

UC San Diego

UC San Diego Previously Published Works

Title

Extensive memory testing improves prediction of progression to MCI in late middle age

Permalink

<https://escholarship.org/uc/item/6cn786c3>

Journal

Alzheimer's & Dementia Diagnosis Assessment & Disease Monitoring, 12(1)

ISSN

2352-8729

Authors

Gustavson, Daniel E
Elman, Jeremy A
Sanderson-Cimino, Mark
et al.

Publication Date

2020

DOI

10.1002/dad2.12004

Peer reviewed

Extensive memory testing improves prediction of progression to MCI in late middle age

Daniel E. Gustavson^{1,2} | Jeremy A. Elman² | Mark Sanderson-Cimino² |
 Carol E. Franz² | Matthew S. Panizzon² | Amy J. Jak^{2,3} | Chandra A. Reynolds⁴ |
 Michael C. Neale⁵ | Michael J. Lyons⁶ | William S. Kremen^{2,3}

¹Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee

²Department of Psychiatry, Center for Behavior Genetics of Aging, University of California, San Diego, La Jolla, California

³Center of Excellence for Stress and Mental Health, Veterans Affairs San Diego Healthcare System, La Jolla, California

⁴Department of Psychology, University of California, Riverside, California

⁵Virginia Institute for Psychiatric and Behavior Genetics, Virginia Commonwealth University, Richmond, Virginia

⁶Department of Psychological and Brain Sciences, Boston University, Boston, Massachusetts

Correspondence

Dr. Daniel Gustavson, Department of Otolaryngology, Vanderbilt University Medical Center, 1215 21st Ave South (10420B MCE), Nashville, TN 37232, USA.

Email: daniel.e.gustavson@vumc.org

Funding information

This research was supported by Grants R01 AG050595, R01 AG022381, and R01 AG059329 from the National Institutes of Health.

Abstract

Introduction: Predicting risk for Alzheimer's disease when most people are likely still biomarker negative would aid earlier identification. We hypothesized that combining multiple memory tests and scores in middle-aged adults would provide useful, and non-invasive, prediction of 6-year progression to MCI.

Methods: We examined 849 men who were cognitively normal at baseline (mean age \pm SD = 55.69 \pm 2.45).

Results: California Verbal Learning Test learning trials was the best individual predictor of amnesic MCI (OR = 4.75). A latent factor incorporating seven measures across three memory tests provided much stronger prediction (OR = 9.88). This compared favorably with biomarker-based prediction in a study of much older adults.

Discussion: Neuropsychological tests are sensitive and early indicators of MCI risk at an age when few individuals are likely to have yet become biomarker positive. The single best measures may appear time- and cost-effective, but 30 additional minutes of testing and use of multiple scores within tests provide substantially improved prediction.

KEYWORDS

Alzheimer's disease, mild cognitive impairment, neuropsychology

The pathogenesis of Alzheimer's disease (AD) begins decades before the onset of dementia, so it is necessary to identify risk factors as early as possible.¹⁻⁴ The recent A/T/N (amyloid/tau/neurodegeneration) framework emphasizes biomarkers in an effort to improve early identification and move toward a biological diagnosis.^{5,6} Preclinical AD is stage 1 of the A/T/N framework staging, defined as being amyloid positive but still cognitively normal (CN). However, being biomarker positive indicates that a significant amount of disease progression has already taken place. The ability to identify individuals at risk before they become biomarker positive would thus be of great potential value. It would also be useful to identify people who are likely to be at ele-

vated risk before embarking on costly and invasive biomarker testing. This view echoes those of several research groups that have noted there is pressing need to identify tests that are non-invasive, low-cost, and can improve earlier identification of risk for AD.^{4,7-9}

Episodic memory is an effective early predictor of progression to Alzheimer's disease (AD).¹⁰⁻¹⁵ It should also be a good predictor of amnesic mild cognitive impairment (MCI). Memory is severely impaired in AD, and MCI diagnoses are in part based on impaired memory performance.¹⁶ However, comprehensive neuropsychological assessment of memory (and other cognitive domains) is often lacking in longitudinal studies aimed at detecting which individuals are at the

greatest risk for MCI/AD. It is also important to recognize that CN individuals are not a homogeneous group. Examining variability within the range of normal cognitive function may be useful for early prediction of MCI.

Efforts to improve longitudinal prediction of MCI or AD using cognitive measures have often focused on evaluating which measure provides the best prediction,^{15,17,18} This approach rests on an unspoken assumption that including the lesser predictors in the model may not improve (or may even hurt) predictive ability. However, we have shown that multiple memory tests (and scores within tests) have both common and unique genetic and environmental influences,^{19,20} suggesting that the right combination of tests might enhance prediction. Here, we tested the hypothesis that combining multiple memory measures (within-test and across multiple tests) provides stronger and more robust prediction than the single best measure of progression to MCI.

We evaluated our hypothesis in a community sample of middle-aged adults from the Vietnam Era Twin Study of Aging (VETSA) who were all cognitively normal (CN) at baseline. The participants completed multiple memory tests at mean ages of 56 and 62. Examining these associations in midlife is important because improving treatment efficacy may depend on early intervention,^{7,21} and cognitive abilities may already be subtly declining by the late 50s.^{20,22–24} Yet there is exceedingly little focus on prediction of MCI, particularly in adults this young. In these analyses, we compared two approaches to aggregating measures: z-score composites and factor scores. We expected the factor score approach to have the strongest prediction because it weighs more strongly the measures that are the best indicators of the latent memory construct.

1 | METHOD

1.1 | Participants

Analyses focused on 849 individuals from the longitudinal VETSA who were CN at wave 1, returned to complete the wave 2 assessment approximately 6 years later, and had data for all covariates. Participants were recruited randomly from a previous large-scale study of Vietnam-Era Twin (VET) Registry participants.²⁵ All served in the U.S. Armed Forces at some time between 1965 and 1975; nearly 80% did not serve in combat or in Vietnam.^{26,27} Participants are generally representative of American men in their age group with respect to health and lifestyle characteristics.²⁸ All participants provided written informed consent and the study was approved by local institutional review boards at the University of California, San Diego and Boston University.

Of the 1237 individuals who completed the VETSA protocol at the first wave, 107 (8.6%) were excluded from the analysis of progression to MCI because they had MCI at wave 1. This left 1130 (91.4%) CN individuals, 906 (80.2%) of whom returned for wave 2. Of these 906, a total of 57 were excluded for missing covariates. This left 849 (93.7% of 906) and 75.1% of the total CN individuals at wave 1 for the analyses of progression to MCI. To compute standardized memory scores used

RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed longitudinal studies on progression to mild cognitive impairment (MCI). Few of these studies have focused on midlife predictors of early MCI and none have systematically examined the predictive efficacy of multiple measures within the same cognitive domain.
2. Interpretation: Results demonstrate that compared to single memory tests, an additional 30 minutes of memory testing can result in two to three times better prediction of 6-year progression to amnesic MCI in cognitively normal (CN) individuals who are only in their mid-50s at baseline. This finding has important implications for early identification given that prediction compared favorably to biomarker-based prediction in a study of adults who were approximately 15-years older.
3. Future directions: Few individuals are likely to be biomarker positive in their 50s, but episodic memory may also predict later progression to biomarker positivity. Combining more extensive memory testing with biomarker assessment may substantially improve earlier risk prediction.

in the odds ratios (ORs) for the primary analyses, we used data from all 1237 individuals at wave 1, plus an additional 53 attrition replacements who completed the VETSA protocol for the first time during the wave 2 assessment but were in the age range of participants at wave 1 (total N = 1290). All participants with available data were used so that ORs should better reflect those who would be obtained from a large population sample.

1.2 | Episodic memory measures

Episodic memory was measured at both waves with the logical memory (LM) and visual reproductions (VR) subtests of the Wechsler Memory Scale–Third Edition (WMS-III)³⁰ and the California Verbal Learning Test–Second Edition (CVLT-II).³¹ For LM and VR, we examined immediate recall and delayed recall measures. For the CVLT, we examined short delay and long delay free recall, and total score for learning trials 1 through 5. In analyses involving MCI, all memory measures were z-scored and transformed so that ORs reflect the increase in odds of MCI for every decrease of one standard deviation (SD) in memory performance (in relation to the full sample of N = 1290 at wave 1). These standardization procedures were also conducted for the aggregated memory measures described next.

In addition to examining the predictive ability of each memory measure alone we combined measures in two ways. First, we created z-score composites for measures within a given test (eg, LM

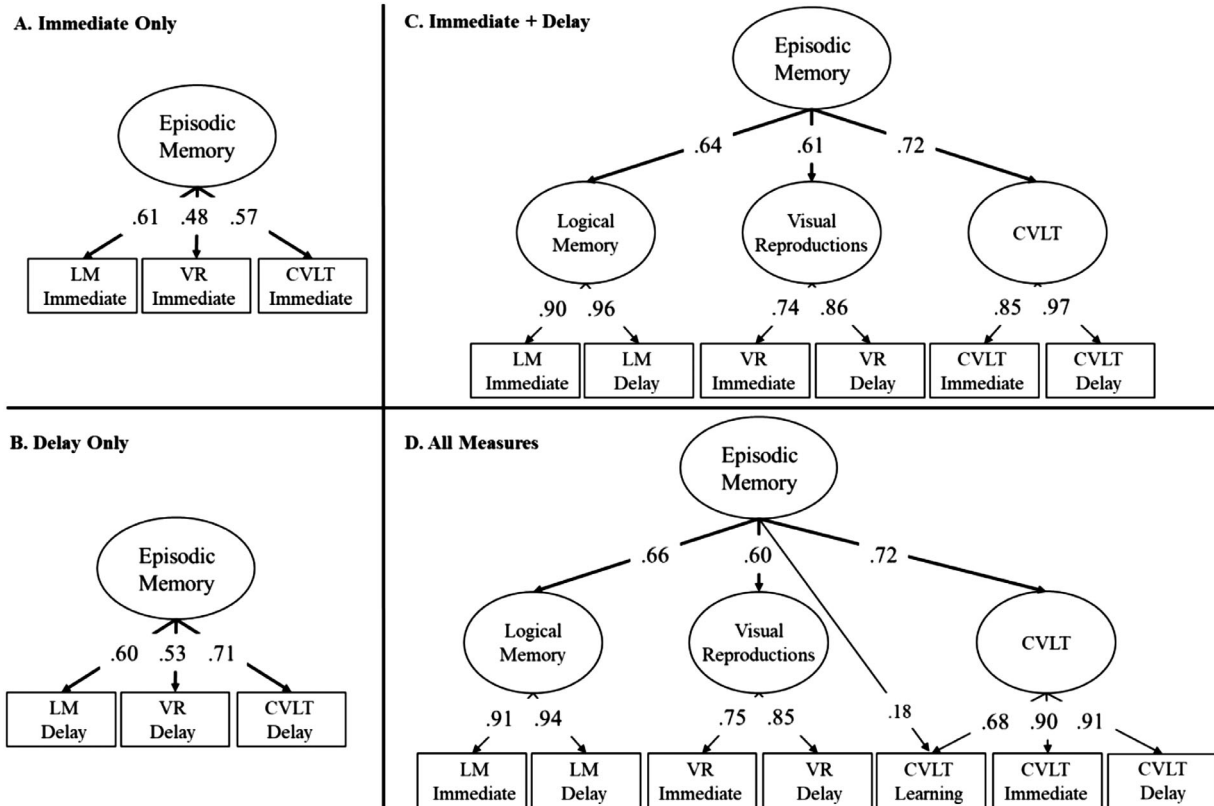


FIGURE 1 Structural equation models used to create factor scores for episodic memory at wave 1. In each model, variance explained in a given memory measure (rectangle) by latent factors (ovals) can be computed by squaring the factor loading on that factor. Factor scores for the highest-level episodic memory latent factor in each model were exported in Mplus and used as a continuous variable in separate logistic regression analyses involving MCI. All factor loadings were significant ($P < 0.05$) and all models fit the data well

immediate and delayed recall), and comparable measures across tests (eg, LM immediate recall, VR immediate recall, CVLT short delay free recall). We also created z-score composites that combined the six short/long delay conditions across all tests or all seven memory measures (including CVLT learning trials).

Second, we created factor scores from latent memory variables. Latent factors were exported from structural equation models in MPlus version 7.2,³² based in part on those reported in earlier work from this sample (see supplement for more information).^{19,20} These factor scores are similar to the z-score composites, but measures are weighted according to their factor loadings on the latent memory factor. Latent factors are displayed in Figure 1 and were also based on the full wave 1 sample ($N = 1289$). Each model had good fit to the data based on standard structural equation metrics, including root mean square error of approximation values of < 0.06 , and the Comparative Fit Index values of > 0.95 .³³

1.3 | Mild cognitive impairment diagnoses

MCI was diagnosed using the Jak-Bondi approach.^{4,16,34} Impairment in a cognitive domain was defined as having at least two tests > 1.5

SD below the age- and education-adjusted normative means after accounting for "premorbid" cognitive ability by adjusting neuropsychological scores for performance on a test of general cognitive ability that was taken at a mean age of 20 years (see Supplemental Method for more information).^{35,36} This test is correlated $r = .84$ with Wechsler Adult Intelligence Scale.^{35,36} The adjustment for age 20 cognitive ability ensures that the MCI diagnosis is capturing a decline in function rather than long-standing low ability. The validity of the VETSA MCI diagnoses is supported in the present sample by evidence of reduced hippocampal volume in those diagnosed with amnesic MCI.³⁷ In addition, higher AD polygenic risk scores were associated with significantly increased odds of MCI in this sample,³⁸ indicating that the MCI diagnosis is genetically related to AD. MCI diagnoses at wave 2 were also based on measures that were adjusted to account for practice effects,³⁹ leveraging data from attrition replacement participants who completed the task battery for the first time at wave 2 ($N = 179$) to estimate the increase in performance expected in returnees who completed the tests twice.

Because we were interested in transition to MCI, analyses included only individuals who were CN at wave 1 and had data for all covariates. Of the 849 returnees meeting this criterion, 45 (5.3%) progressed to amnesic MCI, and 41 (4.8%) progressed to non-amnesic MCI.

TABLE 1 Demographic and clinical characteristics and covariates included in analyses involving mild cognitive impairment (MCI)

Demographic variable	Remained cognitively normal (N = 763)		Progressed to amnestic MCI (N = 45)		Progressed to non-amnestic MCI (N = 41)		p (CN vs aMCI)	p (CN vs nMCI)
	M	SD	M	SD	M	SD		
Lifetime education (years)	14.03	2.16	13.58	1.75	13.49	1.83	0.191	0.098
Age (at wave 1)	55.69	2.45	56.37	2.48	56.68	2.41	0.102	0.024
Age interval (wave 2–wave 1)	5.75	0.70	5.76	0.67	5.65	0.54	0.907	0.400
Depression symptoms (at wave 1)	7.69	7.14	7.32	7.03	9.71	9.66	0.858	0.152
Ethnicity (% white non-Hispanic)	91.34	-	84.44	-	80.49	-	0.119	0.031
APOE status (% ε4 positive)	30.80	-	33.33	-	21.95	-	0.480	0.193
Diabetes (% yes at wave 1)	10.35	-	11.11	-	17.07	-	0.868	0.267
Hypertension (% yes at wave 1)	58.32	-	64.44	-	73.17	-	0.230	0.069

Note: The final two columns display the *p* value for comparisons between individuals who remained cognitively normal (N = 763) and either progressed to amnestic MCI (aMCI; N = 45) or non-amnestic MCI (nMCI; N = 41), controlling for clustering of data within families. Significant group differences are displayed in bold (*p* < 0.05). Lifetime education was the number of years of school completed. Depression symptoms were measured with the Center for Epidemiologic Studies Depression Scale,⁴¹ with scores above 15 indicating risk for clinical depression. Rate of APOE ε4 positivity across the entire sample (30.5%) is similar to 28.9% in the UK Biobank, the largest population sample of APOE prevalence to date (n = 326,535).⁵⁰

1.4 | Data analysis

All statistical analyses were conducted using R version 3.5.1. Analyses involving MCI were conducted with mixed-effects logistic regression using the lme4 package.⁴⁰ In these analyses, we controlled for wave 1 age, the time interval between assessments, education, race/ethnicity (white non-Hispanic vs other), wave 1 diabetes (yes/no), wave 1 hypertension (yes/no), apolipoprotein E (APOE) ε4 allele status (ε4+ vs ε4-), and wave 1 depression symptoms based on Center for Epidemiologic Studies Depression Scale (CES-D).⁴¹ Diabetes and hypertension status were based on whether the participant (1) reported being diagnosed by a doctor, (2) reported that they were currently taking medication for diabetes or high blood pressure, and/or (3) reported whether they had high blood pressure on the day of testing (hypertension only). Finally, twin pair identification number was included as a random effect to account for the clustering of data within families. The lme4 package uses list-wise deletion with missing observations, and reports profile-based 95% confidence intervals (95% CIs).

The predictive utility of all models was assessed with average area under the curve (AUC) calculated from fourfold cross-validation with 10 repeats. Random effects were not included in the cross-validated models due to difficulty obtaining model convergence with a decreased sample size of MCI cases in each fold. However, we found that removing random effects from the full models resulted in decreased odds ratios of cognitive scores. Therefore, these AUC values may represent conservative estimates.

2 | RESULTS

2.1 | Descriptive statistics

There were no significant differences between CN and amnestic MCI groups in any demographic or clinical characteristics (Table 1). Compared to CN returnees, those diagnosed with non-amnestic MCI at

wave 2 were older at baseline (*p* = 0.024) and a smaller proportion were white non-Hispanic (*p* = 0.031). APOE ε4 was not a significant predictor, but it was in the expected direction for amnestic MCI (33% in amnestic MCI vs 30% in controls). Descriptive statistics for memory measures at wave 1 are displayed in Table 2.

2.2 | Six-year prediction of progression to mild cognitive impairment

The results of the primary analyses are displayed in Table 3. Each cell displays an OR from a separate longitudinal logistic regression in which a memory measure predicts progression to amnestic MCI (Table 3A) or non-amnestic MCI (Table 3B) 6 years later. As shown in Table 3A, each of the seven individual memory measures at wave 1 significantly predicted progression to amnestic MCI at wave 2. The learning trials measure of the CVLT provided the strongest individual prediction (OR = 4.75, 95% CI 2.59–12.44), whereas the immediate recall condition of VR was the weakest individual predictor (OR = 1.91, 95% CI 1.21–3.43). However, the strongest prediction came from the latent factor score that incorporated all seven dependent measures at wave 1 (OR = 9.88, 95% CI 4.39–37.72). Cross-validated receiver operating characteristic (ROC) curves for these three models are displayed in Figure 2A. VR immediate recall had a significantly lower AUC (.570) than CVLT learning trials (.741) or the full latent factor score (.796), both *Z* > 2.65, *p* < 0.008, but the factor score outperformed the CVLT learning trials at a trend level only (*Z* = 1.81, *p* = 0.070).

Although most CIs overlapped, long delay recall conditions tended to be stronger predictors of progression to MCI than short delay conditions (average increase of 11.5%, range 4.4%–23.0%). OR estimates for composites greatly outperformed the individual measures. Aggregating measures within the same test resulted in a small increase in ORs (average increase of 21.2%, range 2.1%–55.1%). Aggregating measures of the same type across tests resulted in a larger increase in

TABLE 2 Descriptive statistics for episodic memory measures at baseline (mean age 56)

Memory variable	N	M	SD	Range	Skewness	Kurtosis
Logical memory						
Immediate recall	1281	23.47	6.16	4, 44	-0.04	-0.10
Delayed recall	1279	20.01	6.63	0, 41	-0.10	-0.13
Visual reproductions						
Immediate recall	1284	78.24	12.41	21, 103	-0.58	0.48
Delayed recall	1283	54.75	19.51	0, 100	-0.15	-0.44
California Verbal Learning Test						
Learning trials	1270	42.84	8.51	18, 74	0.09	-0.06
Short delay free recall	1270	8.64	2.74	1, 16	0.07	-0.21
Long delay free recall	1269	9.06	2.89	0, 16	0.00	-0.30

Note: In all analyses involving MCI, dependent measures were standardized and reverse scored so that odds ratios reflect increase in risk of MCI at lower levels of cognitive ability. Shown here are descriptive statistics for all individuals who completed the wave 1 protocol (used to standardize data and create z-score composites and latent factor scores). Distributional characteristics remained acceptable for the subset of 849 individuals in the analyses involving progression to MCI (range skewness = $-.45$ to $.22$; range kurtosis = $-.36$ to $.05$; see Supplementary Table S1).

TABLE 3 Odds ratios and 95% confidence intervals from logistic regressions of mild cognitive impairment (MCI) predicted by baseline memory measures

Dependent measure	Logical memory	Visual reproductions	CVLT	z-score composites	Latent factor scores
A. Prediction of amnesic MCI					
Immediate recall only	3.01 [1.79, 6.44]	1.91 [1.21, 3.43]	3.61 [2.00, 9.39]	5.27 [2.78, 15.29]	5.67 [2.92, 16.98]
Delayed recall only	3.22 [1.89, 6.83]	2.35 [1.50, 4.38]	3.77 [2.14, 8.48]	5.86 [3.09, 15.83]	6.45 [3.30, 17.99]
Learning trials only	-	-	4.75 [2.59, 12.44]	-	-
Immediate and delay recall ^a	3.28 [1.93, 7.05]	2.40 [1.51, 4.56]	4.26 [2.29, 10.92]	6.57 [3.32, 19.91]	6.91 [3.47, 20.20]
Immediate and delay recall + learning trials ^a	-	-	5.60 [2.83, 16.06]	8.82 [4.04, 32.61]	9.88 [4.39, 37.72]
B. Prediction of non-amnesic MCI					
Immediate recall only	1.10 [.71, 1.72]	2.02 [1.33, 3.28]	1.29 [.85, 1.99]	1.71 [1.08, 2.85]	1.63 [1.03, 2.69]
Delayed recall only	1.11 [.71, 1.77]	1.55 [1.01, 2.47]	1.78 [1.15, 3.01]	1.72 [1.07, 2.96]	1.86 [1.16, 3.21]
Learning trials only	-	-	1.03 [.66, 1.57]	-	-
Immediate and delay recall ^a	1.11 [.71, 1.76]	1.96 [1.27, 3.25]	1.53 [1.01, 2.47]	1.78 [1.11, 3.04]	1.91 [1.19, 3.31]
Immediate and delay recall + learning trials ^a	-	-	1.36 [.89, 2.15]	1.64 [1.02, 2.76]	1.56 [.99, 2.60]

Note: Each cell displays an odds ratio (OR) from a separate analysis in which that memory measure (or combination of measures) predicts progression to amnesic MCI (A) or non-amnesic MCI (B). Significant ORs are displayed in bold ($P < 0.05$). Like all individual measures, z-score composites and factor scores were scored such that higher ORs indicate greater risk for amnesic MCI (aMCI) or non-amnesic MCI (nMCI) at -1 SD for that variable.

^aIndicates measures in this row were also based on z-score composites (eg, LM immediate and LM delayed recall in the first column), except for the final column, which was based on factor scores (from the models displayed in Figure 1C and D).

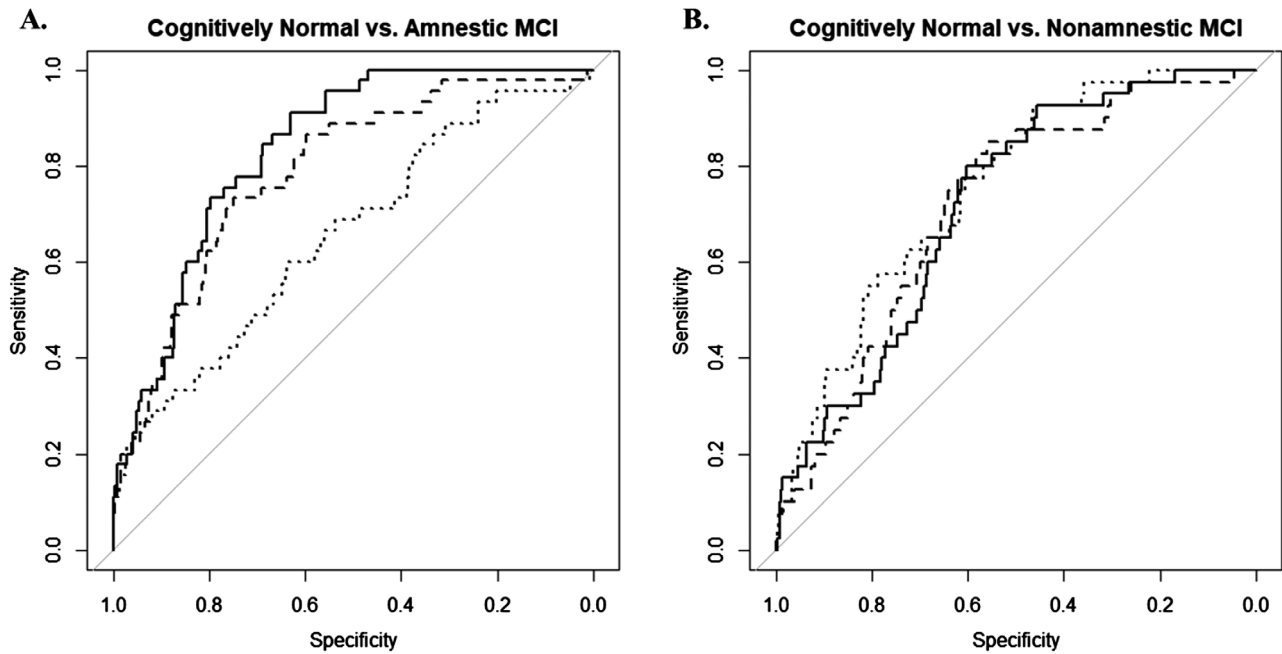


FIGURE 2 Receiver operating characteristic (ROC) curves for the logistic regressions displayed in Table 3 for mild cognitive impairment (MCI) at wave 2 predicted by memory measures at wave 1. Three models are displayed here: the worst individual predictor of amnestic MCI (visual reproductions [VR] immediate recall; dotted line), the best individual predictor of amnestic MCI (California Verbal Learning Test [CVLT] learning trials; dashed line), and the best overall predictor of amnestic MCI (latent factor score of all measures; black line). Area under the curve (AUC) estimates were obtained by doing fourfold cross-validation repeated 10 times. Delong tests revealed significant AUC differences for amnestic MCI between VR immediate recall (AUC = .570) and both CVLT learning trials (AUC = .741) and the full latent factor score (AUC = .796), both $Z > 2.65$, $p < 0.008$, but not between CVLT-learning trials and the latent factor score, $Z = 1.81$, $p = 0.070$. There were no significant AUC differences for non-amnestic MCI, Z 's < 1.36 , $p > 0.175$ (AUCs for VR immediate recall = .686, CVLT learning trials = .607, latent factor score = .637)

ORs (average increase of 105.9% range 46.0%–296.9%). Aggregating across all seven measures resulted in the largest increase in ORs compared to individual measures (average increase of 213.1%, range 85.6%–517.2%). Finally, ORs for latent factors scores were always larger than z-score composites (average increase of 8.7%, range 5.1%–12.0%).

A different pattern of results was observed for non-amnestic MCI. Some memory measures at wave 1 predicted progression to non-amnestic MCI at wave 2 (three of seven measures), the strongest of which was VR immediate recall (OR = 2.02, 95% CI 1.33–3.28). However, the predictive ability was weaker than it was for amnestic MCI. Although aggregating measures sometimes resulted in larger ORs than those for individual measures, none of these estimates was stronger than that for VR immediate recall alone. Thus, there was some evidence that baseline memory predicted later non-amnestic impairment, but combining a mix of informative and non-informative predictors did not appear to strengthen prediction.

3 | DISCUSSION

Memory performance among CN adults who were only in their 50s predicted 6-year progression to amnestic MCI. The results confirmed our primary hypothesis that combining multiple measures of memory improved prediction of progression to amnestic MCI. Although it

will be important to cross-validate these findings in an independent sample, these results have strong implications for prospective studies aimed at early identification of individuals at greatest risk for MCI and AD. They suggest that longitudinal prediction of amnestic MCI can be improved substantially by both administering multiple memory tests and utilizing multiple scores that are available within each test. The use of a single best measure may appear to be time- and cost-effective. However, in this study, administering neuropsychological tests of both verbal and non-verbal memory added only about 30 minutes of additional test administration time and resulted in a dramatic increase in the OR. Even when more than one test has been used, many studies seldom include more than one score from each test. Calculating additional scores from already completed tests adds nothing to participant burden, but does increase predictive power.

Predictors of 3-year progression to MCI or dementia in CN adults in the Australian Imaging, Biomarker & Lifestyle Study of Ageing (AIBL) serve as a sample comparison.⁴² In AIBL (mean age = 72.0 years), Amyloid β ($A\beta$) positivity was the best individual predictor (OR = 4.8), similar to our CVLT trials 1–5 (OR = 4.75). AIBL examined CVLT delayed recall (OR = 4.2), which was similar to that demonstrated in VETSA here (OR = 3.77). The best prediction in AIBL came from combining $A\beta$ and CVLT delayed (OR = 15.9) compared with the prediction based on all memory measures in VETSA (OR = 9.88). Thirteen percent of AIBL progressed to MCI or dementia, whereas 5% of VETSA participants progressed to amnestic MCI. However, the memory tests in VETSA are

accounting for substantially earlier prediction because the average age was 16 years younger than in AIBL.

Although measures of $A\beta$ may have increased prediction of progression to MCI in VETSA, we found that cognitive performance alone has good predictive utility. Obtaining biomarkers of amyloid accumulation in a sample with such a low number of individuals who are expected to be $A\beta$ positive⁴³ would be highly cost-ineffective. Non-invasive neuropsychological tests are cost-effective and sensitive measures for very early risk identification. Using cognitive testing as a first-line screening in clinical trials to reduce unnecessary numbers of positron emission tomography (PET) scans and cerebrospinal fluid (CSF) draws could save millions of dollars and reduce participant burden associated with obtaining biomarker measures of amyloid.⁴⁴ Moreover, being $A\beta$ negative at this age does not necessarily mean non-AD-related, as there is also evidence that $A\beta$ at subthreshold levels can still have negative effects on cognition.¹³ In any case, it will be important for future work on younger samples to evaluate the joint prediction to see if AD biomarkers improve prediction over neuropsychological tests alone in participants as young as those in VETSA.

Including tests in other cognitive domains might also further improve prediction. Some examples include the preclinical Alzheimer cognitive composite (PACC), which used four tests common to three samples,⁴⁵ and multiple tests in the Einstein aging study⁴⁶ and the rush memory and aging project.⁴⁷ However, these indices were created by and for older samples, and were based on a limited number of tests. As in the large majority of studies, their focus was on predicting progression to AD rather than MCI. Baseline ages of participants for these analyses were 16 to 24 years older than the VETSA baseline age. VETSA was designed to have an extensive and taxing battery in order to capture heterogeneity and to avoid ceiling effects in our much younger participants.^{27,48} We see our approach as complementary. Instead of including one score per test as in these other studies, we examine more extensive coverage within specific domains.

In contrast, combining multiple memory predictors did not increase overall prediction for progression to non-amnesic MCI (though all 95% CIs overlapped). The results argue against general cognitive deficit as a predictor of MCI. Rather, there appears to be some specificity of cognitive predictors of MCI. Thus, combining multiple cognitive measures appears to be capable of dramatic improvement in prediction if they are domain-relevant.

3.1 | Strengths and limitations

First, amnesic MCI diagnoses were not validated with AD biomarkers, and we do not know how many MCI cases will develop AD. On the other hand, these diagnoses at wave 1 were validated with evidence of reduced hippocampal volume³⁷ and higher AD polygenic risk scores in individuals with amnesic MCI,³⁸ the latter supporting their being AD-related. Moreover, the 9% of individuals excluded for MCI at wave 1 and 10% converting to MCI at wave 2 correspond well to recent estimates that about 10% of individuals in their mid-50s are amyloid positive, with another 10% percent becoming amyloid positive

by age 65.⁴³ We do not expect all of our MCI subjects to be exactly the same as individuals who are amyloid-positive, but this correspondence at least suggests the plausibility of this being AD-related MCI. Some individuals may also be at subthreshold amyloid levels, and there is growing evidence that subthreshold levels may still be associated with reduced cognitive function.¹³ Second, the sample comprised only men, so it will be important to examine whether these findings generalize to women. Third, many CIs overlapped, but a power analysis suggested that a latent factor based on many test scores would substantially reduce the number of participants needed for studies. Given the proportion of amnesic MCI in this sample (5.3%), power estimates using Gpower V3.1.9.2 (without covariates) suggest that a study that administered only LM delayed recall (OR = 3.22) would require 102 subjects to significantly predict progression to amnesic MCI at the .05 level (with 80% power). With the latent factor based on all seven memory scores (OR = 9.88), the same power analysis indicates that only 35 subjects would be necessary. Thus, we conclude that the small amount of additional non-invasive time in adding two memory tests is a very worthwhile investment.

Fourth, CN individuals closest to the MCI cutoff being at greatest risk for later MCI might appear to suggest that some prediction could stem from test/retest noise. However, the fact that z-score and factor score approaches (which reduce measurement error) improved prediction argues against this point. Post hoc analyses revealed significant prediction by the factor score (OR = 3.42) even when the amnesic MCI group was reduced to only 22 (of 45) individuals who also declined by >1 SD on the full memory factor score between waves 1 and 2. Fifth, as is common for subjects who do not return, the 242 dropouts had significantly lower memory factor scores at wave 1 ($P = 0.009$) than the returnees. Thus, we may have lost some individuals who were at the greatest risk for later memory impairment, but that would suggest that our findings regarding predictive ability are conservative. Finally, fitting the latent factor model and conducting the logistic regression simultaneously might lead to further improvement of prediction. Indeed, prediction of amnesic MCI by latent factors in the same model resulted in 37% to 57% larger ORs than the factor scores displayed in Table 3 (as high as 15.54; see Table S2). However, we presented the results of the two-step procedure here (exporting factor scores and then running a logistic regression) because they are more conservative and are more generalizable in that they can be used by other researchers and/or clinicians to generate risk probabilities in new samples without having to refit a new latent variable model.

3.2 | Concluding remarks

Baseline memory measures can be very useful for early identification as they strongly predict progression to amnesic MCI in middle-age adults. All individuals were CN at baseline, but individual differences still effectively predicted MCI 6 years later. Prospective studies of MCI designed to identify those at greatest risk should administer multiple memory tests and utilize multiple scores from each test as early as possible to maximize their ability to predict change. The Alzheimer's

Association has projected that diagnosing individuals in the MCI stage, as opposed to the dementia stage, could improve quality of life and massively reduce the financial impact of the disease.⁴⁹ Thus, even moderate additional gain in prediction by using additional tests could save large amounts of money when applied to large populations. It is also worth examining whether and when tests in other cognitive domains might further improve prediction. Nevertheless, the results indicate that neuropsychological assessment can be a sensitive predictor of risk for MCI even at an age when few individuals are likely to have become biomarker positive.

ACKNOWLEDGMENTS

The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging/National Institutes of Health (NIA/NIH), or the Veterans Administration. The U.S. Department of Veterans Affairs has provided financial support for the development and maintenance of the Vietnam-Era Twin (VET) Registry. Numerous organizations have provided invaluable assistance in the conduct of the VET Registry, including: Department of Defense; National Personnel Records Center, National Archives and Records Administration; Internal Revenue Service; National Opinion Research Center; National Research Council, National Academy of Sciences; the Institute for Survey Research, and Temple University. This material was, in part, the result of work supported with resources of the VA San Diego Healthcare Center, Center of Excellence for Stress and Mental Health (CESAMH). Most importantly, the authors gratefully acknowledge the continued cooperation and participation of the members of the VET Registry and their families as well as the contributions of many staff members and students.

REFERENCES

- Bateman R, Xiong C, Benzinger TL, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N Engl J Med*. 2012;367:795-804.
- Bennett DA, Schneider JA, Arvanitakis Z, et al. Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology*. 2006;66:1837-1844.
- Aizenstein HJ, Nebes RD, Saxton JA, et al. Frequent amyloid deposition without significant cognitive impairment among the elderly. *Arch Neurol*. 2008;65:1509-1517.
- Kremen WS, Jak AJ, Panizzon MS, et al. Early identification and heritability of mild cognitive impairment. *Int J Epidemiol*. 2014;43:600-610.
- Jack CR, Jr, Bennett DA, Blennow K, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14:535-562.
- Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:280-292.
- Sperling RA, Mormino E, Johnson K. The evolution of preclinical Alzheimer's disease: implications for prevention trials. *Neuron*. 2014;84:608-622.
- Wang G, Coble D, McDade EM, et al. Staging biomarkers in preclinical autosomal dominant Alzheimer's disease by estimated years to symptom onset. *Alzheimers Dement*. 2019.
- Vos SJB, Duara R. The prognostic value of ATN Alzheimer biomarker profiles in cognitively normal individuals. *Neurology*. 2019;94:643-644.
- Edmonds EC, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW. Alzheimer's disease neuroimaging I. Subtle cognitive decline and biomarker staging in preclinical Alzheimer's disease. *J Alzheimers Dis*. 2015;47:231-242.
- Ewers M, Walsh C, Trojanowski JQ, et al. Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiol Aging*. 2012;33:1203-1214.
- Gomar JJ, Conejero-Goldberg C, Davies P, Goldberg TE. Alzheimer's disease neuroimaging I. Extension and refinement of the predictive value of different classes of markers in ADNI: four-year follow-up data. *Alzheimers Dement*. 2014;10:704-712.
- Insel PS, Mattsson N, Mackin RS, et al. Accelerating rates of cognitive decline and imaging markers associated with beta-amyloid pathology. *Neurology*. 2016;86:1887-1896.
- Jedynak BM, Lang A, Liu B, et al. A computational neurodegenerative disease progression score: method and results with the Alzheimer's Disease Neuroimaging Initiative cohort. *Neuroimage*. 2012;63:1478-1486.
- Tabert MH, Manly JJ, Liu X, et al. Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Arch Gen Psychiatry*. 2006;63:916-924.
- Bondi MW, Edmonds EC, Jak AJ, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *J Alzheimers Dis*. 2014;42:275-289.
- De Jager CA, Hogervorst E, Combrinck M, Budge MM. Sensitivity and specificity of neuropsychological tests for mild cognitive impairment, vascular cognitive impairment and Alzheimer's disease. *Psychol Med*. 2003;33:1039-1050.
- Lange KL, Bondi MW, Salmon DP, et al. Decline in verbal memory during preclinical Alzheimer's disease: examination of the effect of APOE genotype. *J Int Neuropsychol Soc*. 2002;8:943-955.
- Kremen WS, Panizzon MS, Franz CE, et al. Genetic complexity of episodic memory: a twin approach to studies of aging. *Psychol Aging*. 2014;29:404-417.
- Panizzon MS, Neale MC, Docherty AR, et al. Genetic and environmental architecture of changes in episodic memory from middle to late middle age. *Psychol Aging*. 2015;30:286-300.
- Golde TE, Schneider LS, Koo EH. Anti- β therapeutics in Alzheimer's disease: the need for a paradigm shift. *Neuron*. 2011;69:203-213.
- Gustavson DE, Panizzon MS, Elman JA, et al. Stability of genetic and environmental influences on executive functions in midlife. *Psychol Aging*. 2018;33:219-231.
- Salthouse TA. When does age-related cognitive decline begin. *Neurobiol Aging*. 2009;30:507-514.
- Salthouse TA. Are individual differences in rates of aging greater at older ages. *Neurobiol Aging*. 2012;33:2373-2381.
- Tsuang MT, Bar JL, Harley RM, Lyons MJ. The Harvard twin study of substance abuse: what we have learned. *Harv Rev Psychiatry*. 2001;9:267-279.
- Kremen WS, Panizzon MS, Xian H, et al. Genetic architecture of context processing in late middle age: more than one underlying mechanism. *Psychol Aging*. 2011;26:852-863.
- Kremen WS, Thompson-Brenner H, Leung YM, et al. Genes, environment, and time: the Vietnam era twin study of aging (VETS). *Twin Res Hum Genet*. 2006;9:1009-1022.
- Schoenborn CA, Heyman KM. Health characteristics of adults aged 55 years and over: United States, 2004-2007. *Natl Health Stat Report*. 2009;16:1-31.
- Wechsler D. Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corp.; 1955.

30. Wechsler D. Wechsler Memory Scale (WMS-III). San Antonio, TX: Psychological Corporation; 1997.
31. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test (CVLT-2). San Antonio, TX: Psychological Corporation; 2000. 2nd ed.
32. Muthén LK, Muthén BO. Mplus Version 7 User's Guide. Los Angeles, CA: Muthén & Muthén; 2010-2012.
33. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6:1-55.
34. Jak AJ, Bondi MW, Delano-Wood L, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry*. 2009;17:368-375.
35. Lyons MJ, Panizzon MS, Liu W, et al. A longitudinal twin study of general cognitive ability over four decades. *Dev Psychol*. 2017;53:1170-1177.
36. Lyons MJ, York TP, Franz CE, et al. Genes determine stability and the environment determines change in cognitive ability during 35 years of adulthood. *Psychol Sci*. 2009;20:1146-1152.
37. Jak AJ, Panizzon MS, Spoon KM, et al. Hippocampal atrophy varies by neuropsychologically defined MCI among men in their 50s. *Am J Geriatr Psychiatry*. 2015;23:456-465.
38. Logue MW, Panizzon MS, Elman JA, et al. Use of an Alzheimer's disease polygenic risk score to identify mild cognitive impairment in adults in their 50s. *Mol Psychiatry*. 2019;24:421-430.
39. Elman JA, Jak AJ, Panizzon MS, et al. Underdiagnosis of mild cognitive impairment: a consequence of ignoring practice effects. *Alzheimers Dement*. 2018;10:372-381.
40. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67.
41. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977;1:385-401.
42. Rowe CC, Bourgeat P, Ellis KA, et al. Predicting Alzheimer disease with beta-amyloid imaging: results from the Australian imaging, biomarkers, and lifestyle study of ageing. *Ann Neurol*. 2013;74:905-913.
43. Jansen WJ, Ossenkoppele R, Knol DL, et al. Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *Jama-J Am Med Assoc*. 2015;313:1924-1938.
44. Insel PS, Palmqvist S, Mackin RS, et al. Assessing risk for preclinical beta-amyloid pathology with APOE, cognitive, and demographic information. *Alzheimers Dement*. 2016;4:76-84.
45. Donohue MC, Sperling RA, Salmon DP, et al. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol*. 2014;71:961-970.
46. Zammit AR, Hall CB, Lipton RB, Katz MJ, Muniz-Terrera G. Identification of heterogeneous cognitive subgroups in community-dwelling older adults: a latent class analysis of the Einstein aging study. *J Int Neuropsychol Soc*. 2018;24:511-523.
47. Zammit AR, Muniz-Terrera G, Katz MJ, et al. Subtypes based on neuropsychological performance predict incident dementia: findings from the rush memory and aging project. *J Alzheimers Dis*. 2019;67:125-135.
48. Kremen WS, Franz CE, Lyons MJ. VETSA: the Vietnam era twin study of aging. *Twin Res Hum Genet*. 2013;16:399-402.
49. Association As. Alzheimer's disease facts and figures. *Alzheimers Dement*. 2018;14:367-429.
50. Lyall DM, Celis-Morales C, Lyall LM, et al. Assessing for interaction between APOE epsilon4, sex, and lifestyle on cognitive abilities. *Neurology*. 2019;92:e2691-e8.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Gustavson DE, Elman JA, Sanderson-Cimino M, et al. Extensive memory testing improves prediction of progression to MCI in late middle age. *Alzheimer's Dement*. 2020;12:e12004. <https://doi.org/10.1002/dad2.12004>