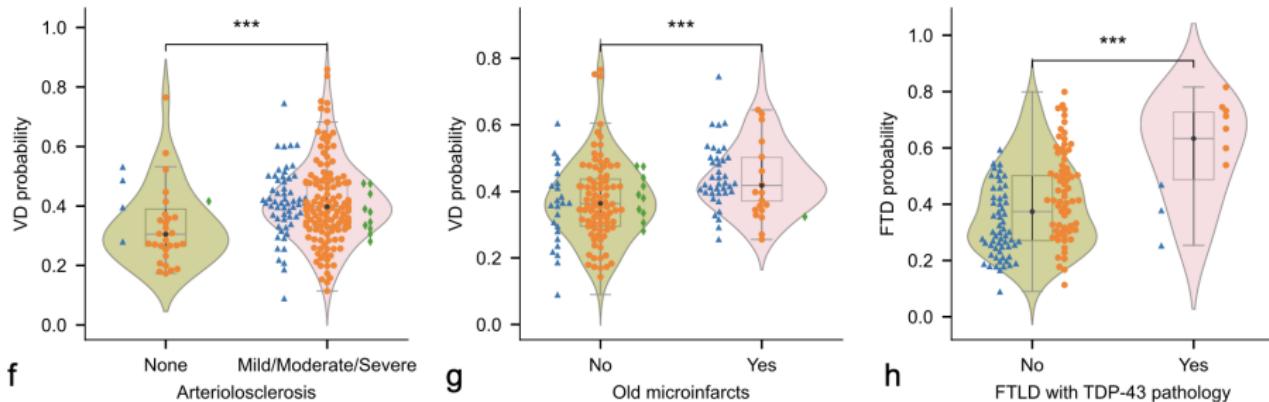
Alzheimer's Disease (AD): Thal, Braak, CERAD

- Higher P(AD) with advanced Thal, Braak, and CERAD stages ($p < 0.0001$).
- Higher P(AD) in cases with CAA and arteriolosclerosis ($P < 0.05$).

Validation with Neuropathological Evidence (contd.)



Validation with Neuropathological Evidence (contd.)



Vascular Dementia (VD): Arteriolosclerosis, microinfarcts

- Significant variation in $P(VD)$ with arteriolosclerosis and microinfarcts ($P < 0.001$).

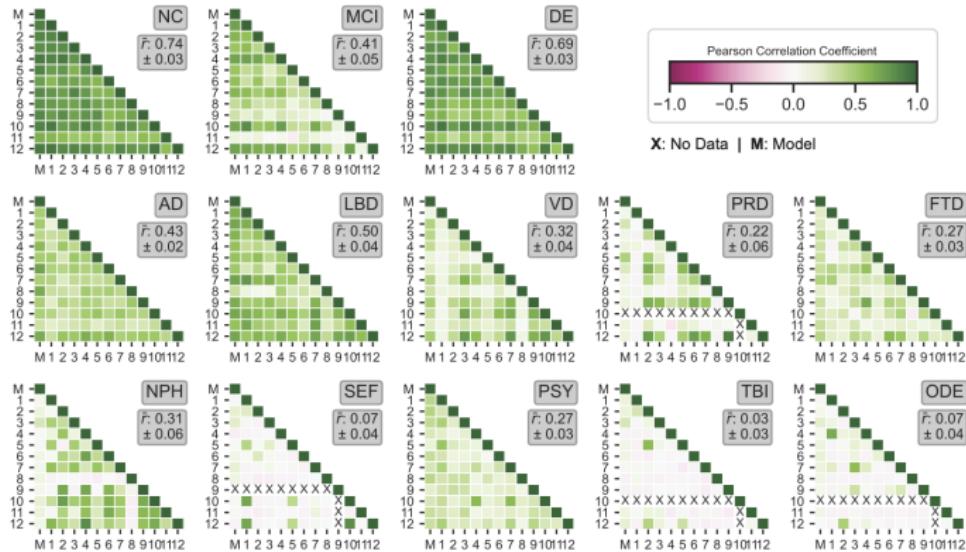
Frontotemporal Dementia (FTD): TDP-43 pathology

- Significant difference in $P(FTD)$ with TDP-43 pathology ($P < 0.001$).

Clinical Experiment:

- 12 neurologists, 7 neuroradiologists
- Neurologists: 100 cases (NC, MCI, various dementia subtypes)
- Radiologists: 70 cases (Dementia subtypes only)
- Data: Demographics, medical history, neuropsych tests, MRI scans
- Clinicians rated confidence (0-100) per diagnosis
- AI model provided predictions on the same cases

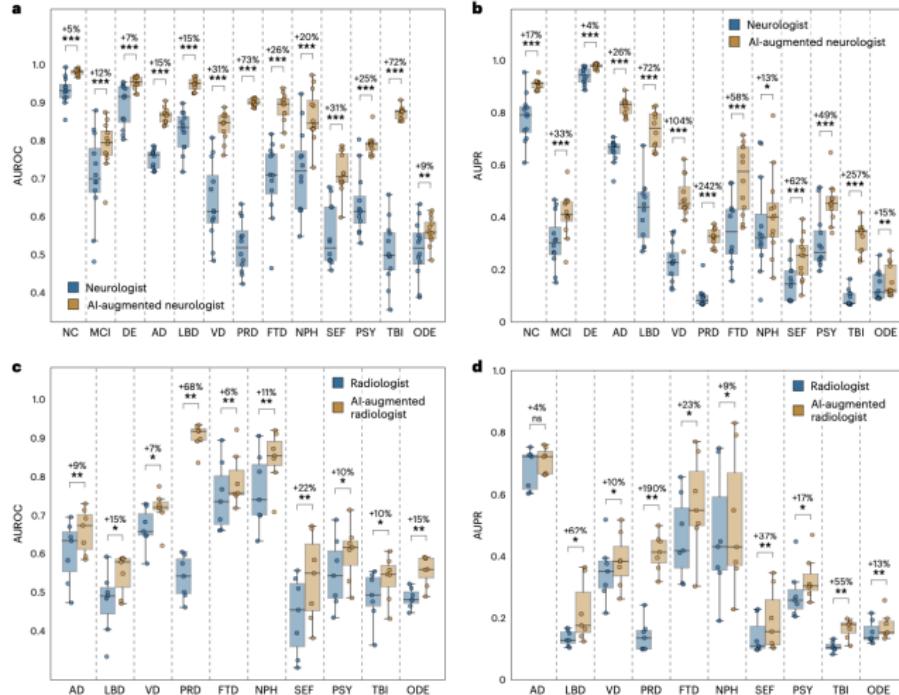
Clinical Assessments: Results



Interrater Agreement:

- Pearson correlation between clinicians' scores and AI predictions.
- NC, DE > MCI, AD, LBD, VD, FTD, PSY > PRD, NPH, SEF, TBI, ODE (More to less agreement)

Clinical Assessments: Results (contd.)



Average, across all etiologies

Neurologists:

- AUROC +26.25%
- AUPR +73.23%

Neuroradiologists:

- AUROC +16.19%
- AUPR +41.79%

Figure: AI-augmented: Average score

Quick Review

Better differential diagnosis of dementia etiologies will help improve misdiagnosis and delayed diagnosis leading to better treatment planning and patient outcomes.

Doctors + AI = Enhanced Diagnoses

Discussion

In this section, we will explore:

- Limitations of Study
- Clinical Implementation
- Potential Takeaways

What are some possible limitations?

Data Bias and Generalisability:

- Predominantly white population, heavy focus on NACC and common dementias (AD)
- Missing imaging data

Data Quality:

- Variability in data quality across datasets could impact model accuracy - PSY (Bad labelling)

Interpretability:

- Difficulty in understanding and explaining the model's decision-making process

Can this model be implemented?

Steps in the right direction: Easier to sell to clinicians.

- Probability outputs enable systematic prioritisation of cognitive impairment drivers.
- Minimal disruption (useful for initial screenings).
- Flexibility of inputs (from GP to specialist).

Obstacles:

- Processing power.
- Trust among clinicians.
- Generalisability and data availability.
- Choosing right threshold for each etiology.

Potential Takeaways for the Lab

- Multimodal integration.
- Masking of missing data.
- Understanding of clinical challenges.

Swin UNETR

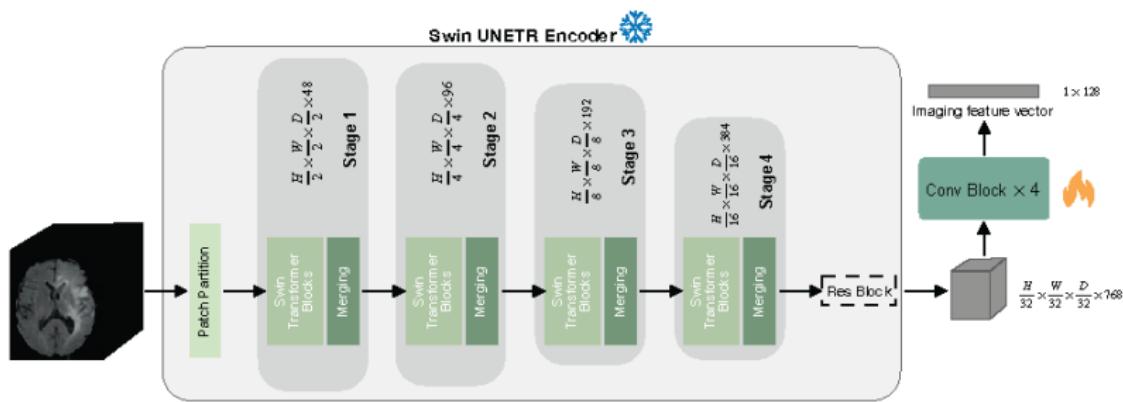


Figure: Swin UNETR: Sliding Window U-Net Transformers