

A CLINICAL EVALUATION OF A DECISION SUPPORT TOOL FOR DEMENTIA DIAGNOSIS USING HMPAO SPECT IMAGING

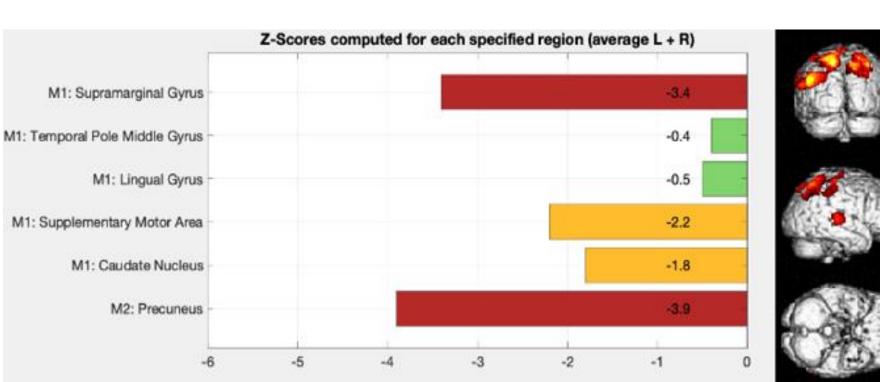


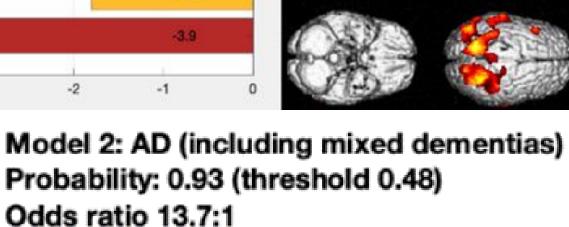
University Hospital Southampton NHS Foundation Trust

Dr Dilan Sanli, Dr Sofia MichopoulouUniversity Hospital Southampton NHS Foundation Trust

PROJECT

Over 950,000 people currently live with dementia in the UK, rising to 1.4 million in 2040¹. Early, accurate diagnosis remains a critical challenge. This project evaluates a novel, NIHR-funded, in-house machine learning tool developed to support the interpretation of brain HMPAO SPECT scans for dementia diagnosis. The tool comprises two logistic regression-based models: Model 1 classifies scans as normal versus abnormal, while Model 2 assesses the likelihood of Alzheimer's Disease (AD)/ Lewy Body Dementia (LBD). Both models generate a probability score and odds ratio to assist clinical decision-making.

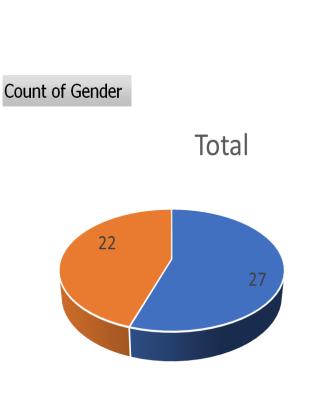




Model 1: Abnormal Scan Probability: 0.99 (threshold 0.85) Odds ratio 107.0:1

METHODS

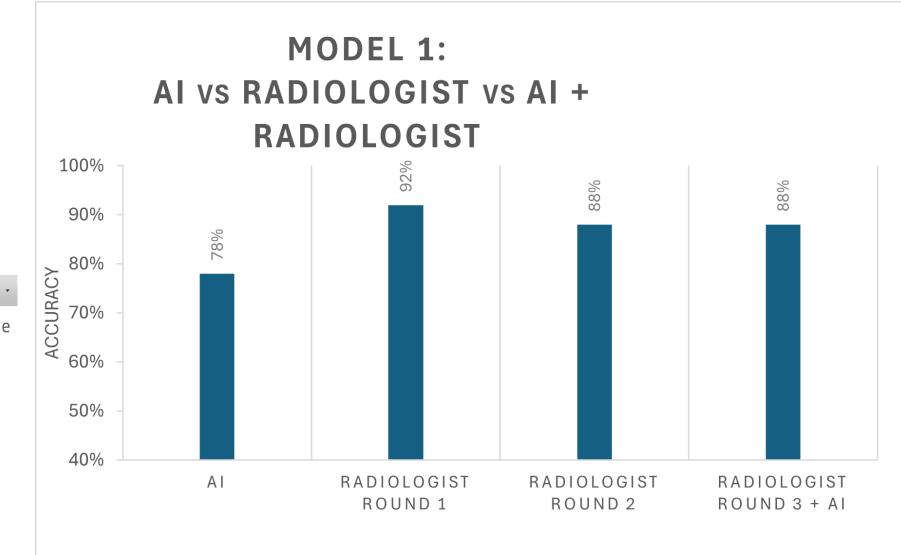
The evaluation was conducted by a Clinical AI Fellow, also a radiology trainee, who received structured training in HMPAO scan interpretation under the guidance of a Professor of Neurology. The AI outputs were embedded on the perfusion to diagnosis reports, providing a holistic view of the scan. The Radiologist independently assessed 49 patient scans acquired in September 2024, across three phases: an initial review (Round 1), a reassessment following further training (Round 2), and a final assessment incorporating the AI model's output (Round 3). Each scan was rated on two dimensions—normal vs abnormal and AD/LBD present vs absent—alongside a five-point diagnostic confidence score (ranging from very low [Level 1] to very high [Level 5])

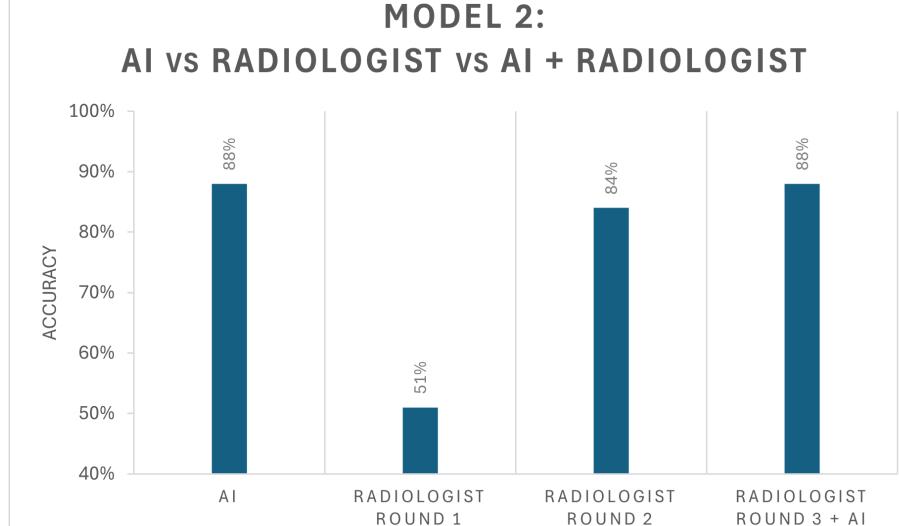


Gender

Female

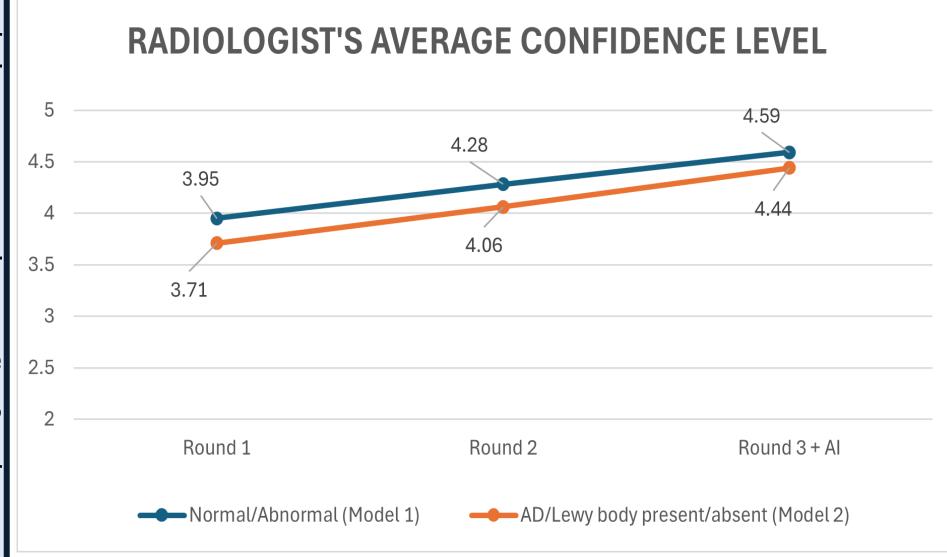
Male





RESULTS

Model 1 demonstrated 78% concordance with expert opinion, while Model 2 achieved 88%. The Fellow's accuracy in classifying scans as normal or abnormal was 92% in Round 1, decreasing slightly to 88% in Rounds 2 and 3. In contrast, accuracy in detecting AD/LBD improved markedly—from 51% in Round 1 to 84% in Round 2, rising to 88% in Round 3 with AI support. Diagnostic confidence also increased over time: for normal/abnormal classification, it rose from 3.95 to 4.28 in Round 2, and further to 4.59 in Round 3 with AI support; for AD/LBD assessment, the scores rose from 3.71 to 4.06 in Round 2 and reaching 4.44 in Round 3 with AI support. Importantly, analysis of eight "borderline" cases in Model 1, which showed a mean probability score of 0.73 and odds ratio >1, suggests that lowering the probability threshold from 0.85 to 0.73 in Model 1 could improve model sensitivity.



CONCLUSION

This evaluation study demonstrates that incorporating AI model output significantly improved a trainee's performance in diagnosing AD/LBD and enhanced diagnostic confidence. These findings highlight the potential of AI tools to support more accurate dementia diagnosis and to facilitate learning and confidence-building among junior clinicians.

NEXT STEPS

- Fine tuning the Model 1 output performance by reducing the Threshold for improving diagnostic sensitivity.
- Model 1 retraining and further evaluation of the re-trained model.
- Further evaluation of both Model 1 and 2 with multiple readers.
- Regulatory approval.
- Deployment and Monitoring.

REFERENCES

The Economic Impact of Dementia (https://www.alzheimers.org.uk/sites/default/files/2024-05/the-annual-costs-of-dementia.pdf)