Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 6 (2017) 171-181

Alzheimer's & Dementia

Cognitive & Behavioral Assessment

A comparison of theoretical and statistically derived indices for predicting cognitive decline

Holly Wilhalme[a,1], Naira Goukasian[b,1], Fransia De Leon[b], Angie He[b], Kristy S. Hwang[c], Ellen Woo[b], David Elashoff[a], Yan Zhou[b], John M. Ringman[d], Liana G. Apostolova[b,e,f,g,*]

[a]*Department of Medicine Statistics Core, University of California, Los Angeles, Los Angeles, CA, USA*
[b]*Department of Neurology, University of California, Los Angeles, Los Angeles, CA, USA*
[c]*Oakland University, William Beaumont School of Medicine, Rochester, MI, USA*
[d]*Department of Neurology, University of Southern California, Los Angeles, CA, USA*
[e]*Department of Neurology, Indiana University, Indianapolis, IN, USA*
[f]*Department of Radiology, Indiana University, Indianapolis, IN, USA*
[g]*Department of Medical and Molecular Genetics, Indiana University, Indianapolis, IN, USA*

Abstract

Introduction: Both theoretical and statistically derived approaches have been used in research settings for predicting cognitive decline.

Methods: Fifty-eight cognitively normal (NC) and 71 mild cognitive impairment (MCI) subjects completed a comprehensive cognitive battery for up to 5 years of follow-up. Composite indices of cognitive function were derived using a classic theoretical approach and exploratory factor analysis (EFA). Cognitive variables comprising each factor were averaged to form the EFA composite indices. Logistic regression was used to investigate whether these cognitive composites can reliably predict cognitive outcomes.

Results: Neither method predicted decline in NC. The theoretical memory, executive, attention, and language composites and the EFA-derived "attention/executive" and "verbal memory" composites were significant predictors of decline in MCI. The best models achieved an area under the curve of 0.94 in MCI.

Conclusions: The theoretical and the statistically derived cognitive composite approaches are useful in predicting decline in MCI but not in NC.

Keywords: Alzheimer's disease; AD; Factor analysis; Neuropsychological test; Mild cognitive impairment; MCI; Cognitive decline; Memory decline; Executive decline; Conversion; Progression; Prognosis

1. Introduction

Mild cognitive impairment (MCI) is a risk state for dementia [1]. Patients with MCI invariably manifest more cognitive difficulties than one might expect given their age but can still live independently [1], thus failing to meet the diagnostic criteria for dementia [2,3]. Although an estimated 15% of people living with MCI progress to dementia each year [4], some remain stable or even revert to exhibit normal cognition [5]. Early identification of cognitively normal (NC) individuals who will convert to MCI and MCI individuals who will convert to dementia with confidence and high sensitivity will provide the opportunity to intervene at early stages and have greater potential for modifying the disease course.

Neuropsychological (NP) testing is an essential tool for assessing cognitive function in both the prodromal and

[1]Denotes co-first authorship.

*Corresponding author. Tel.: +1 317 963 7173; Fax: +1 317 963 7325.
E-mail address: lapostol@iu.edu

http://dx.doi.org/10.1016/j.dadm.2016.10.002

dementia stages [6–8]. It has been previously shown that NP testing can capture areas of cognitive decline relatively early in the disease course [9–11]. NP testing can effectively capture areas of cognitive decline before the observation of any clinical symptoms [12]. NP evaluations obtained in the prodromal stages predicted Alzheimer's disease (AD) pathology with 89% accuracy, as later confirmed by autopsy [13].

Comprehensive NP batteries are often comprised of different measures that tap into various cognitive domains. The common practice is to derive composite scores (i.e., domains) with a normalized distribution that can be more easily compared with one another [14]. Traditionally, the tests that comprise these domains have been chosen on a theoretical basis. Individual test results are grouped and interpreted in theoretically derived cognitive domains, such as attention, language, memory, visuospatial, and executive functions. For example, measures that require recall are often included in the memory domain and measures that require concentration are often included in the attention domain.

Theoretically derived cognitive domains can be used to predict future cognitive decline [14–17]. The memory domain has been observed to discriminate best between NC and MCI, and to be most predictive of progression to AD [8,12,18,19]. Free recall, recognition memory, and paired-associate learning have been shown to be impaired in patients with AD and MCI [8,12,20]. One study showed that memory assessment predicted diagnosis of normal cognition and dementia with 94.5% and 66.7% accuracy, respectively, in a large cohort of older adults [21]. Welsh et al. [22] demonstrated that the amount of information recalled after a 10-minute delay on the Consortium to Establish a Registry for AD [23] differentiated MCI from healthy normal controls with >90% accuracy. These studies demonstrate that NP assessments that track deficits in the memory domain help detect symptom markers essential for early diagnosis.

Executive functions in everyday life, such as planning a vacation or creating a grocery list, are often affected early in the disease course. Such decline is reflected in executive deficits in NP testing [24,25]. Performance in the executive domain is highly predictive of future conversion from MCI to AD [14,15,26]. The executive domain includes tests that require planning and organization. The decision whether to include measures into the theoretical executive domain as opposed to the attention domain is driven by the additional need to formulate a plan of goal-directed action or to inhibit an overlearned response. Notably, some executive tests require problem-solving skills rather than just processing speed and concentration [24,27]. The Wisconsin Card Sorting Test [28], a novel problem solving measure, and the Stroop tests [29], a measurement of response inhibition, are therefore typically categorized as executive functioning tasks and aid in the prediction of future cognitive decline [30].

NP test groupings can also be derived using statistical methods [6,7,14,31–35]. Such methods have been shown to have high accuracy, specificity, and sensitivity in diagnosing dementia [6,31,33]. By combining individual test scores into composite scores using multivariate techniques, diagnoses are less prone to errors caused by chance and the minimum number of tests needed to detect cognitive decline decreases [14]. Exploratory factor analysis (EFA) is a widely used technique that is commonly used to identify the internal relationships among variables and to create interpretable composite scores from large sets of variables. When compared with the theoretical model, where each of the constituent parts are analyzed separately, composite scores in EFA may be more powerful, in that they may increase measurement precision, help avoid specific characteristics of a particular test that may be influenced by chance, and limit the number of statistical tests needed to derive a conclusion [14]. This method groups specific tests within a domain when these tests are highly intercorrelated (i.e., have high covariance), ensuring that they tap the same cognitive constructs. The decision of how to group tests is thus based on the data rather than on theory [36]. When NP tests are run through EFA and the resulting factor models are produced, incorrectly classified models that demonstrate a negative error variance can also be revised and compared with better models [33]. This is important in analyzing and confirming the best model comprising the most accurate factors. Previous studies of EFA with NP tests have yielded different groupings of tests than the theoretical model [6,33,37]. Yet, EFA has been shown to accurately predict cognitive decline in older adults and to improve diagnostic classification and predict cognitive decline [6,31,33,37].

Two large studies, one using data from the National Alzheimer's Coordinating Center (N = 12,020) [6] and the other using data from the Alzheimer's Disease Neuroimaging Initiative (N = 819) [7], used EFA to develop models for predicting cognitive changes over time and differences within diagnostic groups. Hayden et al. [6] derived a four-factor solution including memory, executive function, language, and attention composites, whereas Park et al. [7] also included a fifth visuospatial factor. Another EFA study of 1288 middle-aged adults with and without family risk factors for AD derived a different five-factor model comprising verbal ability, visuospatial, speed and executive function, working memory, and verbal and memory factors that was able to explain 63% of the cognitive variance. These factors were invariant across groups defined by age, gender, family history of AD, and *APOE* ε4 carrier status [33]. Taken together, these data-driven analytic studies suggest that cognition is similarly organized across the geriatric cognitive spectrum and that factor scores resulting from these cognitive domains can be used stably across all groups. The aforementioned studies also suggest that EFA can effectively differentiate unique aspects of samples that are relevant, given that different factors emerged across different cohorts. These

group differences would not be inherently considered as they are when using EFA if only the theoretical method had been used. For example, although executive functioning tasks are generally "problem solving" tasks, they also feature tasks that exhibit aspects of attention, verbal reasoning, visuospatial skills, and more [27]. Many tasks, such as Stroop C [29], Trailmaking B [38], and Rey Osterrieth Complex Figure Copy [39], are often grouped into the executive domain, but can also be categorized in different theoretical domains [24,40,41]. This can sometimes lead to inaccuracy in comparing scores stably across all groups because it is difficult to validate the grouping if a test can be theoretically assigned to two different domains [41].

Applying principal component analysis (PCA), a similar variable reduction technique, to a data set of 43 MCI subjects with longitudinal follow-up, Chapman et al. [31] reduced 49 NP measures into six-component scores. Five of these components—comprising various memory test subcomponents (episodic memory, recognition memory, visuospatial memory, and visuospatial episodic memory) and two executive measures (speeded executive functioning and processing speed)—achieved 86% sensitivity and 83% specificity for predicting conversion from MCI to dementia, despite the study's small sample size [31].

There are various ways to conduct an EFA depending on the goals of the researcher and underlying assumptions of your data set, but there are few absolute guidelines [42]. Although the most common statistical data reduction technique used in the scientific community is PCA with varimax rotation, it has been argued that this technique is not optimal given the intricacies in "real-world" social sciences data and the fact that PCA is only a data reduction technique, which is computed without regard to the structure of the underlying latent variables and therefore can produce inflated values of variance accounted for by the components [42]. Although EFA may be the preferred method of data reduction, there are many decisions, based on the assumptions one makes about the data, which must be made along the way that can influence the resulting factors. First, a factor extraction method must be selected such as maximum likelihood or principal axis factoring, depending on the distribution of the data. Then it must be decided how many factors to retain before rotation. This is most commonly done using the Kaiser criterion (i.e., retaining all factors with eigenvalues >1.0) and the scree test, which can often produce conflicting results. Then the results must be rotated to simplify and create interpretable results from the factors. The researcher must decide whether to use an orthogonal rotation, which generates independent, uncorrelated factors or an oblique rotation, which allows for correlation among factors. Once the factor analysis has been conducted, the researcher may then use the factor loadings to produce factor scores or, in certain cases where identification of the basic subdimensions of the data is the primary concern, the researcher may opt to use the factor loadings to guide the creation of the composites and choose an alternative method to compute the scores (e.g., by averaging the variables, which load onto each factor). Each method has its advantages and disadvantages and depending on how different the resulting factors are from each method, the simplest and most interpretable method may be preferred. In this study, we explore the use of three different methods of EFA to create composite scores from our battery of NP tests and compare them based on their ability to predict cognitive decline from NC to MCI and MCI to dementia.

So far the literature suggests that both the theoretical and experimental variable reduction techniques are useful and predictive of clinical diagnosis. However, which of these two approaches is more powerful in predicting cognitive outcomes should be determined in the same sample. The goal of our study was to compare the ability of theoretical and EFA-based groupings of NP tests for predicting cognitive decline among NC and MCI subjects in an effort to improve early diagnosis and to determine whether the statistical multivariate technique more efficiently and effectively predicts cognitive decline compared with the theoretical model. As factor analysis helps eliminate tests and factors that do not necessarily improve the accuracy of diagnosis, we hypothesized that EFA will achieve better classification accuracy in predicting cognitive decline than the theoretically derived cognitive domain approach.

2. Methods

2.1. Study sample

The Imaging and Genetic Biomarkers for AD study (ImaGene) is a prospective longitudinal study that collects annual comprehensive clinical and NP data for up to 5 years from NC and MCI subjects. A total of 160 subjects were initially enrolled from two sources: (1) the ongoing longitudinal study from University of California, Los Angeles (UCLA) Alzheimer's Disease Research Center (ADRC) and (2) referring neurologists from within and outside UCLA. Initial inclusion criteria required that participants were aged 50 years or older with a Mini-Mental State Examination (MMSE) score $\geq$24 without other neurological or major psychiatric diseases and no recent history of drug or alcohol abuse. All eligible subjects were able to live independently and were free of visual or hearing impairments that could interfere with cognitive testing. Research subjects with significant white matter changes, large vessel strokes, or strategically placed lacunes that may result in cognitive decline were excluded from participation. Informed consent was obtained according to the restrictions and policies of the UCLA institutional review board.

Initial and longitudinal diagnoses were determined by consensus decision by a group of neurologists and

Table 1
Demographic characteristics for ImaGene normal controls and UCLA ADRC normal controls

Variable	ImaGene normal controls	UCLA ADRC normal controls	*P* value
Sample size	58	62	
Age (y)	69 (8)	69 (6)	NS
Education (y)	18 (2)	17 (2)	NS
Sex (M/F)			
Male	32 (55%)	31 (50%)	NS
Female	26 (45%)	31 (50%)	
MMSE	29 (1)	29 (1)	NS
GDS	1.43 (1.96)	1.37 (2.21)	NS

Abbreviations: ADRC, Alzheimer's Disease Research Center; GDS, Geriatric Depression Scale; MMSE, Mini-Mental State Examination; NS, not significant; UCLA, University of California, Los Angeles.

NOTE. Continuous variables are represented as the mean (standard deviation).

neuropsychologists using the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders criteria for AD [9] and Petersen criteria for MCI [16]. Diagnosis of MCI required a score of 1.5 standard deviations less than the age- and/or education-adjusted norms on at least one NP test and intact functional abilities corroborated by the informant. Diagnosis of NC required a score >1.5 standard deviations compared with age- and/or education-adjusted norms. The presence of depression was assessed using the Geriatric Depression Scale (GDS) [43].

For the purposes of our study, subjects were categorized into four groups: NC subjects who remained stable (NCstable), NC subjects who progressed to MCI (NCdecliner), MCI subjects who remained stable (MCIstable), and MCI subjects who progressed to dementia (MCIdecliner). Of the 160 enrolled subjects, 12 discontinued and one passed away before the first follow-up. An additional 18 reverted back from MCI to NC during follow-up and were excluded from our analyses. Our final sample consisted of 129 subjects (58 NC and 71 MCI). Length of follow-up varied (Table 1) because of unforeseen circumstances, such as health, geographic location, and other factors that led to the loss of follow-up. Two MCI subjects developed parkinsonism, hallucinations, and rapid eye movement behavior disorder in follow-up. These subjects were diagnosed with dementia with Lewy bodies based on the dementia with Lewy bodies Consortium criteria [44]. One of these subjects passed away and his autopsy results were consistent with hippocampal sclerosis and early stages of AD pathology. These two subjects were included in the MCIdecliner group. The rest of the subjects in the MCIdecliner group were diagnosed with AD dementia.

2.1.1. Normative sample

A sample of 62 NC subjects enrolled in the longitudinal study at the UCLA ADRC who were not part of the ImaGene study sample described previously and who received the same NP test battery served as a normative sample. Their cognitive data were used to estimate the relative contribution of age for each individual test for age-adjustment of the ImaGene cognitive data.

2.2. NP assessment

Subjects were tested annually with an NP battery consisting of MMSE [45], Wechsler Test of Adult Reading [46], Wechsler Adult Intelligence Scale, third edition [47] subtests, Wechsler Memory Scale—third edition [48] subtests, Trailmaking A and B [38], Stroop Color-Word Interference Test [29], Boston Naming Test [49], Controlled Oral Word Association Test [50], Rey Osterrieth Complex Figure Copy and 3-minute recall [39], California Verbal Learning Test—second edition [51], and Wisconsin Card Sorting Test [28].

2.3. Statistical methods

The goal of this analysis was to predict cognitive decline from NC or MCI using composite scores derived from the battery of cognitive tests administered at baseline. Each cognitive test score was standardized for age to an independent sample of NC subjects (N = 62) from the UCLA ADRC longitudinal research cohort (Table 1) using linear regression. For each cognitive test, linear regression was conducted with age as the predictor in both the ADRC cohort and the normative. The standardized score was then calculated by subtracting the predicted score in the ADRC cohort (as estimated by the linear regression) from the original score and dividing by the standard deviation of the residuals from that model. Then the regression model from the normative controls was applied to the score resulting in the age-standardized score. Then, all NP measures were converted to standard *z*-scores before being combined into composites. For most cognitive tests, a higher score indicates better performance, however, for some of the timed tests (i.e., Stroop Color-Word Interference, Trailmaking A and B), a higher score indicates poorer performance; therefore, scores for these tests were multiplied by −1. Standardization was performed before entering the data into both theoretically and EFA-derived composite indices.

2.4. Data reduction techniques

Two theoretical and three statistical methods were used in the creation of the composite indices. The former consisted of (1) averaging the standardized scores for all tests in each of five theoretically driven domains and (2) using the minimum standardized score of all the tests included in that theoretical domain for each subject. For the latter, we compared three methods using EFA – varimax rotation (average), varimax rotation (factor scores), and oblique rotation (factor scores). For all methods, we used factor analysis with squared multiple correlations on the diagonal for the correlation matrix. The number of factors retained was determined using the Kaiser criterion (eigenvalue >1) and examination of the scree plot. Interpretability of the factors was also considered,

Table 2
Demographic characteristics at baseline visit

Variable	NCstable, N = 48	NCdecliner, N = 10	*P* value	MCIstable, N = 41	MCIdecliner, N = 30	*P* value
Age (y)	69 (8)	72 (10)	NS	69 (9)	74 (8)	**.01**
Education (y)	17 (2)	19 (2)	**.03**	15 (3)	16 (3)	NS
Sex (M/F)						
Male	25 (52%)	7 (70%)	NS	17 (41%)	12 (40%)	NS
Female	23 (48%)	3 (30%)		24 (59%)	18 (60%)	
Race						
White	45 (94%)	9 (90%)	NS	40 (98%)	28 (93%)	NS
African-American	1 (2%)	1 (10%)		0 (0%)	1 (3%)	
Asian	1 (2%)	0 (0%)		1 (2%)	0 (0%)	
Unknown	1 (2%)	0 (0%)		0 (0%)	1 (3%)	
Length of follow-up (y)	4.4 (1.3)	3.9 (1.8)	NS	3.0 (1.2)	3.2 (1.2)	NS
MMSE	29 (1.1)	28 (1.9)	**.04**	28 (1.6)	26 (3.1)	**.0002**
GDS	0.8 (1.1)	2.1 (2.4)	**.009**	1.7 (1.6)	2.3 (2.3)	NS
Antidementia medications	0 (0%)	1 (10%)	NS	5 (12%)	8 (27%)	NS
Started AD medications during follow-up period	3 (6%)	1 (10%)	NS	5 (12%)	24 (80%)	**<.0001**

Abbreviations: AD, Alzheimer's disease; GDS, Geriatric Depression Scale; MCI, mild cognitive impairment; MCIdecliner, MCI subjects who progressed to dementia; MCIstable, MCI subjects who remained stable; MMSE, Mini-Mental State Examination; NC, cognitively normal; NCdecliner, NC subjects who progressed to MCI; NCstable, NC subjects who remained stable; NS, not significant.

NOTE. Continuous variables are represented as the mean (standard deviations).

Bold text indicates significance.

in that only factors in which at least two variables had significant loadings (>0.30) were retained. For the first EFA-based method, the factors were rotated using the varimax procedure to produce independent and uncorrelated factors and cognitive tests were retained on the factor on which they loaded highest. Cognitive variables comprising each factor were averaged to form the first set of EFA-derived composites. In the case where a variable had similar high loadings on more than one factor, we selected the factor, which made the most sense theoretically. The second set of EFA-derived composites was calculated using the factor scores from this procedure. For the third EFA-based method, the factors were rotated using the promax procedure, which allows the resulting factors to be correlated.

Logistic regression models were used to predict decliner status for each of the two subject groupings: NCstable versus NCdecliner and MCIstable versus MCIdecliner. For each subject group, a series of models were constructed, including a single cognitive factor. Each model was adjusted for age, sex, education, and length of follow-up. C-statistics were used to determine the classification accuracy of each model. Leave-one-out cross validation was conducted on the MCI cohort and the two receiver operating characteristics curves (ROCs) were compared. Demographics were compared between the cohorts using the Student *t* test for continuous variables and the chi-square test (and Fisher's exact test where appropriate) for categorical variables. SAS version 9.4 was used for all statistical analyses. *P* value <.05 was considered statistically significant.

3. Results

Table 1 shows the demographic characteristics for the ImaGene normal controls and UCLA ADRC normal controls. No significant differences in age, education, sex, MMSE, GDS, or antidementia medications at baseline were seen between the two groups. Among those that were MCI at baseline, 12% of subjects that remained stable and 80% of subjects who declined over time started antidementia medications during the study (*P* <.0001; Table 2).

Table 2 shows the subject demographics at baseline for each of the diagnostic progression groups. There were no significant differences in sex, race, or length of follow-up in years between the groups. Ten of the 58 NC subjects (17%) progressed to MCI (mean follow-up: 3.9 ± 1.8 years). Thirty of the 71 MCI subjects (42%) progressed to dementia (mean follow-up: 3.2 ± 1.2 years). MCI subjects who remained stable were younger than those who declined (MCIstable vs. MCIdecliner 69 ± 9 vs. 74 ± 8 years; *P* = .01). There was no significant age difference between NC who declined and those who remained stable. MMSE was significantly higher in the NC and MCI that remained stable than those who declined (NCstable vs. NCdecliner: 29 ± 1.1 vs. 28 ± 1.9, *P* = .04; MCIstable vs. MCIdecliner: 28 ± 1.6 vs. 26 ± 3.1, *P* = .0002). Compared with NCstable, NCdecliners were more highly educated (NCstable vs. NCdecliner: 17 ± 2 vs. 19 ± 2, *P* = .03). There were no differences in GDS score in the MCIstable and MCIdecliner groups. The NCdecliner group had a significantly higher GDS score compared with the NCstable group (*P* = .009).

The raw and age-standardized cognitive test scores at baseline for each of the theoretically derived domains per group are presented in Table 3. The factor loadings for the varimax-rotated EFA are shown in Table 4. Factors 1 to 3 had eigenvalues >1 and accounted for 84% of the cumulative variance. The scree plot showing the eigenvalues for each factor is shown in Fig. 1. Factors 1 to 3 had eigenvalues >1.0. The scree plot appears to level

Table 3
Cognitive group means (standard deviations) for each test at baseline

		Raw scores				Age-normalized scores			
Domains	Variables	NCstable	NCdecliner	MCIstable	MCIdecliner	NCstable	NCdecliner	MCIstable	MCIdecliner
Attention	Digit symbol	68 (14)	68 (19)	54 (12)	44 (11)	66 (11)	67 (16)	56 (10)	48 (9)
	Digit span forward	11 (2)	10 (2)	10 (2)	9 (2)	10 (2)	10 (2)	9 (3)	9 (2)
	Digit span backward	8 (2)	8 (2)	6 (2)	5 (1)	8 (2)	8 (1)	6 (2)	6 (1)
	Trails A time	25 (9)	24 (8)	35 (15)	46 (18)	26 (6)	28 (6)	34 (10)	41 (13)
	Stroop Color Naming	61 (10)	70 (16)	69 (18)	90 (18)	59 (7)	67 (12)	63 (12)	77 (12)
	Stroop Word Reading	46 (8)	49 (10)	51 (10)	58 (9)	45 (9)	49 (11)	49 (13)	57 (9)
Language	Boston Naming Test	58 (2)	57 (4)	51 (10)	47 (11)	59 (6)	59 (7)	50 (13)	50 (14)
	Animal fluency	51 (15)	49 (11)	34 (12)	30 (15)	52 (10)	48 (6)	40 (9)	38 (9)
	FAS fluency	22 (5)	22 (5)	18 (5)	14 (4)	24 (4)	23 (5)	21 (4)	18 (4)
Visuospatial	ROCF copy	33 (3)	33 (2)	29 (4)	29 (5)	34 (2)	35 (2)	32 (3)	32 (3)
	Block design	13 (3)	13 (3)	10 (3)	10 (3)	NA	NA	NA	NA
Memory	Logical memory I	45 (9)	39 (8)	31 (11)	15 (8)	50 (5)	48 (3)	43 (7)	35 (6)
	Logical memory II	29 (7)	24 (7)	17 (9)	5 (5)	33 (4)	31 (3)	26 (6)	20 (4)
	Visual reproduction I	86 (11)	78 (12)	68 (17)	52 (16)	89 (9)	84 (10)	78 (12)	68 (11)
	Visual reproduction II	67 (23)	55 (24)	34 (20)	10 (12)	72 (18)	52 (20)	44 (16)	29 (12)
	Trails 1–5 total score	51 (12)	47 (12)	36 (10)	25 (7)	56 (8)	51 (8)	45 (7)	39 (7)
	Short delay free recall	11 (3)	9 (3)	6 (4)	2 (2)	12 (2)	11 (2)	9 (2)	7 (2)
	Long delay free recall	12 (3)	10 (4)	5 (4)	2 (2)	13 (2)	12 (2)	10 (2)	8 (2)
	ROCF 3-min delay	21 (6)	18 (6)	12 (7)	7 (6)	22 (4)	21 (5)	15 (4)	13 (5)
Executive	Wisconsin Card Sorting Test categories	5 (1)	4 (1)	3 (2)	2 (1)	5 (1)	4 (1)	3 (2)	3 (1)
	Trails B time	62 (23)	74 (36)	117 (61)	173 (83)	54 (10)	65 (19)	77 (25)	100 (37)
	Stroop interference	110 (21)	139 (44)	140 (36)	189 (65)	104 (15)	118 (28)	121 (21)	140 (42)

Abbreviations: FAS, F-A-S Test for Verbal Fluency; MCI, mild cognitive impairment; MCIdecliner, MCI subjects who progressed to dementia; MCIstable, MCI subjects who remained stable; NC, cognitively normal; NCdecliner, NC subjects who progressed to MCI; NCstable, NC subjects who remained stable; ROCF, Rey Osterrieth Complex Figure.

off at factor 6; however, only one cognitive test loaded onto factors 4 to 6 and, therefore, these factors were excluded from further analyses. The factor loadings for the promax-rotated EFA are shown in Table 5. The scree plot for the EFA with oblique rotation (also shown in Fig. 1) is very similar to the plot for the orthogonal

Table 4
Factor loadings for the 6-factor EFA solution—varimax rotation

Cognitive variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Stroop Color Naming	**0.82**	−0.37	−0.03	0.11	−0.11	−0.09
Trails A time	**0.76**	−0.27	−0.24	−0.11	0.08	−0.07
Trails B time	**0.73**	−0.29	−0.26	−0.17	−0.13	0.01
Digit symbol	**−0.73**	0.30	0.26	0.17	0.16	−0.10
Stroop Word Reading	**0.70**	−0.24	−0.13	0.26	−0.22	−0.20
Stroop interference	**0.64**	−0.19	−0.05	−0.27	−0.11	−0.14
Verbal fluency—FAS	**−0.51**	0.39	0.15	0.13	0.28	0.20
Verbal fluency—animal	**−0.46**	0.43	0.17	0.12	0.09	−0.01
WCST categories	**−0.46***	**0.46***	0.36	−0.23	0.10	0.02
Short delay free recall	−0.35	**0.88**	0.18	−0.01	0.10	−0.03
Long delay free recall	−0.30	**0.88**	0.20	0.14	0.10	0.00
Trails 1–5 total score	−0.37	**0.79**	0.18	0.09	0.08	0.05
Logical memory I	−0.29	**0.61**	0.19	0.52	0.12	0.10
Logical memory II	−0.38	**0.61**	0.20	0.49	0.14	0.09
ROCF copy	−0.06	0.14	**0.64**	0.14	0.20	−0.17
Visual reproduction II	−0.37	0.55	**0.61**	−0.02	−0.01	0.12
Block design	−0.16	0.09	**0.60**	0.23	0.01	0.13
ROCF 3-min delay	−0.38	0.48	**0.60**	0.10	0.08	0.06
Visual reproduction I	**−0.50***	0.41	**0.48***	−0.15	0.02	0.14
Boston Naming Test	0.00	0.08	0.17	**0.76**	0.05	0.20
Digit span backward	−0.28	0.20	0.21	0.10	**0.78**	0.27
Digit span forward	−0.16	0.02	−0.02	0.35	0.26	**0.73**
Variance explained by each factor	**66%**	**10%**	**8%**	**6%**	**4%**	**0.5%**

Abbreviations: EFA, exploratory factor analysis; ROCF, Rey Osterrieth Complex Figure; WCST, Wisconsin Card Sorting Test.

*Visual Reproduction 1 and WCST categories showed high loadings on two factors. In all further analyses WCST categories were grouped with the rest of the executive measures in factor 1 and visual reproduction 1 with the rest of the visuospatial measures in factor 3.

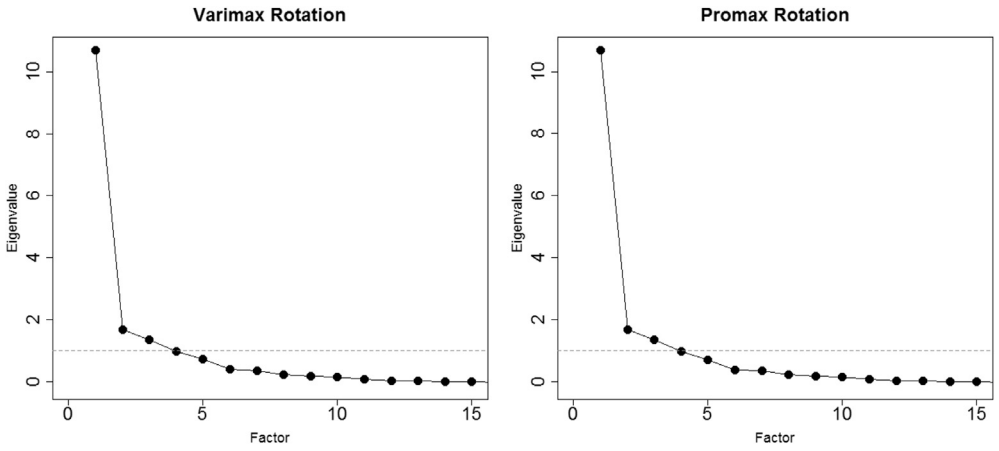

Fig. 1. Scree plot showing the eigenvalues for each factor of the orthogonal rotation (left) and oblique rotation (right).

rotation. On the basis of the break after the factor 5, five factors were retained for further analysis.

The results from the logistic regression analyses are presented in Table 6. Using the average of the theoretically derived domains, we found no significant predictors of decline in the NC group. Within the MCI group, all domains except for the visuospatial were significant predictors of decline to dementia: memory (P = .002, area under the curve [AUC] = 0.94), executive (P = .03, AUC = 0.75), attention (P = .002, AUC = 0.82), and language (P = .03, AUC = 0.76). Using the test with the minimum standardized score as the predictor, we found no significant predictors of decline in the NC group. Within the MCI group, the memory (P = .0003, AUC = 0.87), executive (P = .03, AUC = 0.77), and attention (P = .03, AUC = 0.77) domains were significant predictors of decline to dementia.

The EFA-derived composites produced no significant predictors of decline for subjects who were normal at baseline across all three methods. Using the average of the scores, which loaded highest on each factor, factor 1 (attention/executive) had the best prediction accuracy of 0.77 for the NC group, but did not reach statistical significance. For subjects who were MCI at baseline, factors 1 and 2 (verbal memory and attention/executive, respectively) were significant predictors of decline. Factor 2 (verbal memory) had the

Table 5
Factor loadings for the 6-factor EFA solution—promax rotation

Cognitive test	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Stroop Color Naming	0.86	−0.11	0.21	0.20	−0.01	−0.02	0.19
Trails A time	0.86	−0.01	−0.10	−0.12	0.20	−0.01	−0.12
Trails B time	0.80	−0.07	−0.13	−0.14	−0.03	0.10	−0.21
Stroop interference	0.73	0.05	0.12	−0.21	0.00	−0.09	0.09
Stroop Word Reading	0.68	−0.03	0.03	0.31	−0.15	−0.14	−0.09
Verbal fluency—FAS	−0.39	0.29	−0.01	0.08	0.19	0.14	0.10
Verbal fluency—animal	−0.40	0.36	0.02	0.08	0.00	−0.07	0.06
Digit symbol	−0.80	−0.02	0.14	0.07	0.08	−0.20	−0.11
Long delay free recall	0.03	1.02	−0.03	0.09	0.00	−0.05	0.04
Short delay free recall	−0.03	1.00	−0.06	−0.08	0.01	−0.08	−0.03
Trails 1–5 total raw score	−0.08	0.90	−0.05	0.05	−0.02	0.01	0.08
ROCF copy	0.09	−0.07	0.81	0.08	0.22	−0.25	0.09
Block design	−0.04	−0.05	0.73	0.30	−0.04	0.10	0.46
ROCF 3-min delay	−0.12	0.25	0.58	0.04	0.02	0.01	−0.07
WMS-III visual reproduction II	−0.06	0.34	0.58	−0.07	−0.08	0.08	−0.05
WMS-III visual reproduction I	−0.30	0.19	0.42	−0.19	−0.05	0.10	0.00
Boston Naming Test	0.04	0.03	0.18	0.79	−0.03	0.19	0.09
WMS-III-logical memory (LM) I	−0.08	0.59	0.01	0.46	0.01	0.06	−0.14
WMS-III-LM II	−0.18	0.51	0.02	0.39	0.03	0.03	−0.27
WAIS-III digit span backward	−0.04	0.02	0.16	−0.05	0.78	0.18	−0.02
WAIS-III digit span forward	0.00	−0.08	−0.12	0.36	0.17	0.77	0.02
Categories	−0.29	0.44	0.27	−0.23	0.04	−0.04	0.36
Variance explained by each factor	**68.0%**	**10.0%**	**8.0%**	**6.0%**	**4.0%**	**2.0%**	**2.0%**

Abbreviations: EFA, exploratory factor analysis; ROCF, Rey Osterrieth Complex Figure; WCST, Wisconsin Card Sorting Test; WMS, Wechsler Memory Scale.

Table 6
AUC for all measures at baseline—univariate models

Methods	NCdecliner		MCIdecliner	
	P value	AUC (95% CI)	*P* value	AUC (95% CI)
Method 1: theoretical grouping (average)				
Memory	NS	75% (0.59–0.93)	**.002**	**94% (0.87–1.0)**
Executive	NS	80% (0.66–0.95)	**.03**	**75% (0.62–0.88)**
Attention	NS	75% (0.56–0.93)	**.002**	**82% (0.71–0.92)**
Language	NS	76% (0.57–0.94)	**.03**	**76% (0.65–0.88)**
Visual spatial	NS	73% (0.53–0.93)	NS	73% (0.62–0.85)
Method 2: theoretical grouping (minimum)				
Memory	NS	73% (0.52–0.94)	**.0003**	**87% (0.79–0.95)**
Executive	NS	75% (0.57–0.93)	**.03**	**77% (0.66–0.88)**
Attention	NS	74% (0.55–0.93)	**.03**	**77% (0.66–0.88)**
Language	NS	74% (0.54–0.93)	NS	74% (0.62–0.86)
Visual spatial	NS	78% (0.61–0.96)	NS	73% (0.61–0.85)
Method 3: exploratory factor analysis—varimax rotation (average)				
Factor 1 (attention/executive)	NS	77% (0.60–0.95)	**.008**	**80% (0.67–0.91)**
Factor 2 (verbal memory)	NS	75% (0.58–0.93)	**.003**	**92% (0.83–1.0)**
Factor 3 (visuospatial/visual reproduction)	NS	74% (0.56–0.93)	NS	74% (0.62–0.85)
Method 4: exploratory factor analysis—varimax rotation (factor scores)				
Factor 1 (attention/executive)	NS	72% (0.54–0.91)	**.02**	**83% (0.68–0.97)**
Factor 2 (verbal memory)	NS	76% (0.56–0.95)	NS	80% (0.63–0.96)
Factor 3 (visuospatial/visual reproduction)	NS	72% (0.52–0.92)	NS	72% (0.53–0.91)
Method 5: exploratory factor analysis—oblique rotation (factor scores)				
Factor 1	NS	75% (0.57–0.91)	**.009**	**85% (0.73–0.99)**
Factor 2	NS	75% (0.56–0.93)	**.006**	**89% (0.76–1.00)**
Factor 3	NS	74% (0.53–0.94)	NS	77% (0.61–0.94)
Factor 4	NS	73% (0.53–0.93)	**.03**	**83% (0.68–0.98)**
Factor 5	NS	74% (0.54–0.94)	NS	73% (0.55–0.93)

Abbreviations: AUC, area under the curve; MCI, mild cognitive impairment; MCIdecliner, MCI subjects who progressed to dementia; NC, cognitively normal; NCdecliner, NC subjects who progressed to MCI; NS = not significant; WAIS-III, Wechsler Adult Intelligence Scale, third edition.

best classification accuracy of 0.94 ($P = .003$), followed by factor 1 (attention/executive, $P = .008$, AUC = 0.80). Using the factor scores from the varimax rotation, the only significant predictor of decline in the MCI group was factor 1, which comprised attention/executive ($P = .02$, AUC = 0.83). Three factors from the EFA with promax rotation significantly predicted decline to dementia: factor 1 had a classification accuracy of 85%, factor 2 had a classification accuracy of 89%, and factor 4 had a classification accuracy of 83%. Results of the leave-one-out cross validation for the MCI group are presented in Table 7.

4. Discussion

The goal of our study was to compare the ability of theoretical and statistically derived cognitive composites in predicting cognitive decline from NC to MCI and from MCI to dementia in a longitudinal cohort of well-characterized subjects. Overall, the composite indices were better at classifying subjects who were MCI at baseline than those who were normal. For subjects who were normal at baseline, classification accuracy was best for composites that included executive function tests, except in the case where the factor scores from the varimax rotation were used in which the verbal memory tests produced the best classification accuracy. However, none of these results reached statistical significance. Nevertheless, this observation is in line with previous findings that decline in nonmemory domains can be useful in predicting cognitive decline from NC to MCI [34].

Statistically significant classification accuracy was seen in MCI subjects with both EFA and theoretical composites. The theoretically derived memory domain (including both verbal and nonverbal memory) and factor 2 (verbal memory alone) using the average scores, which loaded highest on each factor had the best classification accuracy (AUC = 0.94 and 0.92, respectively). This observation is consistent with previous reports [19,31]. Although the EFA composites did not outperform the theoretically derived composite, only six cognitive tests were needed in the EFA-derived memory composite (factor 2) compared with eight cognitive tests used in the theoretically derived memory domain which to achieve similar accuracy as visual reproduction I and II loaded highest on factor 3. Performance on the executive and attention measures also achieved significance in outcome prediction in MCI. These were considered independent domains in the theoretical composites but were grouped together by the EFA as has been seen previously in other studies [32,33,35].

On close examination of the factor loadings of various tests, several additional interesting observations could be made. The attention/executive domain included, in addition to the classic tests tapping into these two functions, two

Table 7
Comparison of the AUC for MCIdecliner models—full data set versus cross validation

Methods	Full data set	Cross validation
Method 1: theoretical grouping (average)		
Memory	94% (0.87–1.0)	85% (0.74–0.96)
Executive	75% (0.62–0.88)	61% (0.46–0.77)
Attention	82% (0.71–0.92)	71% (0.59–0.84)
Language	76% (0.65–0.88)	66% (0.53–0.79)
Visual spatial	73% (0.62–0.85)	63% (0.49–0.76)
Method 2: theoretical grouping (minimum)		
Memory	87% (0.79–0.95)	81% (0.72–0.91)
Executive	77% (0.66–0.88)	67% (0.54–0.79)
Attention	77% (0.66–0.88)	69% (0.56–0.82)
Language	74% (0.62–0.86)	63% (0.50–0.76)
Visual spatial	73% (0.61–0.85)	61% (0.48–0.75)
Method 3: exploratory factor analysis—varimax rotation (average)		
Factor 1 (attention/executive)	80% (0.67–0.91)	68% (0.53–0.83)
Factor 2 (verbal memory)	92% (0.83–1.0)	86% (0.74–0.98)
Factor 3 (visuospatial/visual reproduction)	74% (0.62–0.85)	62% (0.48–0.75)
Method 4: exploratory factor analysis—varimax rotation (factor scores)		
Factor 1 (attention/executive)	83% (0.68–0.97)	64% (0.43–0.84)
Factor 2 (verbal memory)	80% (0.63–0.96)	58% (0.36–0.80)
Factor 3 (visuospatial/visual reproduction)	72% (0.53–0.91)	44% (0.23–0.66)
Method 5: exploratory factor analysis—oblique rotation (factor scores)		
Factor 1	85% (0.73–0.99)	70% (0.52–0.88)
Factor 2	89% (0.76–1.00)	77% (0.59–0.96)
Factor 3	77% (0.61–0.94)	56% (0.36–0.76)
Factor 4	83% (0.68–0.98)	64% (0.44–0.85)
Factor 5	73% (0.55–0.93)	44% (0.24–0.65)

Abbreviations: AUC, area under the curve; MCI, mild cognitive impairment; MCIdecliner, MCI subjects who progressed to dementia.

tests that are traditionally grouped in the language domain (animal and phonemic fluency). This is most likely because both of these tests have a significant executive component. To successfully perform these two tasks, the subject must understand and keep in mind the instructions he or she was given and conduct an efficient search of the mental lexicon, all the while suppressing irrelevant responses and repetition [52]. Two executive strategies commonly used to facilitate speedy retrieval of relevant items on these tasks are clustering and switching. Clustering relies on systematic exploration of subcategories (e.g., farm animals) or in the case of letter fluency tasks, by orthographic similarity [53]. Switching refers to the conscious shift to another subcategory (e.g., from farm animals to forest animals, then to aquatic animals, and so forth). Thus, performance of these tasks is dependent on both verbal ability and executive function and, of these two cognitive domains, executive abilities are affected earlier in the disease course.

It is also worth noting that EFA-derived factors 1 to 3 (varimax rotation) explained 84% of the variance between the groups with the attention/executive factor alone explaining 66% of the variance. Several studies have over the years reported the significant role executive function plays in MCI progression [14,33,34,54,55]. Using the ADNI cohort, Gibbons et al. [14] previously demonstrated that the executive domain is a powerful predictor of conversion from MCI to AD. Another Alzheimer's Disease Neuroimaging (ADNI) study reported that subjects with amnestic MCI show faster decline in executive function compared with memory [34].

There are a variety of approaches to factor analysis based on assumptions made about a particular data set. Although using an orthogonal rotation produces more easily interpretable results, in our data we would expect some correlation between the results of the NP tests, whether they are considered memory, executive function, attention, language, or visuospatial. Therefore, using an oblique rotation may be more appropriate. In our data set, we found that the factor scores produced from the promax rotation indeed outperformed those from the varimax rotation. Three factors—factor 1, factor 2, and factor 4—significantly predicted decline to dementia with AUCs of 85%, 89%, and 83% respectively, whereas only factor 1 from the varimax rotation significantly predicted decline (AUC = 0.83). Although these results are good evidence for using the oblique rotation, both methods using the factor scores did not perform as well as the composites calculated from averaging, which is a more straightforward procedure.

There are some limitations to this study that should be considered. Our subjects were predominantly Caucasian and highly educated. Several cross-sectional studies have shown that there are significant differences in cognitive performance across races [56–59]. These studies found higher rates of cognitive impairment, dementia, and AD among minorities compared with Whites. Some studies have also shown that the racial differences also extend to cognitive decline over time. Looking at telephone interview for cognitive status scores for more than 3 years of 113 MCI subjects from the Memory and Medical Care Study, Lee et al. (2012) found that the rate of cognitive decline appears to be faster in African Americans than non–African Americans in the community [60]. Another study, however, showed that the rate of cognitive decline in 452 clinically diagnosed AD patients appears to be slower in African Americans than non–African Americans [61]. Taken together, these studies suggest that cognitive ability varies across races. As such, our findings may not generalize well to the broad community. Another limitation, which is likely responsible for the lack of significance in the predictive model in NC, is the relatively small sample size of the NC group, especially the NCdecliner group.

In conclusion, none of the EFA-derived or theoretical domains were able to predict decline in NC with sufficient accuracy. Equally excellent predictive accuracy of future decline was seen with both the EFA and the theoretical memory composites in MCI; however, the EFA solution achieved this accuracy using fewer cognitive tests relative to the theoretical approach. Therefore, in research settings in which a briefer battery is desirable, the EFA approach might be preferable. In addition to memory, the executive

function composite also achieved significance in predicting cognitive decline in MCI.

Acknowledgments

This work was generously supported by R01 AG040770, K02 AG048240, P50 AG16570, Easton Consortium for Alzheimer's Drug Discovery and Biomarker Development.

RESEARCH IN CONTEXT

1. Systematic review: Both theoretical and data-driven composite scores have been used in research settings for predicting cognitive decline. Our literature review included extensive PubMed database search for original research examining the performance of theoretical and data-driven approaches in diagnosing cognitive impairment and predicting future cognitive decline associated with Alzheimer's disease.
2. Interpretation: Our objective was to compare the accuracy of theoretically and statistically derived neuropsychological composites for predicting cognitive decline. Both the theoretical and the exploratory factor analyses resulted in composite measures in the memory and executive domains that achieved high accuracy in predicting progression from mild cognitive impairment to dementia. Neither method was sensitive enough to predict future cognitive decline in cognitively normal individuals.
3. Future directions: Future studies including new cognitive measures highly sensitive to subtle cognitive decline might be needed to identify cognitively normal individuals in the asymptomatic stages of Alzheimer's disease.

References

[1] Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, et al. Current concepts in mild cognitive impairment. Arch Neurol 2001; 58:1985–92.
[2] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:270–9.
[3] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.
[4] Alzheimer's Association. 2014 Alzheimer's disease facts and figures. Alzheimers Dement 2014;10:e47–92.
[5] Ganguli M, Snitz BE, Saxton JA, Chang CC, Lee CW, Vander Bilt J, et al. Outcomes of mild cognitive impairment by definition: a population study. Arch Neurol 2011;68:761–7.
[6] Hayden KM, Jones RN, Zimmer C, Plassman BL, Browndyke JN, Pieper C, et al. Factor structure of the National Alzheimer's Coordinating Centers uniform dataset neuropsychological battery: an evaluation of invariance between and within groups over time. Alzheimer Dis Assoc Disord 2011;25:128–37.
[7] Park LQ, Gross AL, McLaren DG, Pa J, Johnson JK, Mitchell M, et al. Confirmatory factor analysis of the ADNI neuropsychological battery. Brain Imaging Behav 2012;6:528–39.
[8] Weintraub S, Wicklund AH, Salmon DP. The neuropsychological profile of Alzheimer disease. Cold Spring Harb Perspect Med 2012; 2:a006171.
[9] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939–44.
[10] Diehl J, Monsch AU, Aebi C, Wagenpfeil S, Krapp S, Grimmer T, et al. Frontotemporal dementia, semantic dementia, and Alzheimer's disease: the contribution of standard neuropsychological tests to differential diagnosis. J Geriatr Psychiatry Neurol 2005;18:39–44.
[11] Tierney MC, Yao C, Kiss A, McDowell I. Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. Neurology 2005;64:1853–9.
[12] Backman L, Jones S, Berger AK, Laukka EJ, Small BJ. Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. Neuropsychology 2005;19:520–31.
[13] Salmon DP, Thomas RG, Pay MM, Booth A, Hofstetter CR, Thal LJ, et al. Alzheimer's disease can be accurately diagnosed in very mildly impaired individuals. Neurology 2002;59:1022–8.
[14] Gibbons LE, Carle AC, Mackin RS, Harvey D, Mukherjee S, Insel P, et al. A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. Brain Imaging Behav 2012;6:517–27.
[15] Clark LR, Schiehser DM, Weissberger GH, Salmon DP, Delis DC, Bondi MW. Specific measures of executive function predict cognitive decline in older adults. J Int Neuropsychol Soc 2012;18:118–27.
[16] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999;56:303–8.
[17] Salmon DP, Bondi MW. Neuropsychological assessment of dementia. Annu Rev Psychol 2009;60:257–82.
[18] Irish M, Lawlor BA, Coen RF, O'Mara SM. Everyday episodic memory in amnestic mild cognitive impairment: a preliminary investigation. BMC Neurosci 2011;12:80.
[19] Vliet EC, Manly J, Tang MX, Marder K, Bell K, Stern Y. The neuropsychological profiles of mild Alzheimer's disease and questionable dementia as compared to age-related cognitive decline. J Int Neuropsychol Soc 2003;9:720–32.
[20] Delis DC, Massman PJ, Butters N, Salmon DP, Cermak LS, Kramer JH. Profiles of demented and amnestic patients on the California Verbal Learning Test: implications for the assessment of memory disorders. J Consult Clin Psych 1991;3:19–26.
[21] Herlitz A, Hill RD, Fratiglioni L, Backman L. Episodic memory and visuospatial ability in detecting and staging dementia in a community-based sample of very old adults. J Gerontol A Biol Sci Med Sci 1995;50:M107–13.
[22] Welsh K, Butters N, Hughes J, Mohs R, Heyman A. Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. Arch Neurol 1991;48:278–81.
[23] Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 1989;39:1159–65.
[24] Perry RJ, Hodges JR. Attention and executive deficits in Alzheimer's disease. A critical review. Brain 1999;122:383–404.
[25] Marshall GA, Rentz DM, Frey MT, Locascio JJ, Johnson KA, Sperling RA, et al. Executive function and instrumental activities of

daily living in mild cognitive impairment and Alzheimer's disease. Alzheimers Dement 2011;7:300–8.
[26] Rozzini L, Chilovi BV, Conti M, Bertoletti E, Delrio I, Trabucchi M, et al. Conversion of amnestic mild cognitive impairment to dementia of Alzheimer type is independent to memory deterioration. Int J Geriatr Psychiatry 2007;22:1217–22.
[27] Patterson MB, Mack JL, Geldmacher DS, Whitehouse PJ. Executive functions and Alzheimer's disease: problems and prospects. Eur J Neurol 1996;3:5–15.
[28] Heaton RK, Chelune GJ, Talley JL, Kay GG, Curtiss G. Wisconsin card sorting test manual: revised and expanded. Odessa, FL: Psychological Assessment Resources Inc; 1993.
[29] Stroop JR. Studies in interference of serial verbal reactions. J Exp Psychol 1935;18:643–62.
[30] Binetti G, Magni E, Padovani A, Cappa SF, Bianchetti A, Trabucchi M. Executive dysfunction in early Alzheimer's disease. J Neurol Neurosurg Psychiatry 1996;60:91–3.
[31] Chapman RM, Mapstone M, McCrary JW, Gardner MN, Porsteinsson A, Sandoval TC, et al. Predicting conversion from mild cognitive impairment to Alzheimer's disease using neuropsychological tests and multivariate methods. J Clin Exp Neuropsychol 2011;33:187–99.
[32] Chapman RM, Mapstone M, Porsteinsson AP, Gardner MN, McCrary JW, DeGrush E, et al. Diagnosis of Alzheimer's disease using neuropsychological testing improved by multivariate analyses. J Clin Exp Neuropsychol 2010;32:793–808.
[33] Dowling NM, Hermann B, La Rue A, Sager MA. Latent structure and factorial invariance of a neuropsychological test battery for the study of preclinical Alzheimer's disease. Neuropsychology 2010;24:742–56.
[34] Johnson JK, Gross AL, Pa J, McLaren DG, Park LQ, Manly JJ. Longitudinal change in neuropsychological performance using latent growth models: a study of mild cognitive impairment. Brain Imaging Behav 2012;6:540–50.
[35] Mitchell MB, Shaughnessy LW, Shirk SD, Yang FM, Atri A. Neuropsychological test performance and cognitive reserve in healthy aging and the Alzheimer's disease spectrum: a theoretically driven factor analysis. J Int Neuropsychol Soc 2012;18:1071–80.
[36] Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods 1999;4:272–99.
[37] Siedlecki KL, Honig LS, Stern Y. Exploring the structure of a neuropsychological battery across healthy elders and those with questionable dementia and Alzheimer's disease. Neuropsychology 2008;22:400–11.
[38] Reitan RM, Wolfson D. The Halstead-Reitan neuropsychological test battery: therapy and clinical interpretation. Tucson, AZ: Neuropsychological Press; 1985.
[39] Osterrieth PA. Le test de copie d'une figure complexe. Arch Psychol 1944;30:206–356.
[40] Kafer KL, Hunter M. On testing the face validity of planning/problem-solving tasks in a normal population. J Int Neuropsychol Soc 1997; 3:108–19.
[41] Strauss E, Sherman EMS, Spreen O. A compendium of neuropsychological tests: administration, norms, and commentary. 3rd ed. New York: Oxford University Press; 2006.
[42] Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval 2005;10.
[43] Yesavage JA. Geriatric depression scale. Psychopharmacol Bull 1988; 24:709–11.
[44] McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology 2005;65:1863–72.
[45] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98.
[46] Wechsler D. Wechsler. San Antonio, TX: The Psychological Corporation; 2001.
[47] Wechsler D. Manual for the Wechsler adult intelligence scale. 3rd ed. San Antonio, TX: The Psychological Corporation; 1997.
[48] Wechsler D. Wechsler memory scale—third edition manual. San Antonio, TX: The Psychological Corporation; 1997.
[49] Kaplan EF, Goodgalss H, Weintraub S. The Boston naming test. Philadelphia: Lea & Febiger; 1983.
[50] Benton AL, Hamsher K. Multilingual aphasia examination. Iowa City, Iowa: AJA Associates.; 1989.
[51] Delis DC, Kramer JH, Kaplan E, Ober BA. California verbal learning test, Form II. 2nd ed. San Antonio, TX: The Psychological Corporation; 1987.
[52] Shao Z, Janse E, Visser K, Meyer AS. What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. Front Psychol 2014;5:772.
[53] Clark DG, Wadley VG, Kapur P, DeRamus TP, Singletary B, Nicholas AP, et al. Lexical factors and cerebral regions influencing verbal fluency performance in MCI. Neuropsychologia 2014; 54:98–111.
[54] Albert MS, Moss MB, Tanzi R, Jones K. Preclinical prediction of AD using neuropsychological tests. J Int Neuropsychol Soc 2001; 7:631–9.
[55] Perri R, Serra L, Carlesimo GA, Caltagirone C, Early Diagnosis Group of the Italian Interdisciplinary Network on Alzheimer's Disease. Amnestic mild cognitive impairment: difference of memory profile in subjects who converted or did not convert to Alzheimer's disease. Neuropsychology 2007;21:549–58.
[56] Gurland BJ, Wilder DE, Lantigua R, Stern Y, Chen J, Killeffer EH, et al. Rates of dementia in three ethnoracial groups. Int J Geriatr Psychiatry 1999;14:481–93.
[57] Mehta KM, Simonsick EM, Rooks R, Newman AB, Pope SK, Rubin SM, et al. Black and white differences in cognitive function test scores: what explains the difference? J Am Geriatr Soc 2004;52:2120–7.
[58] Perkins P, Annegers JF, Doody RS, Cooke N, Aday L, Vernon SW. Incidence and prevalence of dementia in a multiethnic cohort of municipal retirees. Neurology 1997;49:44–50.
[59] Manly JJ, Jacobs DM, Sano M, Bell K, Merchant CA, Small SA, et al. Cognitive test performance among nondemented elderly African Americans and whites. Neurology 1998;50:1238–45.
[60] Lee HB, Richardson AK, Black BS, Shore AD, Kasper JD, Rabins PV. Race and cognitive decline among community-dwelling elders with mild cognitive impairment: findings from the Memory and Medical Care Study. Aging Ment Health 2012;16:372–7.
[61] Barnes LL, Wilson RS, Hebert LE, Scherr PA, Evans DA, Mendes de Leon CF. Racial differences in the association of education with physical and cognitive function in older blacks and whites. J Gerontol B Psychol Sci Soc Sci 2011;66:354–63.