

Classifying Alzheimer's disease, mixed/vascular dementia and other diagnoses in a real-world memory-clinic cohort using routine structural MRI and clinical data

Le Yang²; Daniel Stamate^{4, 5}; Daniel Stahl^{2, 3*}; Ashwin V. Venkataraman^{1, 3*}

¹Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London; ²Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology & Neuroscience, King's College London; ³South London and Maudsley NHS Foundation Trust
⁴School of Health Sciences, University of Manchester; ⁵Data Science and Soft Computing Lab, School of Computing, Goldsmiths, University of London | *Shared senior authorship

Routine MRI and clinical/EHR data supported moderate classification of Alzheimer's disease, mixed/vascular dementia and other diagnoses; WMH added a small, clinically plausible vascular signal.

INTRODUCTION

Clinical problem

- Routine memory clinics must distinguish Alzheimer's disease, mixed/vascular dementia and other non-AD diagnoses.
- Accurate stratification matters for amyloid-therapy pathways and vascular risk management.
- Structural MRI and EHR data remain central when PET or CSF biomarkers are unavailable.

Why routine-care data matter

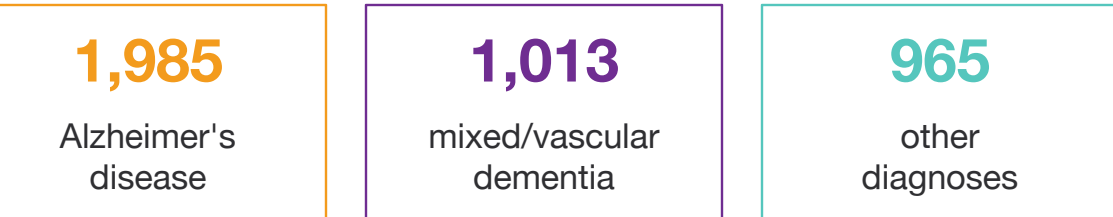
- The cohort reflects symptomatic patients seen in real clinical pathways.
- The data include heterogeneous diagnoses, missingness and variable reporting practice.
- This setting tests whether routine information can support scalable diagnostic stratification.

Study aim

- Evaluate standard memory-service data for three-class diagnostic prediction.
- Test whether WMH adds value beyond MRI volumes, cognition, comorbidity and report NLP.
- Interpret feature patterns for vascular, neurodegenerative and cognitive signals.

Cohort and labels

- Baseline-unique analysis cohort: 3,963 participants.
- Labels mapped to AD, Mixed and Non-AD groups.



METHODS

Study design and data sources

- Retrospective prediction-modelling study using linked EHR and structural MRI data.
- MRI came from the SLaM Image Bank; clinical data came from CRIS.
- The linked scan-level dataset contained 24,418 structural-MRI records.

Outcome definition

- The first eligible structural MRI per participant was retained.
- Structured diagnosis labels were mapped using prespecified whitelists.
- Final outcome groups were AD, Mixed and Non-AD.

Predictor domains

- MRI volumetrics covered temporal, cingulate, precuneus, parietal-occipital and ventricular regions.
- Clinical/EHR predictors included age, sex, MMSE and vascular comorbidities.
- Report NLP and WMH features captured ischaemic, neurodegenerative and vascular-burden signals.

Model development and reporting

- Compared direct three-class and hierarchical two-stage XGBoost models, with and without WMH.
- Missing data were handled inside XGBoost; missingness indicators were retained.
- Participant-grouped nested CV used 5 outer folds and 10 inner folds; model selection optimised macro-F1.
- Reporting focused on macro-AUC, macro-F1, balanced accuracy, overall accuracy and feature interpretation.

RESULTS

Direct + WMH had the highest macro-AUC, macro-F1 and balanced accuracy. Hierarchical + WMH had the highest overall accuracy and Alzheimer's disease recall. Higher WMH burden increased the SHAP-style contribution to Mixed probability.

Figure S3. WMH burden SHAP-style dependence

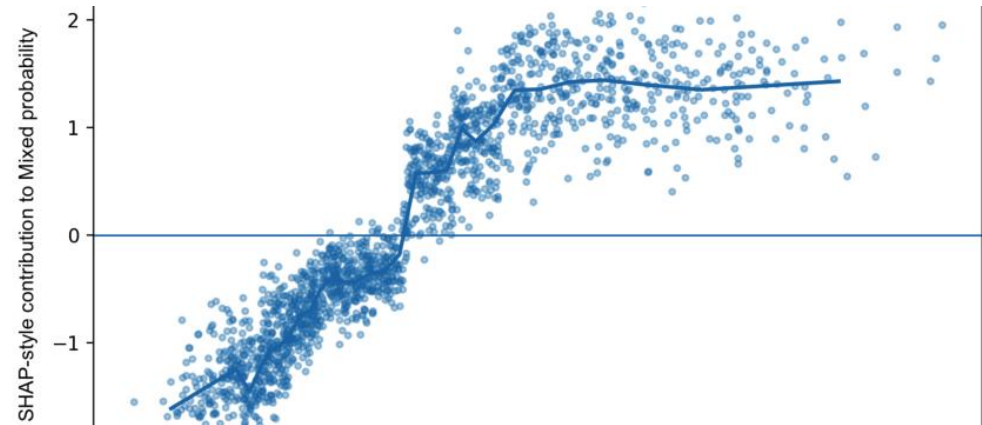


Table 2. Primary nested cross-validation performance in the baseline participant-level subset (n = 3963)

Model	Macro-AUC	Macro-F1	Balanced accuracy	Overall accuracy
Direct three-class model without WMH	0.752 [0.740, 0.765]	0.554 [0.538, 0.569]	0.565 [0.550, 0.582]	0.568 [0.553, 0.583]
Direct three-class model with WMH	0.760 [0.749, 0.773]	0.569 [0.553, 0.584]	0.582 [0.567, 0.599]	0.583 [0.567, 0.597]
Hierarchical model without WMH	0.749 [0.738, 0.762]	0.553 [0.536, 0.568]	0.552 [0.537, 0.569]	0.583 [0.568, 0.598]
Hierarchical model with WMH	0.755 [0.744, 0.768]	0.563 [0.547, 0.579]	0.559 [0.544, 0.576]	0.598 [0.583, 0.613]

Figure 2. Row-normalised confusion matrices

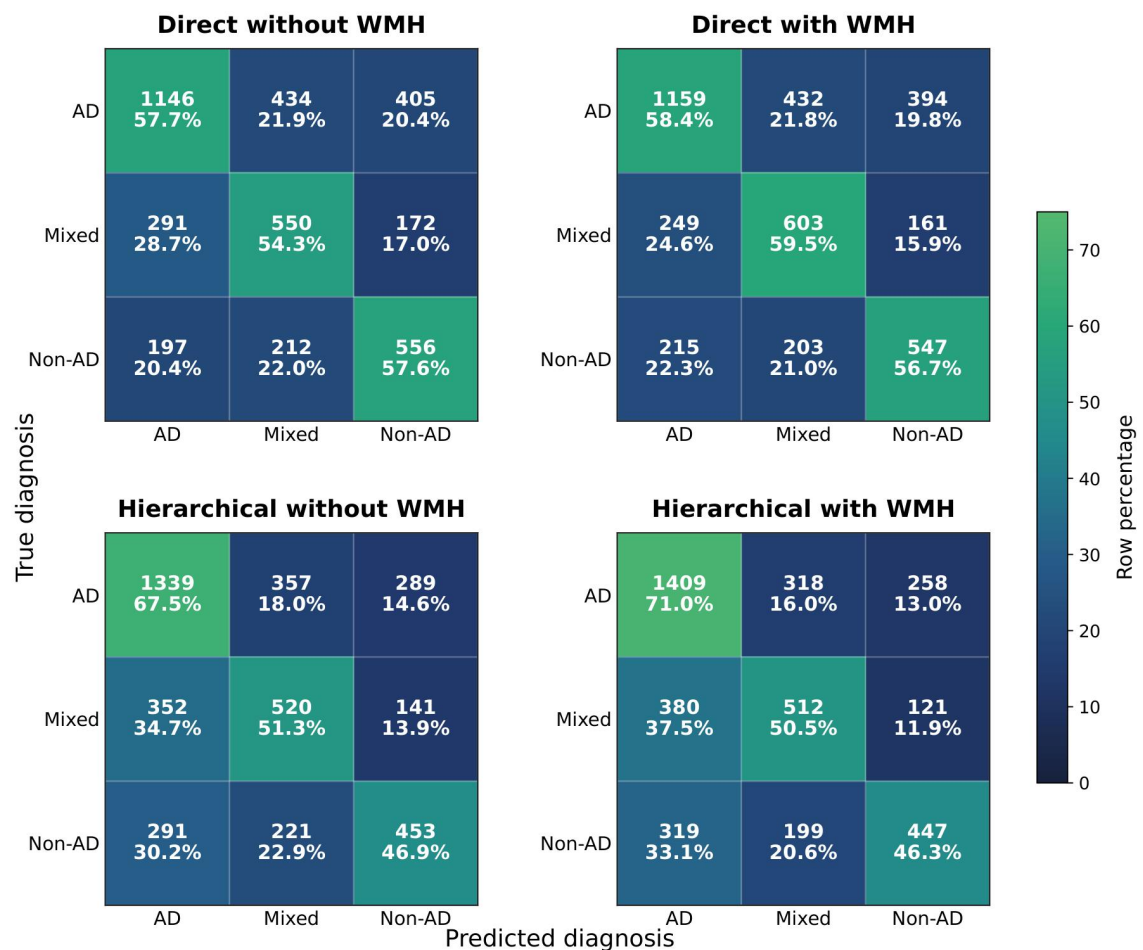
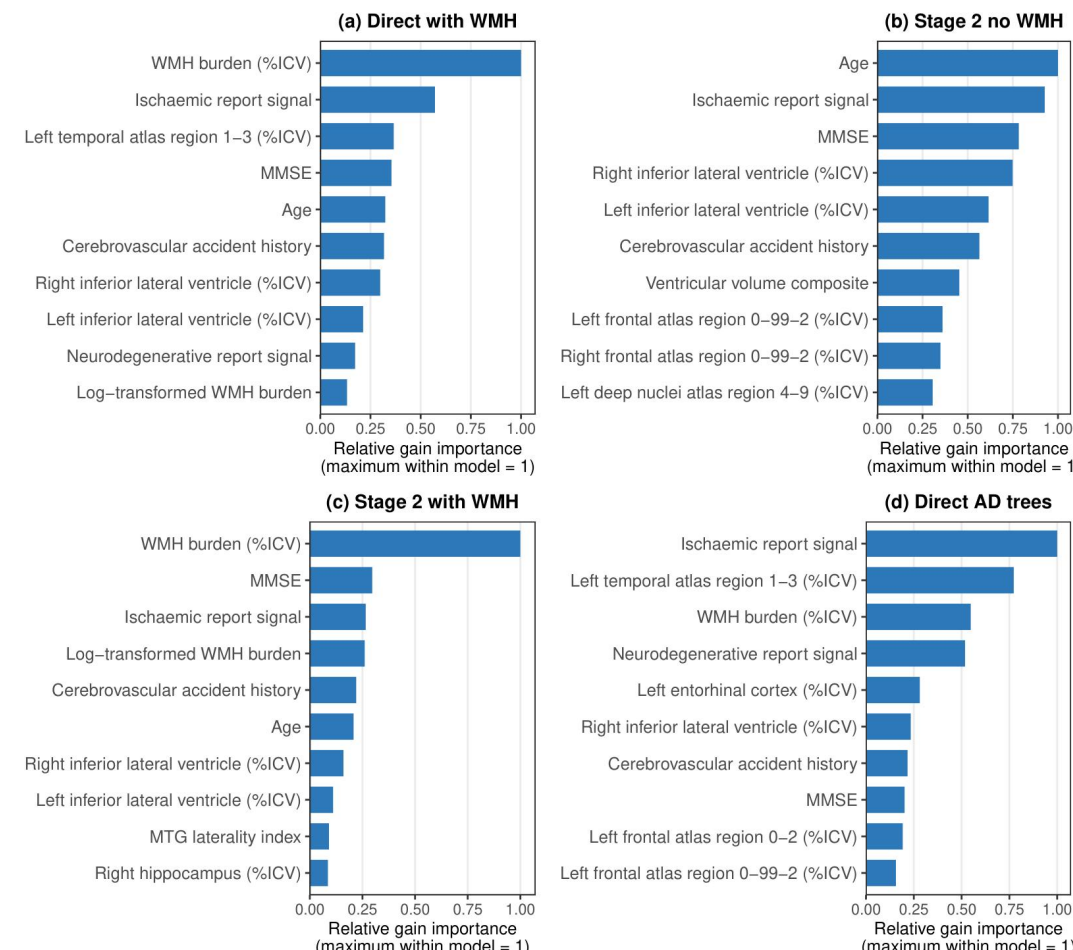


Figure 3. Gain-based XGBoost feature importance



CONCLUSIONS

Main conclusions

- Direct + WMH gave the most balanced classification across Alzheimer's disease, mixed/vascular dementia and other diagnoses.
- Hierarchical + WMH achieved the highest overall accuracy and Alzheimer's disease recall.
- WMH burden, ischaemic report signal and stroke history provided a clinically plausible vascular signal.

Limitations

- Single-site routine-care cohort.
- No external validation or biomarker-confirmed outcome labels.
- Calibration and testing across other NHS settings are still needed.

REFERENCES

1. van Dyck et al. N Engl J Med. 2023;388:9-21.
2. Livingston et al. Lancet. 2020;396:413-446.
3. DeBette & Markus. BMJ. 2010;341:c3666.
4. Wardlaw et al. Lancet Neurol. 2013;12:822-838.
5. Collins et al. BMJ. 2024;385:e078378.
6. Chen & Guestrin. KDD. 2016:785-794.

CONTACT

Le Yang
le.1.yang@kcl.ac.uk
King's College London