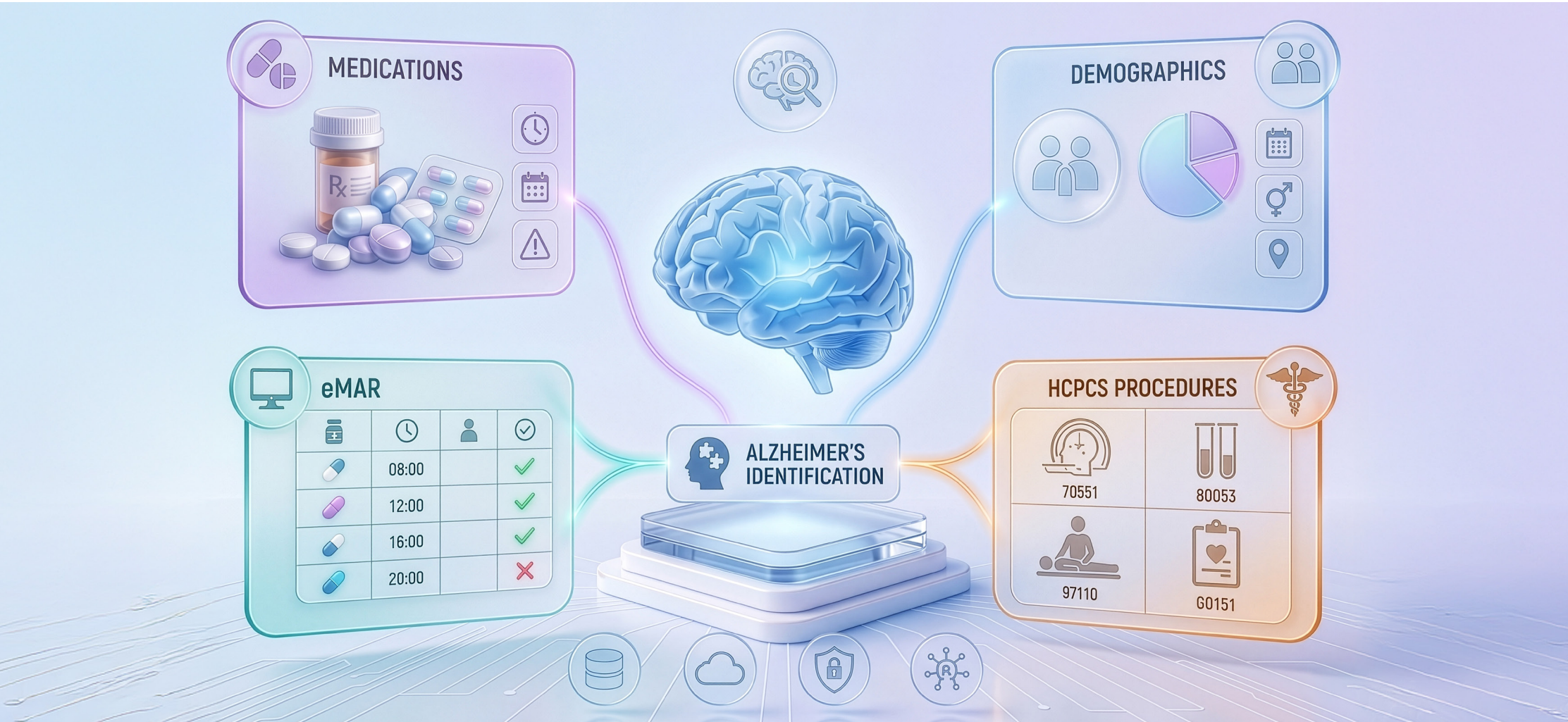




BACKGROUND



Identifying Alzheimer’s disease and related dementias (AD/ADRD) in real-world settings is often limited by the lack of formal neurocognitive assessments^{1, 2}. Inpatient electronic health records (EHRs) provide a scalable yet underutilized resource for detecting cognitive impairment through hospitalization patterns, medication use, and procedural history^{3, 4}.

EHR Modalities Used:
Patient-level features were extracted from inpatient hospitalizations in the MIMIC-IV v3.1 database⁵.

- Medications
- Demographics
- eMAR
- HCPCS Procedures

METHODS

- 1. Feature extraction** from four EHR modalities at the patient level using ICD hospitalization codes^{6, 7}.
- 2. Hierarchical classification framework** (see Table 1): controls versus cognitively impaired patients, followed by MCI versus dementia, and finally classification of dementia subtypes^{8, 9}.
- 3. Random forest classifiers** trained per modality and as an integrated multimodal model at each classification level¹⁰.
- 4. Feature importance analysis** to identify clinically interpretable signatures across modalities.

HIERARCHY STAGE	TASK	CLASSES	DEFINITION
General screening	Binary cognitive impairment detection	Normal cognition vs Cognitive impairment	Identifies the presence of any ICD-coded cognitive impairment
Intermediate stratification	Binary MCI vs dementia classification	Mild cognitive impairment vs Dementia	Differentiates mild cognitive impairment from established dementia
Granular classification	Dementia subtype classification	Alzheimer’s disease, Lewy body or Parkinson-related dementia, Vascular dementia, Frontotemporal dementia, substance– or exposure-related dementia, other or unspecified dementias	Classifies dementia into different subtypes

Table 1. Hierarchical classification framework

RESULTS

Integrated models combining all EHR modalities outperformed single-modality models and improved discrimination across cognitive status and dementia subtype classification. The most salient features and their relative importance are summarized in **Figure 1**.

- 1. Demographics:** Demographic features, particularly age, helped distinguish patients with cognitive impairment from controls¹¹ but showed limited ability to differentiate dementia subtypes.
- 2. eMAR:** Electronic medication administration records captured dynamic inpatient care patterns, including temporal irregularities in medication administration, medication burden, and pharmacy-related features that helped distinguish MCI from dementia¹².
- 3. Medications:** Prescription medication features showed a distributed importance profile across therapeutic class, route of administration, dosing frequency, and medication status, with notable relevance among neurological pharmacy categories^{13, 14}.
- 4. HCPCS Procedures:** HCPCS procedural codes reflected care intensity and system-level involvement, with cardiovascular, digestive, and hospital observation services serving as meaningful markers across all classification levels¹⁵.
- 5. Multimodal Integration:** Integrated multimodal approaches consistently outperformed classifiers based on any single modality, demonstrating superior predictive performance across all levels of the hierarchical classification framework^{16, 17}.

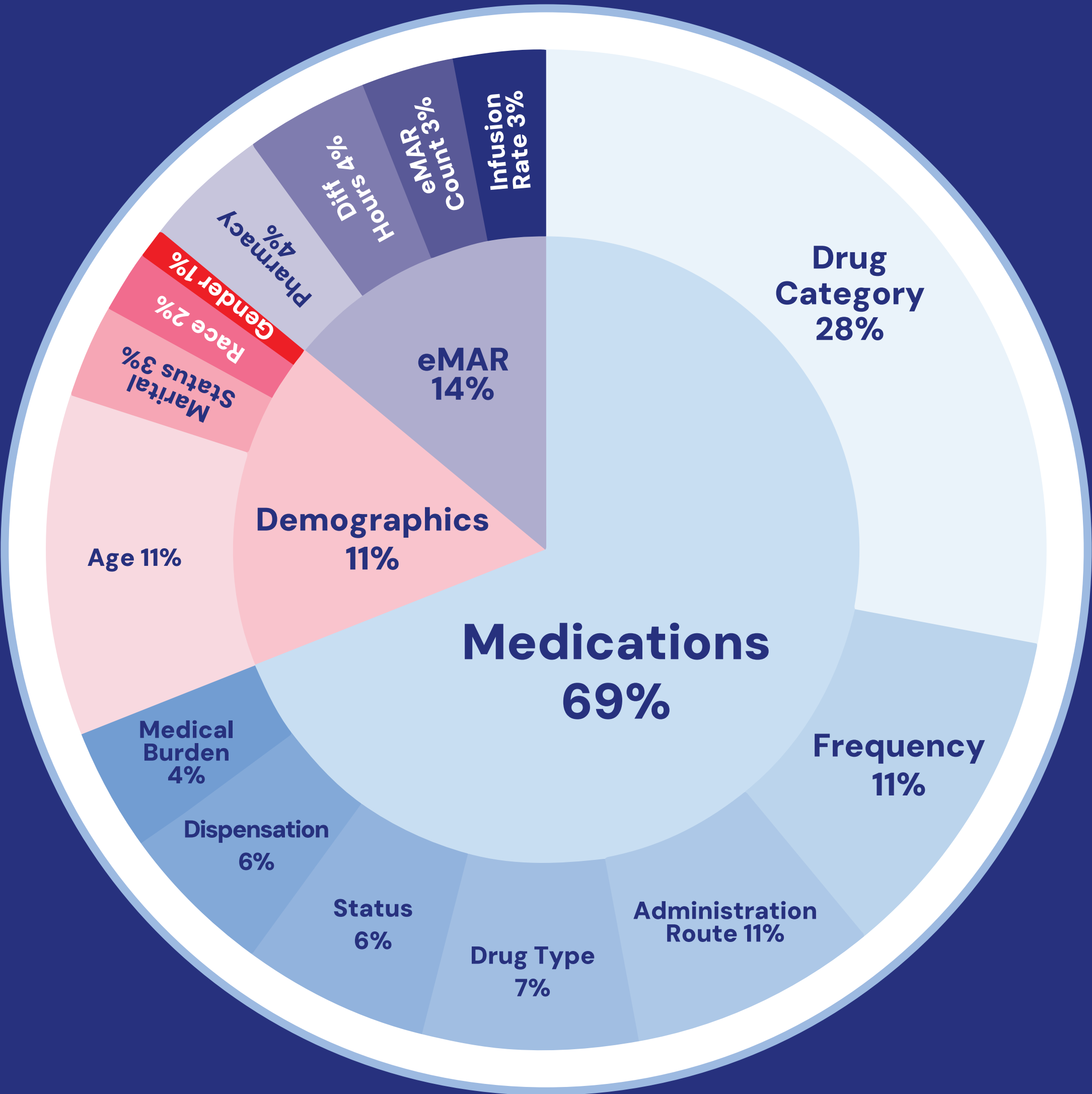


Figure 1. Feature importance per modality for the dementia subtypes classification

CONCLUSIONS

Machine learning models using routinely collected inpatient EHR data can identify interpretable, modality-specific signatures of AD/ADRD. Integrating complementary EHR modalities offers a scalable approach for Alzheimer’s dementia screening and dementia subtype differentiation in real-world settings without specialized cognitive testing.

DISCLOSURE OF FUNDING

This research was funded by IGC Pharma, LLC.

REFERENCES

1. V. G. Valcour, K. H. Masaki, J. D. Curb, and P. L. Blanchette, "The detection of dementia in the primary care setting," Archives of Internal Medicine, vol. 160, no. 19, pp. 2964–2968, 2000, doi: 10.1001/archinte.160.19.2964.

2. S. Irwin, E. Reddy, M. Doehler, C. C. Lima, and J. H. T. Bates, "Correlates of missed or late versus timely diagnosis of dementia in healthcare settings," Alzheimer’s & Dementia, vol. 21, no. 1, e14494, 2025, doi: 10.1002/alz.14494.

3. Q. Li, X. Yang, J. Xu, C. Guo, H. He, J. Bian, Y. Luo, and X. Du, "Early prediction of Alzheimer’s disease and related dementias using real-world electronic health records," Alzheimer’s & Dementia, vol. 19, no. 8, pp. 3506–3518, Aug. 2023, doi: 10.1002/alz.12967.

4. R. Wei, S. S. Buss, R. Milde, M. Fernandes, D. Sumsion, E. Davis, and M. B. Westover, "Automated phenotyping of mild cognitive impairment and Alzheimer’s disease and related dementias using electronic health records," Journal of Biomedical Informatics, vol. 164, p. 104813, Apr. 2025, doi: 10.1016/j.jbi.2025.104813.

5. A. E. W. Johnson, L. Bulgarelli, L. Shen, A. Gayles, A. Shammout, S. Horng, T. J. Pollard, B. Moody, B. Gow, L.–H. Lehman, L. A. Celi, and R. G. Mark, "MIMIC-IV, a freely accessible electronic health record dataset," Scientific Data, vol. 10, no. 1, p. 1, Jan. 2023, doi: 10.1038/s41597-022-01899-x.

6. S. S. Knox, T. Bhatt, T. Bhatt, M. C. Reid, and C. A. Metti, "AI approaches for phenotyping Alzheimer’s disease and related dementias using electronic health records," Alzheimer’s & Dementia: Translational Research & Clinical Interventions, vol. 11, no. 2, e70089, Apr. 2025, doi: 10.1002/trc2.70089.

7. A. M. Raskin, K. M. Beaulieu-Jones, A. B. Walling, C. Ritchie, L. E. Donahue, and R. A. Lasky-Su, "An algorithm to characterize a dementia population by disease subtype," Alzheimer Disease & Associated Disorders, vol. 33, no. 4, pp. 316–322, 2019, doi: 10.1097/WAD.0000000000000336.

8. R. O. Roberts and D. S. Knopman, "Classification and epidemiology of MCI," Clinics in Geriatric Medicine, vol. 29, no. 4, pp. 753–772, 2013, doi: 10.1016/j.cger.2013.07.003.

9. M. Zhang, K. Lofteröd, M. Boland, and A. Fagerbakke, "Cognitive performance classification of older patients using machine learning and electronic medical records," Scientific Reports, vol. 15, p. 6812, Feb. 2025, doi: 10.1038/s41598-025-90460-y.

10. L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

SIGNIFICANCE
AND FUTURE DIRECTIONS.

Routine EHR data provides an accessible and cost-effective pathway for detecting early cognitive impairment at scale across hospital systems. Future work will include additional modalities and validate this framework in external EHR cohorts and assess its integration into real-time clinical decision support for AD/ADRD screening and dementia subtype differentiation.

