The Impact of Test-Scoring Models on MCI Diagnosis

David L. Woods, PhD[1], Juliana Baldo, PhD[2], Kathleen Hall, PhD[1], Michael Blank, PhD[1], Kristin Geraci, PhD[1], Miranda Miranda[1], Jas M. Chok[3], Sandy J. Lwi, PhD[2], Isabella Jaramillo[1], Krista Schendel, PhD[2], , Maria Spinelli[2], , Timothy J Herron[2], and David K Johnson, PhD[4] (1)Neurobehavioral Systems, Inc, Berkeley, CA, USA, (2)Veterans Affairs Northern California Health Care System, Martinez, CA, USA, (3)University of Chicago, Chicago, IL, USA, (4)UC Davis Alzheimer's Disease Center, Walnut Creek, CA, USA

Introduction

The importance of the early detection of mild cognitive impairment (MCI) has increased with the development of amyloid-clearing therapies to slow cognitive decline**.** However, three challenges make traditional scoring models ill-suited for early MCI detection:

- **Insensitivity:** Many demographic factors that significantly influence performance are missing from traditional scoring models, increasing unexplained variance.
- **Racial bias:** MCI diagnoses are more prevalent among patients from minority communities when using traditional scoring models
- **Cognitive-reserve bias:** Patients with high cognitive reserve must experience substantial cognitive decline to score at MCI levels when using traditional scoring models

Methods

- 798 older participants (mean age 65.8 years) completed three 90-minute test sessions using the computerized California Cognitive Assessment Battery (CCAB). CCAB contains 32 tests in a variety of cognitive domains and response modalities.
- Demographic information (see Table 1) was gathered with questionnaires.
- An estimate of premorbid verbal IQ was obtained with an adaptive 4-minute vocabulary test.
- Each CCAB test produces one or more core test measures (e.g., mean span in digit span). For each participant, an omnibus ("OMNI") z-score was obtained by averaging unregressed z-scores from the 120 individual test scores.

Models

Unadjusted omnibus z-scores were corrected for demographic influences using three linear models that differed in the predictors used, and selection procedure:

- An *Age-only* model (A- model)
- An *Age, Education, and Gender* model (AEG-model)
- A *Comprehensive* model (C-model), using 10 possible predictors

Comprehensive (C-) Model

The C-model included 10 predictors, one of which is the vocabulary test score, which serves as an estimate of premorbid verbal IQ. Significant predictors for the C-model were identified with LASSO (least absolute shrinkage and selection operator). The number of LASSO-selected predictors (mean 5.94 selected predictors, range 1-10) varied by test measure. Mean coefficients for the predictors were extracted from linear model results from 1000 random samples of the normative population.

All three models (A-, AEG-, and C-) were then analyzed for goodness of fit.

Results

Factor	Coefficient
Age	-0.44
Education	0.07
Gender (female)	0.18
Vocabulary	0.47
Race (Black)	-0.54
Race (Asian)	-0.11
Race (other)	-0.21
Computer use	0.08
Daily medications	-0.10
Socioeconomic status	0.09

Table 1. COEFFICIENTS IN THE COMPREHENSIVE SCORING MODEL. Predictors of unregressed omnibus z-scores were selected with LASSO at lambda = 1.0 SE with the constraint that they contributed to at least 85% of solutions in random samples. White participants served as the reference population for the race factor.

Predictors	Adj R^2	RMSE
None	0.00	1.00
Age only	0.14	0.93
AEG	0.28	0.85
Comprehensive	0.61	0.62

Table 2. GOODNESS OF FIT METRICS. Adjusted r^2 and Root Mean Square Error (RMSE) for the three different models.

Model	White MCI	Non-white MCI	Ratio
A	1.31%	9.96%	7.62
AEG	2.94%	10.37%	3.52
C	7.52%	6.50%	0.87

Table 3. RACIAL BIAS. The incidence of OMNI MCI scores in white and non-white participants. Non-white participants had higher MCI incidence that whites for A- and AEG models

Model	High Vocab MCI	Low Vocab MCI	Ratio
A	0.76%	12.35%	16.17
AEG	2.04%	12.84%	6.31
C	6.87%	6.91%	1.01

Table 4. MCI DETECTION AND COGNITIVE RESERVE. Incidence of OMNI MCI scores in participants with large (top 50%) and small (bottom 50%) vocabularies. Individuals with smaller vocabularies were overrepresented in A- and AEG MCI populations.

Summary

- Ten predictors made significant contributions to the omnibus z-score model, as shown in Table 1. Vocabulary, Age, and Race had the greatest influence
- The C-Model accounted for more than four times the variance of the A-model, and more than twice the variance of the AEG-model, as shown in Table 2.
- Racial bias in MCI incidence was minimal for the C-model, but substantial for the A- and AEG models, as shown in Table 3.
- A- and AEG- models showed reduced sensitivity to MCI in individuals with high cognitive reserve (estimated from vocabulary scores), while the C-model's MCI-detection sensitivity was unaffected (Table 4).

Discussion

- Conventional scoring models, such as the A (age only) and AEG (age, education, gender), reduce scoring sensitivity by ignoring factors (e.g., vocabulary, race, prescription medications, etc.) that correlate significantly with cognitive test scores.
- Failing to include Race in scoring models inflates racial disparities in MCI classification.
- Finally, failing to include a vocabulary measure in scoring models compromises the detection of MCI in individuals with high cognitive reserve.

Contact us

drdlwoods@neurobs.com for reprints
ccabresearch.com
neurobs.com

**Supported by NIA R44AG062076
NIA R44AG080951**