

#12503 - Artificial Intelligence for Amyloid PET Interpretation in Alzheimer's Disease: A Systematic Review

Callum Hill^{1*}, Huw Morgan^{2*†}, Emmanuel Ogunsakin^{1*}, Evie Day², Tunchanok Ngamsaowaros², Roberto Martinez Camacho³, Mahesan Niranjana², Sofia Michopoulou¹³

¹Faculty of Medicine, University of Southampton, Southampton, UK, ²School of Electronics and Computer Science, University of Southampton, Southampton, UK, ³Medical Physics, University Hospital Southampton NHS Foundation Trust, Southampton, UK, * Authors have contributed equally as joint first authors, † Corresponding author - hwcm1g20@soton.ac.uk

Introduction

- Amyloid PET imaging enables quantification of a patient's $A\beta$ burden.
- Artificial Intelligence (AI) integration with amyloid PET is gaining momentum for its potential to improve diagnostic accuracy, support early detection and predict treatment outcomes. However, the clinical implementation remains limited.

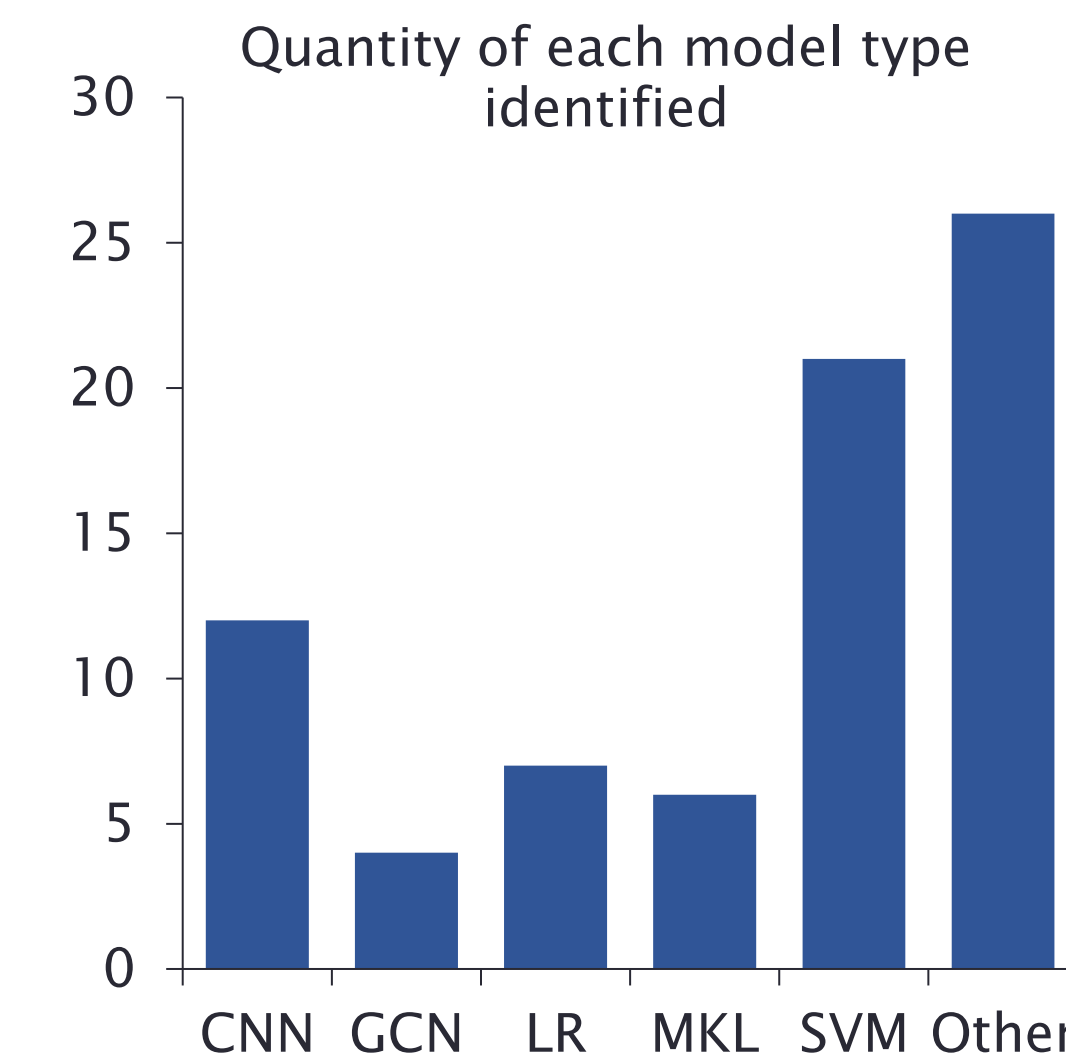
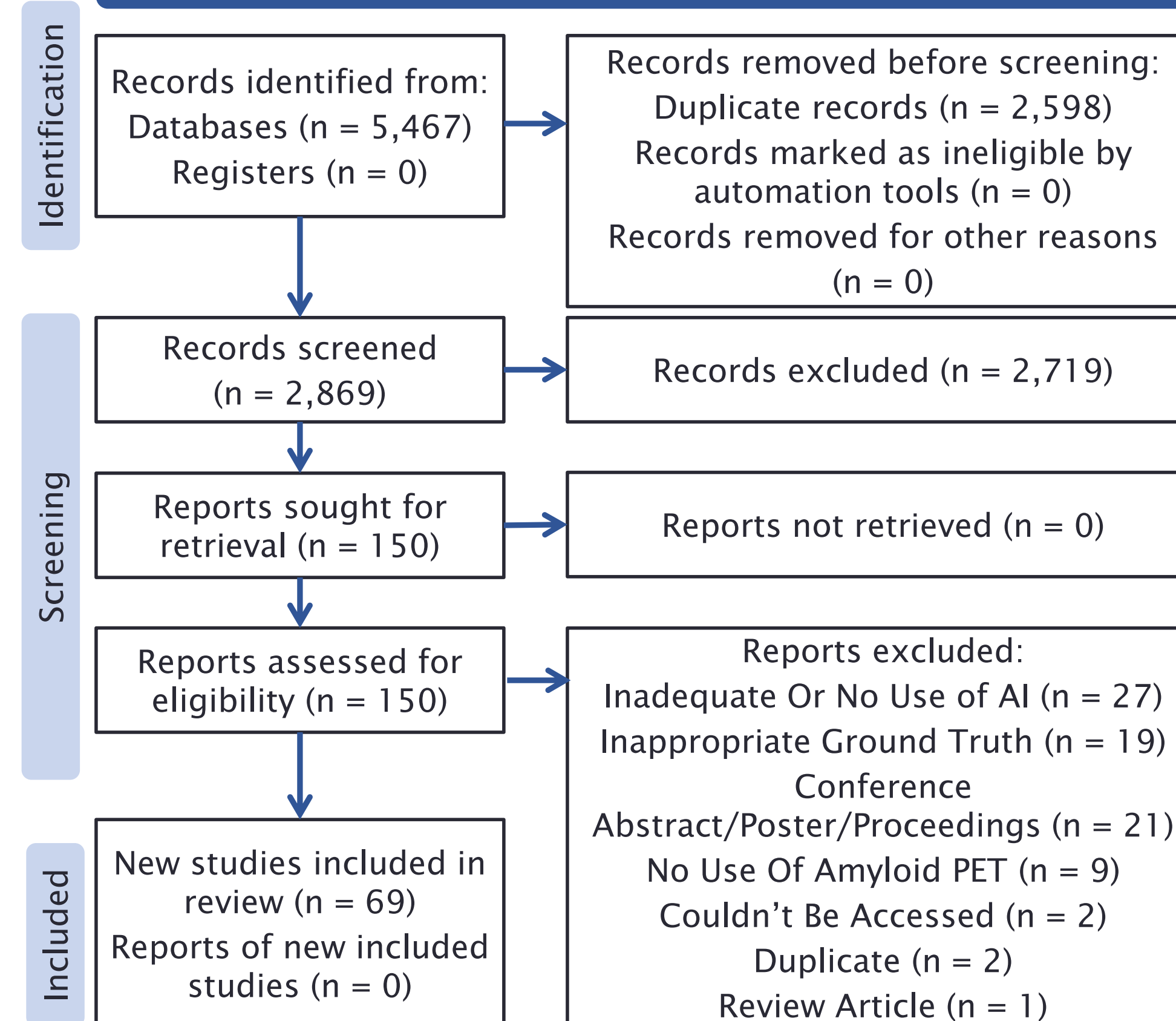
Aims

This systematic review highlights the methodological, technical and validation challenges that need to be addressed to enable clinical implementation.

Methods

- A comprehensive search was conducted across PubMed, EMBASE, Web of Science, Scopus, and IEEE Xplore.
- Included studies underwent narrative synthesis of clinical task, input data, methodology, performance metrics, and reported strengths/limitations.
- Meta-analysis was conducted for clinical task as it was sufficiently homogeneous.

Identification of new studies via databases and registers



Pooled estimates for model performance by task:

Tasks	ACC % (95% CI)	AUC (95% CI)
CN vs. AD	90.0 (85.0-94.9)	0.93 (0.88-0.99)
CN vs. MCI	82.6 (74.3-91.0)	0.72 (0.58-0.86)
MCI vs. AD	77.5 (70.2-84.8)	0.80 (0.66-0.95)

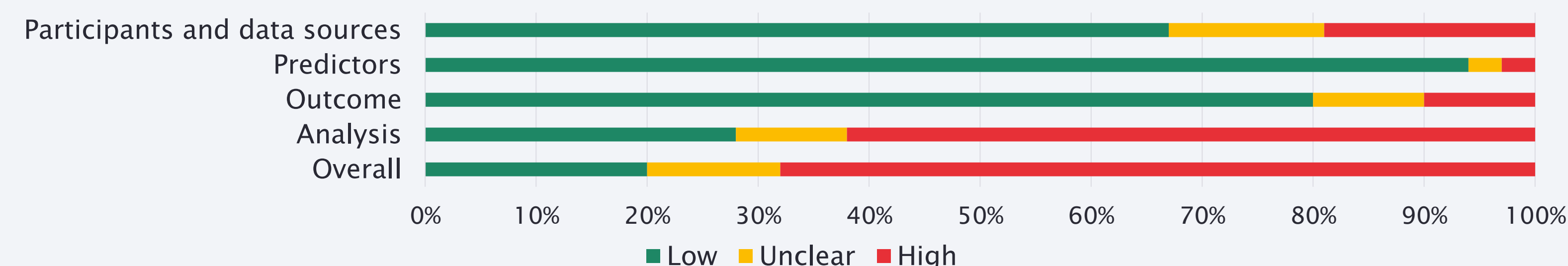
Results

- Most common algorithm used was SVM (30.0%) followed by CNN (17.1%). Most common dataset was ADNI (60.0%) followed by local cohorts (11.4%).
- Multimodality was considered by 15.7% of papers.
- High risk of bias was found in 47/69 papers (68.1%).
- Progress is recorded on PROSPERO database (scan QR to read).



Quality Assessment

- Quality assessments (risk of bias and applicability concern) were conducted using the PROBAST+AI model evaluation framework.
- Overall risk of bias across studies (displayed) was judged as high, largely due to analysis methods used and statistical considerations (Domain 4).
- The overall applicability concern across studies was judged to be low.



Conclusion

Papers reported high accuracy across cohorts but with many caveats. Recommendations for improving confidence include:

- Consistent reporting of metrics (ACC, AUC, SEN, SPE).
- Independent validation.
- Evaluation of uncertainty.
- Mention risks posed by AI and how to address them.