

Enrolling an ADRC Clinical Trial-Ready Cohort: A Machine Learning Prediction Model

Daniel Willie-Permor, MD, MPH¹; Lilcelia A. Williams, PhD^{3,4}; Sarah B. Berman, MD, PhD^{2,4} Beth E. Snitz, PhD^{2,4} Robert A. Sweet, MD^{2,3,4} Ann D. Cohen, PhD^{3,4} Oscar L. Lopez, MD^{2,4}, C. Elizabeth Shaaban, PhD, MS, MPH^{1,4,5}
¹Department of Epidemiology, School of Public Health, University of Pittsburgh ²Department of Neurology, University of Pittsburgh ³Department of Psychiatry, University of Pittsburgh ⁴University of Pittsburgh Alzheimer's Disease Research Center ⁵Department of Health Promotion, School of Nursing, University of Pittsburgh



KEY OBJECTIVE Develop a machine-learning model from routine ADRC screening data to identify applicants without cognitive impairment , enabling more precisely targeted, trial-ready registry enrollment for early-stage ADRD.

TAKEAWAY MESSAGE: A handful of screening items can flag cognitive impairment as well as extensive testing, letting research centers target early-stage, trial-ready enrollment more efficiently

INTRODUCTION

- New treatments for Alzheimer's disease & related dementias (ADRD) are hypothesized to be most effective early, before substantial neurodegeneration and cognitive impairment.
- Research centers must enrich registries with people at the earliest, trial-ready end of the disease spectrum.
- Key question: how to target enrollment toward minimally- or un-impaired applicants without burdensome screening.

METHODS

- 192 participants screened at the University of Pittsburgh ADRC.
- Outcome: CDR Global score as no impairment (CDR 0) vs. any impairment (CDR ≥ 0.5).
- Random-forest classifiers over candidate predictors from routine screening domains.
- Split 50% train / 25% validation / 25% test; trees pruned by cross-validated error; tuned by out-of-bag error.
- Variable importance ranked by mean decrease in Gini (MDG).



Screen 192 applicants



Random forest (mtry = 5)



Held-out test (25%)

TABLE 1. *Baseline characteristics by CDR Global status*

Characteristic	CDR 0 (n = 49)	CDR ≥ 0.5 (n = 143)	p
Age, y - mean (SD)	68 (6)	71 (8)	0.032
Female - n (%)	40 (82)	88 (62)	0.012
Race - n (%)			<0.001
White	18 (37)	109 (76)	
Black or African American	30 (61)	29 (20)	
Other	0 (0)	4 (2.8)	
Education $\geq$ Bachelor's degree - n (%)	25 (51)	87 (61)	-
Living arrangement - n (%)			<0.001
With spouse / partner	20 (41)	99 (69)	
Alone	20 (41)	25 (17)	
Study partner - n (%)			0.036
Spouse / partner	16 (34)	80 (56)	
Child	14 (30)	34 (24)	
Friend / other	17 (36)	28 (20)	
MMSE - mean (SD)	28.0 (1.5)	25.1 (4.6)	<0.001

Values are mean (SD) or n (%). p from Wilcoxon rank-sum, Pearson χ^2 , or Fisher's exact test. "Other" race = Asian and multiracial; education collapsed from 8 categories.

0.87

VALIDATION
ACCURACY

0.81

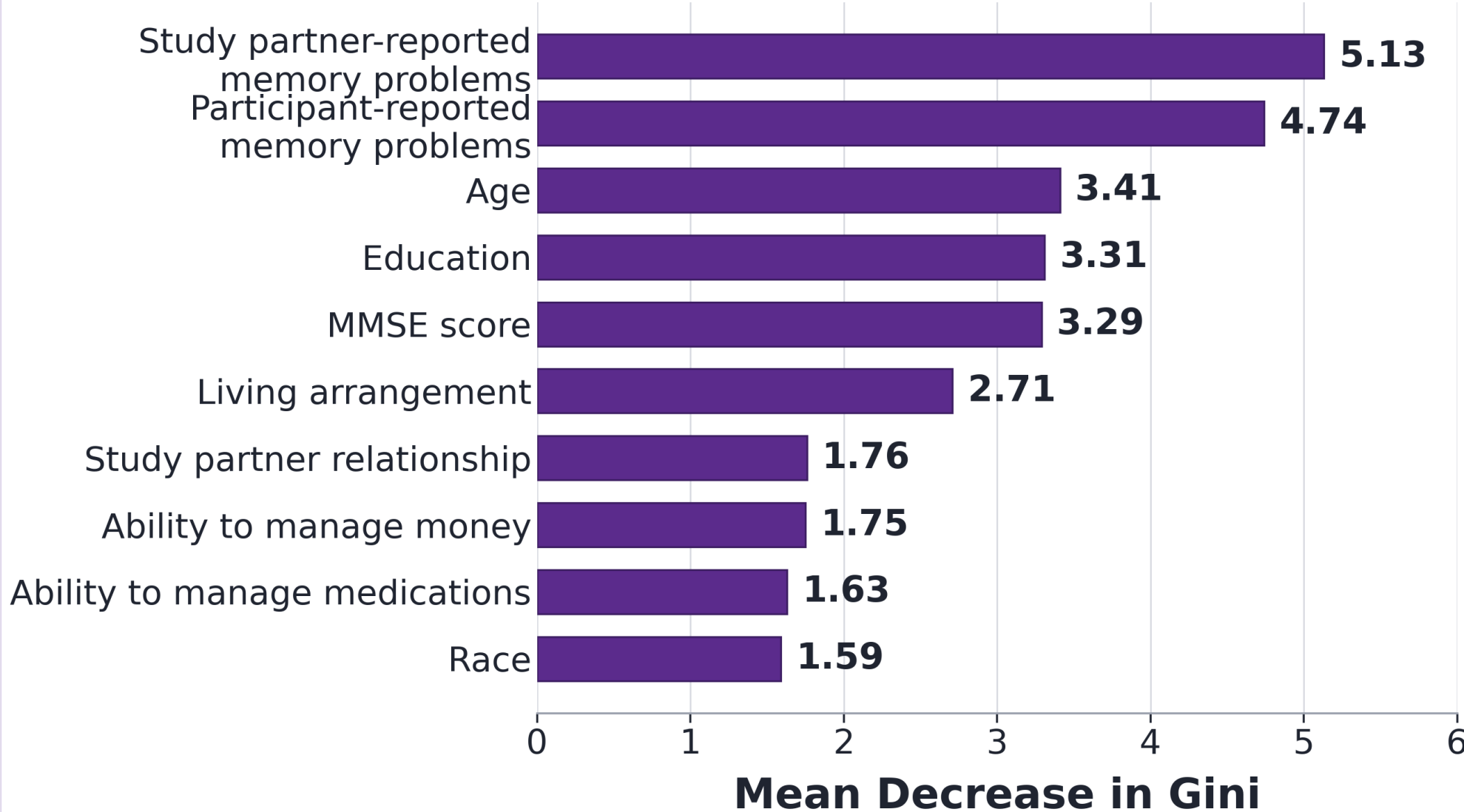
TEST
ACCURACY

0.73

TEST BALANCED
ACCURACY

Figure 1: Variable Importance for Predicting CDR Impairment

Random forest (mtry = 5) — ranked by Mean Decrease in Gini



Higher values indicate greater contribution to classification in the fitted random forest.

*mtry = number of candidate predictors randomly sampled at each tree split (a random-forest tuning parameter)

RESULTS

- Predictive accuracy relied on a small subset of variables, not the full candidate set.
- Strongest contributors to classification: study-partner- and participant-reported memory problems.
- Next most important: age, education, MMSE, and living arrangement.
- Functional items (managing money / medications), study-partner relationship, and race added smaller contributions.

CONCLUSIONS

- Brief, targeted screening - cognitive complaints, living arrangement, and a few functional items - identifies trial-ready applicants about as well as a more extensive screening questionnaire.
- ADRCs could streamline registry recruitment for early-stage ADRD trials using this compact, low-burden item set.



CONTACT



Daniel Willie-Permor, MD, MPH

drwillie.permor@pitt.edu

Scan to follow up

ALZHEIMER'S ASSOCIATION | AAIC >26

Funding: University of Pittsburgh Alzheimer's Disease Research Center, P30-AG06468.