

Vision-language AI model for dementia staging and 12-phenotype classification using routine assessments from real-world memory clinics

aaic2026.3f7070c

Christensen M (1, 2), Shen Y (3), Jiang Y (3), Dennison J (4), Darby D (1, 2), Ge Z (3), Brodtmann A (1, 2)

1 Faculty of Medicine, Nursing and Health Sciences, School of Translational Medicine, Monash University; 2 Eastern Cognitive Disorders Clinic, Department of Neurosciences, Box Hill Hospital, Eastern Health Clinical School; 3 Faculty of IT, Augmented Intelligence and Multimodal Analytics for Health Lab, Monash University; 4 Lived Experience Advisor, Melbourne, Australia

PLAIN LANGUAGE SUMMARY

We have built an artificial intelligence (AI) model to support how dementia is diagnosed in clinic. Unlike most AI tools, which rely on research data, brain scans and other specialised tests, ours is designed for everyday clinical practice.

- It uses information that is already collected in routine care, such as memory tests and questionnaires completed by family.
- Because this information is available early, it is designed to support the doctor from the very first appointment.
- It moves from answering only whether a person has dementia towards also answering what type they have, if they do.
- It explains the reasons behind each result, so the doctor can weigh them up for that person.
- People with lived experience of dementia helped design it, so it reflects what matters to patients and families.

In our study, the model was good at judging whether a person had normal cognition or dementia and showed promise at predicting specific types of dementia, though this is an early proof of concept and the rarer types remain a work in progress.

INTRODUCTION

The emergence of disease-modifying therapies has made timely dementia diagnosis more critical than ever, yet diagnostic delay remains measured in years. Whilst AI-based diagnostic tools show promise in research settings, clinical translation has been limited by reliance on homogeneous research datasets and input features that do not reflect real-world practice, often requiring advanced neuroimaging or expensive and invasive biomarkers unavailable at initial assessment.

RESULTS

Staging model (Figure 2): macro-AUC 0.90, macro-AUPR 0.79. Per-class precision was highest for dementia (0.83) and HC (0.75); lower MCI precision (0.58) reflected the inherent diagnostic uncertainty at this clinical boundary.

Phenotype model (Figure 3): macro-AUC 0.71, macro-AUPR 0.31, top-3 precision 87% (correct diagnosis within the top 3 predictions, chance level 25%). Performance was strongest for prevalent phenotypes (amnesic F1 0.80); rare phenotypes were limited by underrepresentation.

CONCLUSIONS

Our model addressed key translational gaps by leveraging heterogeneous real-world clinic data and offering broad phenotype coverage. The novel multimodal architecture successfully integrated structured clinical data with vision-language encoding of handwritten patient responses. Strong staging and promising phenotype discrimination, with high top-3 accuracy, demonstrated the feasibility of explainable AI-augmented diagnosis using only widely available clinical inputs. Limitations included class imbalance affecting phenotype precision-recall. External validation and prospective studies are needed to establish generalisability and impact on patient outcomes.

METHODS

Figure 1

Patient Cohort

1066 Patients, 57% male, mean onset age 63.3 years.

Private Clinic n = 579

Public Clinic n = 487

Data collection retrospective (2012-2024), Ethics approved

Input Features

Demographics and modifiable risk factors for dementia

ACE-R: all items, including verbal fluency answers, drawings and patient writing

MMSE: subset of the ACE-R

Carer Questionnaires: CBI-R and 3 local questionnaires (function, behaviour, memory)

Prediction Model

VLM Encoder (Qwen3-VL-Embedding-8B)

MLP Decoder

Train/Val/Test: 80/10/10

Explainability: Sequential (leave-one-out) ablation of clinically grouped features

Prediction Outcomes

Model 1: Staging

Healthy Control, Mild Cognitive Impairment, Dementia

Model 2: Phenotype

Amnesic, VCI, PCA, Behavioural, lvPPA, nvPPA, svPPA, rtvFTD, CBS, PSP, DLB, PDD

Implementation Vision

Multimodal Data Collection via Diagnostic Portal

Patient: Clinical Data
Carer: Questionnaires
Practice Nurse: ACE-R

AI predictions prior to Review with Doctor

Preliminary Stage and Phenotype Prediction with Explainability

Targeted AI-Augmented Assessment by Doctor

Timely diagnosis and access to treatment and support services

MONASH University

Eastern Health

CONTACT

maja.christensen@monash.edu

ALZHEIMER'S ASSOCIATION

AAIC>26