

This article was downloaded by: [University of Rochester]

On: 2 February 2011

Access details: Access Details: [subscription number 931557198]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Clinical and Experimental Neuropsychology

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713657736>

Predicting conversion from mild cognitive impairment to Alzheimer's disease using neuropsychological tests and multivariate methods

Robert M. Chapman^a; Mark Mapstone^b; John W. McCrary^a; Margaret N. Gardner^c; Anton Porsteinsson^d; Tiffany C. Sandoval^c; Maria D. Guillily^c; Elizabeth DeGrush^c; Lindsey A. Reilly^c

^a Brain and Cognitive Sciences and Center for Visual Science, University of Rochester, Rochester, NY, USA ^b Neurology, University of Rochester Medical Center, Rochester, NY, USA ^c Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA ^d Psychiatry, University of Rochester Medical Center, Rochester, NY, USA

First published on: 13 August 2010

To cite this Article Chapman, Robert M. , Mapstone, Mark , McCrary, John W. , Gardner, Margaret N. , Porsteinsson, Anton , Sandoval, Tiffany C. , Guillily, Maria D. , DeGrush, Elizabeth and Reilly, Lindsey A.(2011) 'Predicting conversion from mild cognitive impairment to Alzheimer's disease using neuropsychological tests and multivariate methods', Journal of Clinical and Experimental Neuropsychology, 33: 2, 187 – 199, First published on: 13 August 2010 (iFirst)

To link to this Article: DOI: 10.1080/13803395.2010.499356

URL: <http://dx.doi.org/10.1080/13803395.2010.499356>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Predicting conversion from mild cognitive impairment to Alzheimer's disease using neuropsychological tests and multivariate methods

Robert M. Chapman,¹ Mark Mapstone,² John W. McCrary,¹ Margaret N. Gardner,³ Anton Porsteinsson,⁴ Tiffany C. Sandoval,³ Maria D. Guillily,³ Elizabeth DeGrush,³ and Lindsey A. Reilly³

¹Brain and Cognitive Sciences and Center for Visual Science, University of Rochester, Rochester, NY, USA

²Neurology, University of Rochester Medical Center, Rochester, NY, USA

³Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA

⁴Psychiatry, University of Rochester Medical Center, Rochester, NY, USA

Behavioral markers measured through neuropsychological testing in mild cognitive impairment (MCI) were analyzed and combined in multivariate ways to predict conversion to Alzheimer's disease (AD) in a longitudinal study of 43 MCI patients. The test measures taken at a baseline evaluation were first reduced to underlying components (principal component analysis, PCA), and then the component scores were used in discriminant analysis to classify MCI individuals as likely to convert or not. When empirically weighted and combined, episodic memory, speeded executive functioning, recognition memory (false and true positives), visuospatial memory processing speed, and visuospatial episodic memory were together strong predictors of conversion to AD. These multivariate combinations of the test measures achieved through the PCA were good, statistically significant predictors of MCI conversion to AD (84% accuracy, 86% sensitivity, and 83% specificity). Importantly, the posterior probabilities of group membership that accompanied the binary prediction for each participant indicated the confidence of the prediction. Most of the participants (81%) were in the highly confident probability bins (.70–1.00), where the obtained prediction accuracy was more than 90%. The strength and reliability of this multivariate prediction method were tested by cross-validation and randomized resampling.

Keywords: Neuropsychological tests; Predict; Multivariate; Alzheimer's disease; Mild cognitive impairment; Discriminant analysis; Principal component analysis; Posterior probability.

Recent large-scale studies of mild cognitive impairment (MCI) have suggested that not all patients with MCI will convert to Alzheimer's disease (AD; Petersen, 2004). As a result, identifying those with MCI who are likely to

convert to AD is becoming increasingly important. Early identification of MCI patients who will convert to AD is essential to timely administration of pharmacologic and therapeutic interventions as well as to determining with

Maria Guillily is now at the Department of Pharmacology and Experimental Therapeutics at Boston University. Tiffany Sandoval is now at the San Diego State University/University of California at San Diego Joint Doctoral Program in Clinical Psychology. Elizabeth DeGrush is now at the Chicago College of Osteopathic Medicine at Midwestern University. Lindsey Reilly is now at the Springer Publishing Company. We thank: the Memory Disorders Clinic at the University of Rochester Medical Center, Monroe Community Hospital, the Alzheimer's Disease Center, especially Paul Coleman, Charles Duffy, and Roger Kurlan, for their strong support of our research; Susan E. Chapman for help in writing; Robert Emerson and William Vaughn for their technical contributions; Courtney Vargas, Dustina Holt, Jonathan DeRight, Cendrine Robinson, Kristen Morie, Anna Fagan, Michael Garber-Barron, Leon Tsao, and Brittany Huber for technical help; and the many voluntary participants in this research. This research was supported by the National Institute of Health Grants P30-AG08665, R01-AG018880, and P30-EY01319.

Address correspondence to Robert M. Chapman, Center for Visual Science at the University of Rochester, 775 Library Road, Rochester, NY 14627–0270, USA (E-mail: rmc@cvs.rochester.edu).

some confidence which participants with memory disorders are appropriate for various research studies. Amnesic MCI is a clinical diagnosis commonly characterized by a memory deficit, which does not interfere significantly with activities of daily living (Petersen et al., 1999). This memory deficit must fall at least 1.5 standard deviations below age-adjusted performance on standardized tests of memory and should be corroborated by an informant (Petersen, 2004). While the majority of MCI patients have primary memory deficits, the behavioral heterogeneity of MCI is becoming clearer. Numerous reports now suggest that MCI can present initially as a primary impairment in other cognitive domains including language, visuospatial or visuo-perceptual abilities (Mapstone, Steffenella, & Duffy, 2003), executive function, or even affect. The notion of multidomain MCI complicates early diagnosis as there are numerous disorders that may cause subtle cognitive deficits in multiple domains. Because the primary cognitive deficit in most cases of MCI and AD is in the domain of memory, it is not surprising that memory performance, particularly new learning, recall, and retention, are strong predictors of conversion from MCI to AD. However, some studies have suggested that at baseline other cognitive domains, including executive function and lower cognitive abilities, better predict conversion from MCI to AD (e.g., Rozzini et al., 2008). Still other studies suggest that nonamnesic or multidomain MCI patients convert to AD at lower rates than amnesic MCI patients (Maioli et al., 2007). These discrepancies highlight the behavioral heterogeneity of the clinical presentation of MCI and AD and the difficulties of applying findings based on groups of patients to individual patients.

Measuring group differences is an important aspect of understanding the differing cognitive processes between progressing to dementia and remaining stable; however, predicting individual outcomes is essential to early intervention in patients. Despite the great interest in identifying MCI patients who are at high risk for developing AD, presently there are no clinical or imaging markers that predict conversion to AD with certainty or that reliably establish which MCI participants will convert (Brayne, 2007; Marcos et al., 2006). While ongoing research emphasizes biological markers found in analyses of blood and cerebrospinal fluid (Bateman, Wen, Morris, & Holtzman, 2007; Papaliagkas, Anogianakis, Tsolaki, Koliakos, & Kimiskidis, 2009; Simonsen et al., 2007), in anatomical and functional brain imaging studies (Klunk et al., 2004), in event-related potential studies (Chapman et al., in press; Chapman et al., 2007), and by other biomedical techniques, here we will investigate whether behavioral markers measured through neuropsychological testing can be analyzed and combined in multivariate ways to predict conversion to AD. Because cognitive changes are a prominent and early feature of AD, focusing on neuropsychological markers for conversion would seem appropriate. Neuropsychological testing, which is relatively inexpensive and noninvasive to the patient, has long been used in the clinical assessment of AD.

Studies have shown that neuropsychological measures of episodic memory have good power in predicting MCI progression to AD (Albert, Blacker, Moss, Tanzi, & McArdle, 2007; Albert, Moss, Tanzi, & Jones, 2001; Bondi et al., 1994; Lekeu et al., 2010; Marcos et al., 2006; Perri, Serra, Carlesimo, & Caltagirone, 2007). Also, non-memory measures have been studied in the interest of increasing predictive success (Babins, Slater, Whitehead, & Chertkow, 2008; Bäckman, Jones, Berger, Laukka, & Small, 2005; Lekeu et al., 2010; Marcos et al., 2006; Rozzini et al., 2008; Tabert et al., 2006; Tierney, Yao, Kiss, & McDowell, 2005). In this article, we test whether employing multivariate methods to combine data from differing cognitive domains can predict conversion to AD in MCI patients. We study the predictive power of the tests using a multivariate method with two levels of empirically derived weighting. The test measures are first reduced to underlying components (principal component analysis, PCA), and then the component scores are combined in a weighted, linear fashion (discriminant analysis) to classify individuals. PCA resolves a correlation matrix of test measures from a set of participants into underlying components, and each participant receives a component score for each component. This data reduction from many test measures to a few components is also key in reducing concerns about degrees of freedom in the discriminant analyses. We hypothesize that these multivariate composite measures will be able to predict AD in MCI individuals with strong, statistically significant success. The reliability and influence of chance will be assessed through validation and randomized resampling analyses.

METHOD

Study participants

We studied 43 elderly individuals diagnosed with MCI (Table 1). These participants were recruited from the Memory Disorders Clinic at the University of Rochester and other affiliated University of Rochester clinics. The MCI participants were evaluated by memory-disorders physicians and met current consensus criteria for the amnesic subtype of MCI ("a-MCI"; Petersen, 2004; Petersen et al., 1999; Petersen et al., 2001). (In this article, we use the term "MCI" to refer to amnesic MCI.) Each MCI participant was subsequently found either to have converted to clinically defined AD—by the NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association) criteria (McKhann et al., 1984) and *DSM-IV-TR* (*Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition, Text Revision*; American Psychiatric Association, 2000) criteria for dementia of the Alzheimer's type—or to have remained stable with regard to cognitive state. These determinations were made at a later date through clinical follow-ups by the same memory-disorders physicians, who were blind to our study data. Those who converted were given the clinical diagnosis of "probable" AD (but

TABLE 1
Participant demographics

Participants	Group	n	Gender		Age (years)		Education (years)	
			No. females	No. males	Mean	SD	Mean	SD
MCI participants used in prediction analyses	Conversion	14	3	11	76.6	5.4	15.7	2.0
	Stable	29	15	14	75.2	8.6	14.3	2.7
Additional participants used in PCA	AD	55	24	31	76.0	7.4	14.1	3.1
	Controls	78	47	31	70.3	10.0	15.9	2.6
	MCI	35	16	19	71.2	8.5	14.7	2.6
	AAMI	5	2	3	68.6	7.3	17.6	2.5

Note. $N = 216$. MCI = mild cognitive impairment. AD = Alzheimer's disease. AAMI = age-associated memory impairment. PCA = principal component analysis. For the MCI participants used in prediction analyses, there was no significant group effect (conversion or stable), gender effect, or group by gender interaction for the age and education demographics. Demographic information for each participant was collected at baseline evaluations. MCI individuals were placed into the conversion or stable groups based upon subsequent diagnoses at follow-up evaluations.

referred to here as AD for brevity's sake). Of the 43 MCI patients, 14 were subsequently diagnosed with AD (the conversion to AD group, or conversion group), and 29 were not (the stable group). The clinical diagnoses of MCI and AD were based on the history, relevant laboratory findings, and imaging studies routinely performed as part of the clinical assessment of dementia (Petersen et al., 2001). Separate cognitive testing was performed by the memory-disorders physicians to assist with their diagnoses; these tests included the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), a clock face drawing, the Auditory Verbal Learning Test (Rey, 1964; Taylor, 1959), and a category fluency task (animal naming). With the exception of the MMSE, the clock face drawing, and the category fluency task (animal naming), all of which had small weights in the components used in discrimination, no cognitive test used in clinical decision making was repeated as part of our experimental cognitive test battery described below. Thus, our study maintained relative independence between predictors and diagnostic outcomes (Tierney et al., 2005).

The median number of months between the initial diagnosis of MCI and the subsequent diagnosis of AD was 19.7 (interquartile range p_{25} – $p_{75} = 10.1$ – 24.4) for the conversion group. For the stable group, the median number of months between the initial MCI diagnosis and the most recent clinical work-up was 19.6 (interquartile range p_{25} – $p_{75} = 10.1$ – 27.6). The gender, age, and education demographics for each group appear in Table 1. There were no significant group or gender differences for age and education. In the conversion group, 7 of the 14 individuals were taking cholinesterase inhibitors and/or memantine at the time of testing. In the stable group, 13 of the 29 individuals were taking these medications. The proportions taking these medications were not significantly different between these groups: Fisher's Exact Test, $\chi^2(1, N = 43) = 0.04, p = .69$.

Exclusion criteria for all groups included clinical (or imaging) evidence of stroke, Parkinson's disease, HIV/AIDS, and reversible dementias, as well as treatment

with benzodiazepines, antipsychotic, or antiepileptic medications. As an additional inclusion criterion for our study, all clinical participants had a previously administered score of 21 or higher on the MMSE—this criterion included AD participants; the MCI participants used in this study had mean MMSE scores of 25 or 27 (conversion or stable) as shown in Table 2. There was no significant difference between the two MCI subgroups in comorbid depressive symptoms (as shown through the Geriatric Depression Scale) or in impact of disease on daily activities (indicated by the Blessed Dementia Scale)—see Table 2. In general, the mean scores for the Geriatric Depression Scale for each group were considered “normal” for depressive symptoms (Hickie & Snowden, 1987).

Our study received IRB (institutional review board) approval from the University of Rochester Research Subjects Review Board, and informed consent was obtained from each participant.

Additional participants used in PCA

More participants were used to develop the neuropsychological component structure with PCA. In addition to the 43 MCI participants used in discriminant analysis, for the PCA analysis we also included 55 elderly individuals diagnosed with AD, 78 individuals with normal cognition (controls), 5 individuals diagnosed with age-associated memory impairment (Crook et al., 1986), and 35 more MCI participants (totaling 216 participants). These 35 additional MCI individuals did not return for follow-up evaluations, and therefore their subsequent clinical outcomes were not known at the time of this analysis. Enlarging the set of participants for the development of the component structure was done for several reasons. First, increasing the number of observations added stability to the resultant structure. Second, including a variety of participant groups in the creation of the component structure allowed for greater generalizability to the population (Chapman et al., 2010a; John, Easton, Pritchep, & Friedman, 1993). Using data from only one

TABLE 2
MCI conversion to AD and stable group means for each of the 49 neuropsychological test measures

Test measure	Raw scores		Standard (z scores)		<i>p</i> ^c
	Conversion (<i>n</i> = 14)	Stable (<i>n</i> = 29)	Conversion (<i>n</i> = 14)	Stable (<i>n</i> = 29)	
<i>Rey–Osterrieth Complex Figure (Osterrieth, 1944; Rey, 1941)</i>					
Copy Score	27.3 (7.3)	31.0 (3.8)	−1.3 (2.0)	−0.2 (0.9)	.02
Copy Speed ^a	3.7 (2.3) ^b	4.9 (2.7) ^b	−0.6 (0.8)	−0.2 (0.9)	.17
Immediate Recall Score	4.7 (5.2)	11.1 (7.2)	−1.6 (1.4)	−0.3 (1.6)	<.01
Immediate Recall Speed ^a	10.0 (7.6) ^b	8.4 (4.5) ^b	0.2 (1.1)	−0.1 (0.7)	.38
Delayed Recall Score	5.9 (5.5)	10.1 (7.4)	−1.5 (1.4)	−0.5 (1.8)	.07
Delayed Recall Speed ^a	19.2 (21.3) ^b	13.2 (8.8) ^b	0.8 (2.4)	0.1 (1.0)	.20
<i>Mini-Mental State Examination (Folstein et al., 1975)</i>					
Score	25.1 (3.2)	27.1 (2.6)	−1.9 (2.0)	−0.3 (2.1)	.02
<i>WMS–III Digit Span (Wechsler, 1997a)</i>					
Forward Score	6.2 (1.1)	6.2 (1.0)	0.1 (0.8)	0.1 (0.8)	.92
Backward Score	4.6 (1.1)	5.0 (1.3)	0.3 (0.8)	0.6 (1.1)	.35
Letter–Number Score	4.0 (1.4)	5.0 (1.3)	−1.4 (0.6)	−1.1 (0.7)	.13
<i>Geriatric Depression Scale (Yesavage et al., 1983)</i>					
Score	7.3 (7.4)	6.9 (4.8)	−1.6 (3.0)	−1.4 (1.7)	.71
<i>WMS–III Logical Memory I (Wechsler, 1945, 1987, 1997b)</i>					
A Recall Score ^a	8.3 (4.5)	12.0 (4.4)	−1.7 (1.1)	−0.7 (1.1)	.01
B1 Recall Score ^a	6.1 (3.7)	9.5 (3.1)	−1.8 (1.0)	−0.9 (0.9)	<.01
B2 Recall Score ^a	8.9 (4.1)	13.1 (4.4)	−2.1 (1.1)	−1.0 (1.2)	<.01
<i>WMS–III Logical Memory II (Wechsler, 1945, 1987, 1997b)</i>					
A Recall Score ^a	3.6 (5.2)	8.0 (5.6)	−2.2 (1.2)	−1.2 (1.3)	<.01
B Recall Score ^a	5.3 (5.6)	9.9 (5.5)	−2.4 (1.4)	−1.2 (1.4)	<.01
Recognition Score ^a	20.6 (5.2)	24.1 (3.8)	−3.2 (2.5)	−1.5 (1.8)	.02
Percent Retention ^a	42.9 (39.7)	68.2 (28.6)	−3.5 (3.0)	−1.6 (2.2)	.02
<i>Clock Drawing Test (Tuokko, Hadjistavropoulos, Miller, & Beattie, 1992)</i>					
Score	18.5 (1.2)	18.6 (2.2)	0.1 (0.6)	0.2 (1.2)	.83
<i>North American National Adult Reading Test (Grober & Sliwinski, 1991)</i>					
Score	37.6 (7.1)	37.5 (8.2)	−0.9 (3.0)	−0.9 (2.2)	.91
<i>Stroop Test (Golden, 1978)</i>					
Word Score	83.9 (19.6)	95.0 (13.4)	−1.4 (1.5)	−0.4 (1.1)	.02
Color Score	52.0 (13.7)	61.6 (11.5)	−1.8 (1.2)	−0.9 (1.0)	.01
Color–Word Score	26.4 (10.4)	27.4 (9.2)	−0.8 (1.0)	−0.6 (0.9)	.45
<i>Brief Visuospatial Memory Test–Revised (Benedict, 1997; Benedict & Groninger, 1995)</i>					
Trial 1 Score	1.6 (1.2)	2.8 (1.7)	−1.6 (0.6)	−1.0 (0.8)	.02
Trial 2 Score	2.4 (2.0)	4.4 (3.0)	−0.6 (1.2)	0.2 (1.4)	.06
Trial 3 Score	3.2 (2.5)	5.4 (3.3)	−2.3 (1.1)	−1.4 (1.4)	.03
Learning Slope	1.8 (1.9)	2.9 (2.0)	−1.1 (1.0)	−0.5 (1.1)	.11
Delayed Recall	2.0 (2.1)	4.5 (3.3)	−2.4 (0.9)	−1.4 (1.3)	.01
Percent Retention	43.6 (38.7)	72.5 (33.9)	−3.2 (2.8)	−1.0 (2.5)	.02
Hits	5.3 (0.9)	5.1 (0.9)	−0.6 (1.4)	−1.0 (1.4)	.35
False Alarms	1.0 (1.0)	0.5 (0.9)	2.4 (2.8)	1.1 (2.5)	.14
Discrimination Index	4.7 (1.6)	4.4 (1.3)	−0.9 (2.0)	−1.4 (1.7)	.41
<i>Controlled Oral Word Association Test (Benton & Hamsher, 1976)</i>					
F Score ^a	12.0 (6.0)	12.5 (5.1)	−0.5 (1.3)	−0.4 (1.1)	.79
A Score ^a	10.2 (5.2)	11.1 (4.0)	−0.6 (1.2)	−0.4 (0.9)	.57
S Score ^a	11.4 (6.6)	14.3 (4.9)	−0.8 (1.4)	−0.2 (1.0)	.11
<i>Category Fluency (Benton & Hamsher, 1976)</i>					
Animal–Naming Score	13.1 (5.5)	16.3 (5.2)	−1.0 (1.2)	−0.1 (1.4)	.04

(Continued)

TABLE 2
(Continued)

Test measure	Raw scores		Standard (<i>z</i> scores)		<i>p</i> ^c
	Conversion (<i>n</i> = 14)	Stable (<i>n</i> = 29)	Conversion (<i>n</i> = 14)	Stable (<i>n</i> = 29)	
<i>Blessed Dementia Scale</i> (Blessed, Tomlinson, & Roth, 1968; Morris et al., 1989)					
Score	-1.6 (1.8)	-1.4 (1.5)	-1.0 (1.8)	-0.8 (1.4)	.71
<i>Hopkins Verbal Learning Test</i> (Brandt, 1991)					
Trial 1 Score	3.8 (1.9)	5.5 (1.9)	-1.2 (1.1)	-0.3 (1.0)	.01
Trial 2 Score	6.6 (1.9)	8.1 (2.1)	-0.7 (0.9)	-0.1 (1.0)	.05
Trial 3 Score	6.6 (2.4)	9.0 (2.1)	-1.3 (1.2)	-0.1 (1.0)	<.01
Delayed Recall Score	2.7 (3.9)	5.9 (3.5)	-2.1 (1.6)	-1.0 (1.6)	.04
True Positives	10.1 (2.6)	10.9 (1.3)	-1.1 (2.5)	-0.4 (1.4)	.22
Related False Positives	2.4 (1.7)	1.4 (1.5)	1.5 (1.6)	0.6 (1.5)	.07
Unrelated False Positives	1.1 (1.4)	0.2 (0.6)	3.4 (4.6)	0.5 (1.9)	<.01
Discrimination Index	6.6 (3.4)	9.3 (2.2)	-2.3 (2.2)	-0.7 (1.4)	<.01
<i>Boston Naming Test</i> (Kaplan, Goodglass, & Weintraub, 1978; Mack, Freed, Williams, & Henderson, 1992)					
Score	13.9 (1.2)	14.3 (0.8)	-0.9 (1.7)	-0.4 (1.1)	.22
<i>Standardized Road-Map Test of Direction</i> (Money, 1976)					
Score	27.0 (7.2)	25.8 (7.1)	-0.6 (2.1)	-1.2 (3.0)	.49
<i>Trail Making Test</i> (Reitan, 1958)					
A Speed ^a	21.5 (7.0) ^b	27.5 (9.4) ^b	-0.8 (0.5)	-0.4 (0.7)	.04
B Speed ^a	7.1 (3.8) ^b	11.3 (4.6) ^b	-1.1 (0.6)	-0.5 (0.7)	<.01

Note. MCI = mild cognitive impairment. AD = Alzheimer's disease. The conversion group contained 14 MCI individuals who later developed AD. The stable group contained 29 MCI individuals who did not develop AD. Group means shown for raw scores and *z* scores; standard deviations in parentheses.

^a*z* scores for these test measures were generated from laboratory data (normal elderly) because published age/education-corrected normative data were not available. ^bMean speed score and standard deviation are *s*⁻¹ multiplied by 100. ^cThe *p* values for group mean differences (*t*-tests computed on standard scores) are unadjusted for multiple comparisons; using a Bonferroni adjustment resulted in no test measure having a significant group mean difference at *p* < .05.

group also would risk restricting the range in the test measures and attenuating correlations among variables, which could result in falsely low estimates of component loadings (Fabrigar, MacCullum, Wegener, & Stahan, 1999). This risk is reduced by involving data from multiple groups of individuals.

These additional participants were also evaluated by the same memory-disorders physicians from area clinics. Demographic information for the additional participants also appears in Table 1 (for more detailed demographic and neuropsychological information concerning the AD and control participants, see Chapman et al., 2010b).

Neuropsychological assessment

The neuropsychological battery we administered to each MCI participant contained 17 common tests (total of 49 measures; Table 2) that target all eight cognitive domains as defined by the NINCDS-ADRDA criteria, particularly memory. We designed the battery to produce a comprehensive sample of cognitive processes and their degeneration in AD. Among others, the tests included measures of memory retrieval and retention, generative fluency, executive function, visuospatial abilities, and attributes of

mood and daily living. Each participant's battery of raw scores was transformed to standard scores using established age/education-corrected normative data when possible or laboratory-derived data (normal elderly) when published norms were not available. Normalizing the data limited the influence of age, education, and gender effects. The standard *z* scores were used for all statistical analyses described in this paper.

Group mean differences for each of the 49 neuropsychological test measures are included as baseline characterization of the conversion and stable groups.

Developing component scores (PCA)

Principal component analysis (PCA) was used to develop the component structure from the battery of neuropsychological tests. The 216 AD, MCI, and normal participants (observations) and 49 test measures (variables) were submitted to a PCA using the correlation matrix and with varimax rotation (Kaiser, 1958). PCA produced both component loadings and component scores (Chapman & McCrary, 1995). The component loadings (the general underlying structure of the neuropsychological test results) were used to derive interpretations of

the components by relating the test measures to the component structure.

PCA was an important step in our data analysis in three ways. First, it revealed underlying cognitive dimensions implicit in neuropsychological test performance through the component loadings. The loadings related the test measures to the components through each measure's weighted contribution to the component structure (Albert et al., 2007; Carroll, 1993; Chapman et al., 2010a, 2010b; Harman, 1976). Secondly, PCA achieved data reduction by remapping the 49 test measures to a smaller number of component scores via the component structure without being influenced by the group to which the participants belonged. These were important advantages both in organizing similar test measures into components and in reducing the number of variables while retaining the contributions of all the measures. Finally, PCA permits direct and easy computation of the component scores.

Discriminant analysis using PCA component scores

After PCA, the component scores of the 43 MCI participants were retained for discriminant analysis. This analysis developed discriminant functions, based on Bayesian posterior distributions (Ingelfinger, Mosteller, Thibodeau, & Ware, 1983), that predict individuals who will likely convert to AD or likely remain stable. Discriminant analysis provided the posterior probability of group membership for each participant as an integral part of the computation, which adds a key quantitative context when analyzing binary predictions of individuals.

First, a stepwise selection (PROC STEPDISC of SAS) was used to find a subset of the component scores to use as predictors in the analysis. A reduced set of predictors was desirable, and a stepwise discriminant procedure used statistical criteria to determine order of entry. A probability to enter criterion of .20 was used to ensure entry of important variables that best revealed differences between the stable and conversion groups. Then, the subset of selected components was used in a second multivariate procedure (PROC DISCRIM of SAS) to compute linear discriminant functions for classifying individuals. The linear discriminant function was composed of the weights (coefficients) to be used with each of the input variables. The group classification for each participant was dependent on from which group that individual had the smaller generalized squared distance.

Afterward, a jackknifed cross-validation was performed in which the data from each individual were left out when the coefficients used to assign that individual to a group were computed. Thus, a new discriminant function was developed for and tested on every participant individually. Jackknifed cross-validation gives a more realistic estimate of the ability of predictors to separate groups, and bias in classification is eliminated when the same predictors are forced into the equation, as was done here (Hora & Wilcox, 1982; Lachenbruch, 1975; Tabachnick & Fidell, 2001). We chose this method considering sample size limitations and our desire to use

as much data as possible in the development of the discriminant function (Johnson & Wichern, 2002). This solution stability coupled with elimination of bias in classification makes for a better approach given a limited, fixed sample size. Hora and Wilcox (1982) indicated that the one-left-out method is a "superior alternative" to a split-half method, which has an unfortunate effect of reducing the effective sample size.

Required sample size depends upon a number of issues, including expected effect size and number of predictors (Tabachnick & Fidell, 2001). Green (1991) provides a thorough discussion of these concerns and some procedures to determine an appropriate number of cases, including a more complex rule of thumb that takes effect size into account. Expecting the squared multiple correlations to be .2 or greater, we computed a necessary sample size of 42 individuals, given 11 predictor variables. Our sample size of 43 individuals exceeds this rule of thumb. Additionally, the sample size of the smallest group should be larger than the number of predictor variables (Tabachnick & Fidell, 2001), which is true in our sample.

Nevertheless, to substantiate empirically that the sample size is sufficient in the present set of data, a randomized resampling procedure was done to assess baseline discriminant performance to compare with our nonrandomized performance. Classification success after randomizing the data largely depends on capitalizing on chance. We randomized our MCI sample such that each participant was randomly placed in a pseudo-conversion or pseudostable group regardless of his or her clinical diagnosis. The constraint of 14 members in the pseudoconversion group and 29 members in the pseudostable group was maintained. The subset of components best able to discriminate between these pseudogroups was collected by stepwise discriminant analysis (PROC STEPDISC) and was used in classification analysis (PROC DISCRIM), the methods being the same as those used with our nonrandomized (real) data. We randomized our participant groups 50 times, performed stepwise and discriminant analyses on each randomization, and collected the predictive accuracies for the development and cross-validation of the pseudogroups.

Statistical procedures

Statistical analyses were computed with SAS 9.1.3 (SAS Institute Inc., 2002). The primary procedures were the MULTTEST, FACTOR, STEPDISC, and DISCRIM procedures. These have also been applied to neuropsychological tests used to classify AD from normal elderly (Chapman et al., 2010b), as well as applied to brain event-related potentials used to study AD (Chapman et al., 2007) and MCI conversion to AD (Chapman et al., in press). To evaluate the statistical significance of the classification results, we applied Fisher's Exact Test with an alpha level of .05. This test is appropriate because each individual is placed in a cell in a 2×2 contingency table: test classification of conversion or stable by clinical diagnosis of conversion or stable. We corrected for multiple comparisons in the analysis of group mean differences

with Bonferroni adjustments. Also, p values calculated from the Fisher's Exact Tests on classification results were corrected with a Bonferroni adjustment (Shaffer, 1995).

RESULTS

Group means of the neuropsychological test measures

The group mean scores for each of the 49 neuropsychological test measures for the conversion and stable groups are shown in Table 2. Generally the conversion group performed worse than the stable group, particularly on measures of retentive memory. However, none of the 49 test measures had a statistically significant group difference between the stable and conversion groups when adjusted for multiple comparisons.

Neuropsychological components measured with PCA

Using mainly Kaiser's (eigenvalue > 1) criterion (Kaiser, 1960) as a guideline, we obtained 13 distinct, orthogonal, and interpretable components in the component structure. These 13 components accounted for 77% of the total variance of the data and included a General Episodic Memory component, a Generative Fluency component, a Speeded Executive Function component, a Mood/Activities of Daily Living component, and other components representative of learning and recognition memory. These neuropsychological components have been shown to have strong discriminatory power in differentiating AD from normal aging (Chapman et al., 2010b).

Predicting conversion to AD using PCA component scores

Of the 13 PCA components, we retained 11 component scores for each of the 43 MCI participants for discriminant

analysis. The last 2 components (which accounted for little variation in the component solution) were not used in order to maintain a roughly 4:1 ratio between participants and predictor variables entering the stepwise discriminant procedure. These first 11 components accounted for 72% of the total variance of the data. The group mean component scores appear in Table 3. While group differences between the component scores could be examined, it should be noted that the stepwise discriminant procedure classifies individuals and takes the correlations between the components into account when making its determinations; therefore, significant differences between the group means might not necessarily signify strong discriminatory power at the individual level given the rest of the components in the set used. From the 11 component scores entering the stepwise discriminant procedure, 6 component scores were selected as those that had the best discriminability between the conversion and stable groups (Table 4). These 6 components were weighted and combined in linear discriminant functions to classify each individual as a member of either the conversion or stable group (Table 5). Little credence should be placed in the meaning of the particular coefficients found for the sample unless all important variables are known to be included in the analysis or are known to be uncorrelated with the variables already included (Ahlgren, 1986). We show them here because they were used in the discriminant functions as the weights to be multiplied by the neuropsychological component scores of an individual and as a set were assessed to have favorable, statistically significant classification success. Furthermore, they may be used as a tool in analyzing additional data.

The discriminant functions performed well in the development set: A total of 36 of the 43 participants were correctly classified, resulting in 83.7% prediction accuracy: Fisher's Exact Test, $\chi^2(1, N = 43) = 18.2, p < .0001$. Of the 14 members of the conversion group, 2 were incorrectly predicted to have remained stable, resulting in a sensitivity of .86 and a positive predictive value of .71. Additionally, 24 of the 29 members of the stable group were correctly predicted, resulting in a specificity

TABLE 3
Group mean component scores for the conversion to AD and stable groups

Components	Component scores	
	Conversion ($n = 14$)	Stable ($n = 29$)
1: Episodic Memory	-0.77 (1.06)	-0.07 (0.69)
2: Speeded Executive Function	-0.19 (1.10)	0.17 (0.93)
3: Generative Fluency	-0.02 (1.14)	-0.03 (1.00)
4: Recognition Memory (False Positives)	0.14 (1.48)	-0.25 (0.74)
5: Immediate Attention Span	0.10 (0.55)	0.19 (1.14)
6: Recognition Memory (True Positives)	0.32 (0.72)	-0.20 (0.84)
7: Visuospatial Episodic Memory	-0.41 (1.04)	0.14 (1.21)
8: Visuo-construction Abilities	0.07 (0.86)	0.17 (0.65)
9: Visuospatial Learning	0.07 (0.77)	0.05 (1.13)
10: Mood/Activities of Daily Living	-0.36 (1.24)	-0.24 (1.15)
11: Speed in Visuospatial Memory	0.07 (0.82)	-0.10 (0.56)

Note. Standard deviations in parentheses. AD = Alzheimer's disease.

TABLE 4
PCA loadings for the six components selected by the
stepwise discriminant procedure

Component	PCA loading
<i>1: Episodic Memory</i>	
LM-II B Recall Score	.83
LM-II A Recall Score	.82
LM-II Recognition Score	.80
LM-I B2 Recall Score	.80
LM-I A Recall Score	.77
LM-I B1 Recall Score	.77
HVLT Trial 3 Score	.75
HVLT Discrimination Index	.75
LM-II Percent Retention	.74
HVLT True Positives	.70
HVLT Trial 1 Score	.69
HVLT Trial 2 Score	.69
BVMT-R Delayed Recall	.65
BVMT-R Trial 3 Score	.60
BVMT-R Percent Retention	.54
BVMT-R Learning Slope	.52
HVLT Related False Positives	-.49
Rey Delayed Recall Score	.45
Animal Naming	.44
MMSE Score	.44
Rey Immediate Recall Score	.43
<i>2: Speeded Executive Function</i>	
Stroop Color Score	.85
Stroop Word Score	.75
Stroop Color Word Score	.72
Trail Making Test A Speed	.68
Trail Making Test B Speed	.60
<i>4: Recognition Memory (False Positives)</i>	
BVMT-R False Alarms	.74
HVLT Unrelated False Positives	.74
HVLT Related False Positives	.65
HVLT Discrimination Index	-.46
<i>6: Recognition Memory (True Positives)</i>	
BVMT Hits	.80
BVMT Discrimination Index	.74
<i>11: Speed in Visuospatial Memory</i>	
Rey Delayed Recall Speed	.76
Rey Immediate Recall Speed	.75
<i>7: Visuospatial Episodic Memory</i>	
Rey Delayed Recall Score	.76
Rey Immediate Recall Score	.75

Note. The six neuropsychological components are in the order they were selected by the stepwise discriminant procedure. Test measures shown had a loading for the component above an arbitrary threshold of .43 that was selected to highlight more salient loadings. This analysis included 55 early-stage AD, 78 MCI, 78 controls, and 5 AAMI. The table shows the PCA varimax rotated component pattern.

MCI = mild cognitive impairment. AD = Alzheimer's disease. AAMI = age-associated memory impairment. PCA = principal component analysis. Rey = Rey-Osterrieth Complex Figure; LM-I/LM-II = Logical Memory I/Logical Memory II; BVMT-R = Brief Visuospatial Memory Test-Revised; HVLT = Hopkins Verbal Learning Test; MMSE = Mini-Mental State Examination.

TABLE 5

Linear discriminant function coefficients for the conversion to AD and stable groups

Variable (component)	Conversion	Stable
Constant	-1.37	-1.15
1: Episodic Memory	-2.42	0.24
2: Speeded Executive Function	-1.57	0.41
4: Recognition Memory (False Positives)	0.74	-0.35
6: Recognition Memory (True Positives)	0.65	-0.47
11: Speed in Visuospatial Memory	0.93	-0.45
7: Visuospatial Episodic Memory	-0.49	0.15

Note. Discriminant coefficients shown are for the six neuropsychological components selected by the stepwise discriminant procedure (in the order they were selected). AD = Alzheimer's disease.

of .83 and a negative predictive value of .92. The likelihood ratios for a positive and negative test were: LR+ = 5.06 and LR- = -0.17.

Examination of posterior probabilities

To gauge the confidence in these predictions, we analyzed the posterior probabilities of group membership computed by the discriminant procedure for each individual. These probabilities added quantitative context to the binary predictive decision by supplying measures of likelihood that the group into which the participant was placed by our multivariate method was the correct group. The obtained prediction accuracy for each posterior probability bin was plotted (Figure 1). Participants were placed in posterior probability bins by their probability of belonging to the group in which the discriminant function placed them (and placement was determined by the group for which the posterior probability was greater than .50). First, the obtained prediction accuracy dramatically rose with posterior probability. Second, most (35) of the 43 (81%) participants lie in the highly confident probability bins (.70-1.00, where the prediction accuracy curve reached its highest level). Only 8 participants were located in the least confident bins (.50-.69) where the obtained prediction accuracy was near 50% chance.

Cross-validation

The cross-validation provided good results: A total of 34 of the 43 (79.0%) individuals were correctly placed in either the conversion or the stable group: Fisher's Exact Test, $\chi^2(1, N = 43) = 13.2, p < .001$. In the conversion group, 11 participants were correctly classified, resulting in a sensitivity of .79. Likewise, in the stable group, 23 participants were correctly classified, resulting in a specificity of .79.

Randomized resampling analysis

Given the modest sample size (43), there might be some concern that we have arrived at seemingly impressive results solely by chance variation in the sample (Ahlgren, 1986). However, since only 11 variables entered the

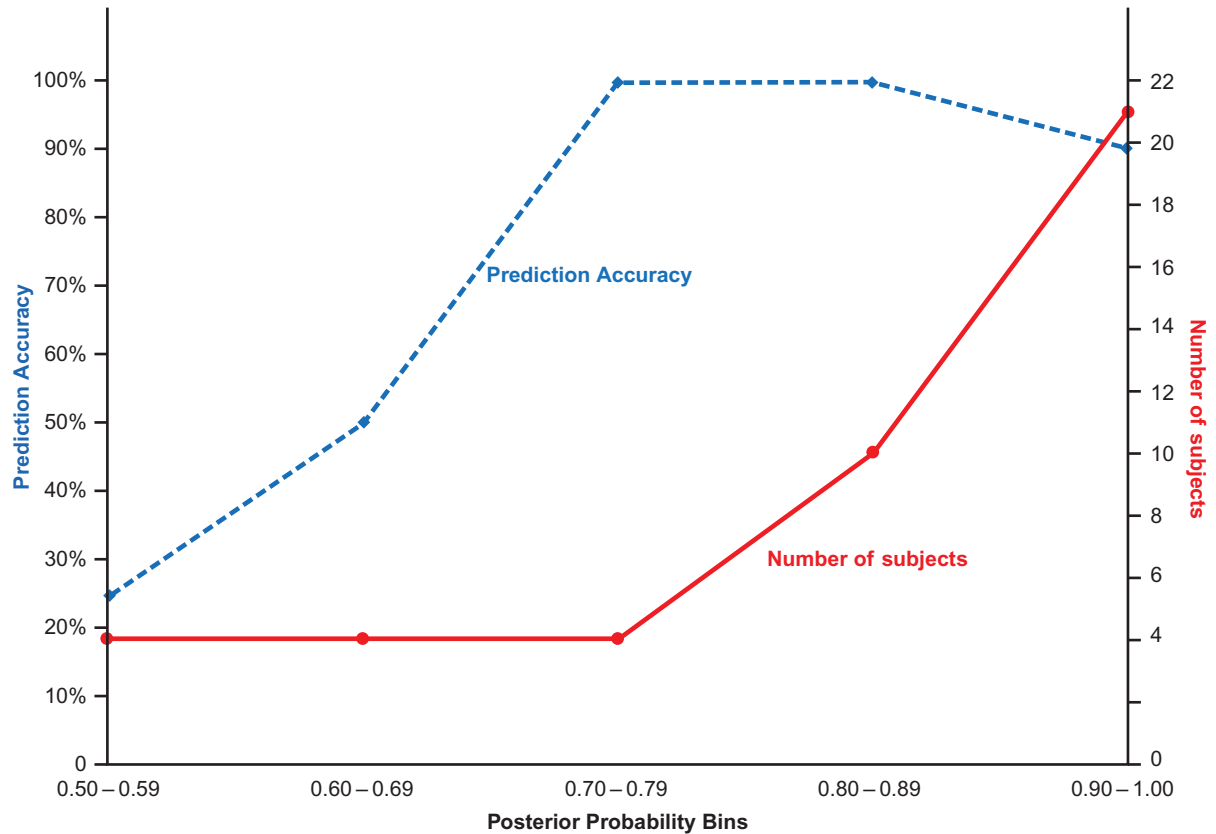


Figure 1. The prediction accuracy for, and number of participants in, each posterior probability bin based on the development set. Prediction accuracy is the percentage of MCI (mild cognitive impairment) individuals correctly predicted to convert to AD (Alzheimer's disease) or to remain stable. Each of the 43 MCI individuals was placed in a bin by the probability of group membership calculated through the discriminant function (placement was determined by the probability greater than .50). To view a color version of this figure, please see the online issue of the Journal.

stepwise discriminant procedure, the ratio of participants (43) to predictor variables was approximately 4:1. Also, our number of participants exceeds the number suggested by analysis of expected effect size (see Method section, and Green, 1991).

Still, to assess empirically that our predictive results were not capitalizing on chance, we measured chance performance with these data randomized. We discriminated randomized, resampled pseudogroups of our MCI participants and determined whether our nonrandomized results were statistical outliers. We randomized our participant groups 50 times, performed stepwise and discriminant analyses on each randomization, and collected the predictive accuracies for the development and cross-validation of the pseudogroups. The mean (with standard deviations in parentheses) percentage accuracies for the development and cross-validation analyses of the pseudogroups were 73.0 (5.7) and 61.3 (9.0), respectively. Our nonrandomized (real) results reported above (83.7% accuracy for the development and 79.0% accuracy for the cross-validation) are nearly two standard deviations above the mean accuracies calculated from the randomized pseudogroups for each analysis. Despite the modest sample size, our real results were statistical outliers ($p < .05$) from the mean predictive accuracies

that chance in the pseudogroups could produce. It should be noted that, because of the constraints on the sample size in each pseudogroup, the mean predictive accuracies were higher than 50% chance (as with each randomization, 15 stable MCI individuals must be placed in the pseudostable group). Additionally, by chance there could have been randomizations where nearly all of the participants were placed in their correct groups. Nonetheless, the prediction accuracies for our nonrandomized (real) groups were substantially higher than the averages of the pseudogroups. This finding indicated that the sample size was large enough for these data and that our results were not due simply to capitalizing on chance.

DISCUSSION

Neuropsychological tests are sensitive to the cognitive deficits of MCI. We have examined whether weighted combinations of neuropsychological test measures derived from a battery of commonly used tests can predict conversion in MCI to AD at the individual level. The prediction results using PCA component scores combined by discriminant analysis were accurate with strong sensitivity and specificity. We now explore the

predictive success of the neuropsychological tests as obtained through our combinatory methods and what these tests may reveal about cognitive decline in MCI.

Predictive power of PCA component scores

We studied whether or not a multivariate approach, where the correlations among all the test measures in the battery were taken into account in the resultant component structure before discrimination, could generate strong, statistically significant predictive results. PCA reorganized the neuropsychological test measures into 11 more interpretable components that were implicit in the test data. An advantage of PCA is that it allows contribution of all cognitive tests in the neuropsychological battery as measured by each test's loading on each component. Bäckman et al. (2005) stated in a meta-analysis that few studies tapped all of the eight cognitive domains suggested by the NINCDS-ADRDA criteria in their analyses of conversion from MCI to AD using neuropsychological tests. Through PCA, we were able to represent all eight cognitive domains in our composite neuropsychological component scores.

The multivariate method discussed in this paper has essentially two layers of weighting: (a) the weighting applied by PCA in reorganizing the neuropsychological test measures via the correlations among them into the component structure (independent of knowledge about the differing groups), and (b) the differential weighting for the component scores added by the discriminant analysis in computing discriminant coefficients that are best able to differentiate between the conversion and stable groups. This method achieved 84% accuracy, a sensitivity of .86, and a specificity of .83 by simultaneously taking into account an individual's performance in all the cognitive domains represented by the battery. The components become a new metric that more parsimoniously represents neuropsychological test performance than the many measures of the original battery, and each participant's component scores, weighted through formal, data-driven methods, can place his or her performance along the dimensions of the metric as closer to (or farther from) conversion to AD.

Additionally, the posterior probabilities of group membership that accompanied the binary prediction for each participant indicated the confidence of the prediction (Figure 1). We have rarely seen these measures used or discussed in the literature on predicting MCI progression in individuals. Featuring the posterior probabilities is important for several reasons. It allows a determination of which participants may be "too close to call" (e.g., the determination of a cut point of a group membership for classification). These participants, because of their low posterior probabilities (near chance), could be labeled indeterminate in their diagnosis.

One could consider the prediction outcomes to be binary (the participant either converted to AD or did not). However, the posterior probabilities might be used to measure disease progression (how similar or dissimilar an MCI patient is to other patients who have converted

to AD). Evaluating the posterior probabilities may allow the physician or researcher both to identify the probable predictions in a group of individuals and to measure the stage of progression for each individual. Posterior probabilities could aid a physician in determining the appropriateness of treatment, and it could benefit researchers when selecting project participants.

There are examples in the literature of analyzing a neuropsychological battery for predicting conversion from MCI to AD. When memory alone was used as a predictor for conversion to AD, sensitivity and specificity were lower (Lekeu et al., 2010; Perri et al., 2007; Tierney et al., 2005). We have shown that the addition of other domains, particularly executive function, produces high predictive accuracy, a finding echoed by other studies (Marcos et al., 2006; Rozzini et al., 2008; Tabert et al., 2006). Removing the speeded executive function component (Component 2) from our discriminant analysis caused a 14% drop in cross-validation predictive accuracy (27% drop in sensitivity and 10% drop in specificity). Tabert et al. used regression analysis on both memory and executive function measures and reached a classification accuracy of 86%. Likewise, Marcos et al. examined the predictive power of another battery of tests, which included similar neuropsychological measures. In their study, conversion to AD or stability was correctly identified in approximately 82% of their MCI sample using primarily the Global Cognitive Subscale of the Cambridge Mental Disorders (CAMCOG) and multiple regression analysis. Their results are roughly similar to our developmental findings using component scores, though an examination of their participant demographics suggested that their MCI individuals had less exposure to education (<10 years) than ours (Table 1). In these studies and in our own, the combination of measures from other cognitive domains with memory measures in a formal, multivariate manner improves predictive accuracy. However, none of these other studies provided validation analyses.

Our multivariate method may be generalizable and could be implemented in other settings. This would be easiest to do for a new participant if the tests administered are the same as those used in the development of the component structure. However, it might be possible to use different tests if their loadings on the same components could be reasonably estimated. This is an important point, considering that different clinics and research centers might wish to employ their own battery of tests. An aid to doing this might be to calibrate the new measures in combination with marker variables that belong to some of the tests we used in this study that have strong loadings. For more information on this topic and a flow diagram depicting the application of this methodology to new individuals, see Chapman et al. (2010a). Perhaps more important than the particular tests are the cognitive dimensions represented by those tests. Our multivariate methodology can be applied to different neuropsychological batteries that represent the same or similar cognitive dimensions.

Additionally, after the component structure has been developed, it may be possible to reduce the number of tests administered and to achieve essentially the same

results. The 49 test measures we used came from 17 neuropsychological tests (Table 2), which were resolved into 11 PCA components (Table 3). Because in the discriminant analyses only 6 of the 11 component scores were used for prediction, some reduction in the number of neuropsychological tests may be possible without greatly harming the results of the analysis. This can be estimated by studying the PCA loadings of the test measures on the 6 components that were used in the discriminant analyses (Table 4). Test measures with high loadings on a component play the largest role in computing its component scores. Perhaps dropping test measures with low loadings might incur only small changes in the component scores. For example, since each of the 6 components selected by the stepwise discriminant procedure had one or more test measures with loadings as high as .74, perhaps test measures with loadings below .43 might be dropped with minor effects on the component scores.

A caveat is that it may be risky to completely remove measures with low loadings, and we have not studied the effects of doing so in these data. Here we are only proposing this as an idea for those who wish to use a smaller neuropsychological battery.

Neuropsychological tests with good predictive power

Our results indicate that measures of memory strongly herald the decline of cognition from MCI to AD, a finding in concordance with much of the literature (e.g., Bäckman et al., 2005; Lekeu et al., 2010; Tabert et al., 2006). The PCA components selected by the stepwise discriminant procedure showed the General Memory component as the first chosen for its discriminatory power between the conversion and stable groups (Table 5). It was not obvious from simple examination of mean differences between the conversion and stable groups that measures of nonmemory cognitive domains might provide predictive power (Table 2). However, our multivariate method revealed that these cognitive processes may be involved in the decline from MCI to AD. These contributions might not be as prominent as the increased impairment evident in memory measures, but they still added discriminatory power to the discriminant functions. Our finding of the utility of executive functioning and perceptual speed (Component 2) supports Bäckman et al. (2005) that episodic memory alone may not be the best predictor of conversion to AD in MCI.

One would naturally expect memory measures to present strong predictive power of conversion from MCI to dementia given the nature of AD. The tests with the highest loadings on the General Memory component generally reflect delayed recall (the Logical Memory II), which has typically shown greater power in predicting conversion to AD (Bäckman et al., 2005; Perri et al., 2007). Recognition memory is impacted to a lesser extent (Bäckman et al., 2005), and in our study, the Recognition Memory (True Positives) component was selected later in the stepwise process. Conversion to AD in amnesic MCI

is marked by increased difficulties with memory (Petersen, 2004). It is worth noting that employing the PCA allowed many memory measures to contribute to the composite component score along this cognitive dimension (Table 4), which can enhance the discriminatory power by including different types of memory (episodic, verbal, and visuospatial) in the component scores without threatening the degrees of freedom in subsequent discriminant analyses.

The second most powerful predictor of conversion to AD in MCI was the Speeded Executive Function Component, which featured high loadings from the Stroop and Trail Making tests. Impairments in quickly switching attention as well as disinhibition may indicate progression toward or conversion to AD. We found this facet of impairment again in the selection of PCA Component 4, the Recognition Memory (False Positives) component. Those MCI who convert to AD seem to have difficulty inhibiting an incorrect response, a finding echoed by others (e.g., Marcos et al., 2006). This component provided an important enhancement of predictive power to the discriminant function as removing it from the analysis caused a sizable drop in prediction accuracy (14%) in the cross-validation.

In addition to emerging difficulties with executive function and response inhibition, individuals with amnesic MCI who convert to AD may also have impairments in processing visuospatial memory (the Visuospatial Episodic Memory component, Component 7). Perhaps visuospatial episodic memory retrieval may be more seriously or quickly impacted as the disease progresses. Component 11, the Speed in Processing Visuospatial Memory component, also added discriminatory power. This measure was derived from the speed at which the participants were capable of reproducing the Rey Complex Figure from memory in the immediate and delayed recall tests. Marcos et al. (2006) suggested that as impairment toward AD progresses, MCI participants exhibit increased difficulty during complex or demanding visuospatial tasks. The notion that visuospatial memory is impacted at an earlier stage or to a greater extent in those MCI individuals who later develop AD could be consistent with the known pathology of posterior brain regions in visuospatial variants of AD (Mapstone et al., 2003).

The development of impairment in a secondary cognitive domain is a diagnostic hallmark of conversion from MCI to AD (Petersen, 2004). We have seen through our study of this neuropsychological battery for this group of amnesic MCI individuals that failures in executive function (specifically attention switching and response inhibition) may represent impairment of the secondary cognitive domain. It would be interesting to relate these results to individuals with multiple-domain MCI or with single-domain MCI that presents without primary memory impairment. Perhaps these individuals may already be impaired in executive function or visuospatial memory and perception (Mapstone et al., 2003) and may develop memory difficulties as a "secondary" domain in a temporal sense. This concept requires further study to determine whether our component scores are applicable to the wide range of cognitive and behavioral heterogeneity of MCI that often presents itself in clinics. Additionally,

our study warrants further validation with a greater number of MCI participants. Although others have indicated that the time between the initial diagnosis and the time of conversion is not a strong predictor (Jelic et al., 2000), our follow-up periods were relatively short and should be studied more extensively. Finally, we believe our battery of neuropsychological tests (found to have 13 dimensions) was representative of the myriad cognitive domains affected by AD. Still, it is possible that other tests that we did not study (for example, measures of olfactory discrimination) may offer better assessment of the decline from MCI to AD, and these tests should be analyzed in such a multivariate way to measure their contributions.

CONCLUSION

This study was intended to test whether or not a multivariate, composite marker of neuropsychological test performance developed with PCA can predict conversion to AD in MCI individuals with strong accuracy, sensitivity, and specificity. Multivariate weighting, achieved sequentially first through PCA and then through discriminant analysis, produced high, statistically significant accuracy, sensitivity, and specificity in predicting which MCI individuals will be members of a stable or conversion group. In addition, the posterior probabilities provided by discriminant analysis added an important measure of confidence to the predictions, allowing the identification of those participants whose predictive diagnosis might be “too close to call.” Our results were tested for reliability by a statistically significant cross-validation and a randomized resampling method. This multivariate approach warrants further study in comparison to and in combination with other biological, neuropsychological, and behavioral methods of identifying the progression of cognitive impairment.

Original manuscript received 10 March 2010

Revised manuscript accepted 27 May 2010

First published online 13 August 2010

REFERENCES

- Ahlgren, A. (1986). Multivariate analysis. *Science*, 234, 530–531.
- Albert, M. S., Blacker, D., Moss, M. B., Tanzi, R., & McArdle, J. J. (2007). Longitudinal change in cognitive performance among individuals with mild cognitive impairment. *Neuropsychology*, 21, 158–169.
- Albert, M. S., Moss, M. B., Tanzi, R., & Jones, K. (2001). Preclinical prediction of AD using neuropsychological tests. *Journal of International Neuropsychological Society*, 7(5), 631–639.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of the mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Babins, L., Slater, M., Whitehead, V., & Chertkow, H. (2008). Can an 18-point clock-drawing scoring system predict dementia in elderly individuals with mild cognitive impairment? *Journal of Clinical and Experimental Neuropsychology*, 30, 173–186.
- Bäckman, L., Jones, S., Berger, A. K., Laukka, E. J., & Small, B. J. (2005). Cognitive impairment in preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology*, 19, 520–531.
- Bateman, R. J., Wen, G., Morris, J. C., & Holtzman, D. M. (2007). Fluctuations of CSF amyloid-beta levels: Implications for a diagnostic and therapeutic biomarker. *Neurology*, 68, 666–669.
- Benedict, R. H. B. (1997). *Brief Visuospatial Memory Test—Revised*. Lutz, FL: Psychological Assessment Resources.
- Benedict, R. H. B., & Groninger, L. (1995). Preliminary standardization and validation of a new visuospatial memory test with six alternate forms. *The Clinical Neuropsychologist*, 9, 11–16.
- Benton, A. L., & Hamsher, K. S. (1976). *Multilingual Aphasia Examination*. Iowa City, IA: University of Iowa.
- Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in cerebral grey matter of elderly subjects. *British Journal of Psychiatry*, 114, 797–811.
- Bondi, M. W., Monsch, A. U., Galasko, D., Butters, N., Salmon, D. P., & Delis, D. C. (1994). Preclinical cognitive markers of dementia of the Alzheimer type. *Neuropsychology*, 8, 374–384.
- Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a memory test with six equivalent forms. *The Clinical Neuropsychologist*, 5, 125–142.
- Brayne, C. (2007). The elephant in the room—healthy brains in later life, epidemiology and public health. *Nature Reviews Neuroscience*, 8, 233–239.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Chapman, R. M., Mapstone, M., McCrary, J. W., Gardner, M. N., Bachus, L. E., DeGrush, E., et al. (2010a). Cognitive dimensions in Alzheimer's disease, mild cognitive impairment, and normal elderly: Developing a common metric. *The Open Geriatric Medicine Journal*, 3, 1–10. doi:10.2174/1874827901003010001
- Chapman, R. M., Mapstone, M., Porsteinsson, A. P., Gardner, M. N., McCrary, J. W., DeGrush, E., et al. (2010b). Diagnosis of Alzheimer's disease using neuropsychological testing improved by multivariate analyses. *Journal of Clinical and Experimental Neuropsychology*. Advanced online publication. doi: 10.1080/13803390903540315
- Chapman, R. M., & McCrary, J. W. (1995). EP component identification and measurement by principal components analysis. *Brain and Cognition*, 27, 288–310.
- Chapman, R. M., McCrary, J. W., Gardner, M. N., Sandoval, T. C., Guillily, M. D., Reilly, L. A., et al. (in press). Brain ERP components predict which individuals progress to Alzheimer's disease and which do not. *Neurobiology of Aging*. doi:10.1016/j.neurobiolaging.2009.11.010
- Chapman, R. M., Nowlis, G. H., McCrary, J. W., Chapman, J. A., Sandoval, T. C., Guillily, M. D., et al. (2007). Brain event-related potentials: Diagnosing early-stage Alzheimer's disease. *Neurobiology of Aging*, 28, 194–201. doi:10.1016/j.neurobiolaging.2005.12.008
- Crook, T. H., Bartus, R. T., Ferris, S. H., Whitehouse, P., Cohen, G. D., & Gershon, S. (1986). Age-associated memory impairment: Proposed diagnostic criteria and measures of clinical change—report of a National Institute of Mental Health work group. *Developmental Neuropsychology*, 2, 261–276.
- Fabrigar, L. R., MacCullum, R. C., Wegener, D. T., & Stahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state.” A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.
- Golden, C. J. (1978). *Stroop Color and Word Test: A manual for clinical and experimental uses*. Chicago: Stoelting.
- Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, 26(3), 499–510.
- Grober, E., & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence

- in the elderly. *Journal of Clinical and Experimental Neuropsychology*, 13, 933–949.
- Harman, H. H. (1976). *Modern factor analysis*. Chicago: University of Chicago Press.
- Hickie, C., & Snowdon, J. (1987). Depression scales for the elderly: GDS, Gilleard, Zung. *Clinical Gerontologist*, 6, 51–53.
- Hora, S. C., & Wilcox, J. B. (1982). Estimation of error rates in several-population discriminant analysis. *Journal of Market Research*, 19, 57–61.
- Ingelfinger, J. A., Mosteller, F., Thibodeau, L. A., & Ware, J. H. (1983). *Biostatistics in Clinical Medicine*. New York: Macmillan Publishing Company.
- Jelic, V., Johansson, S. E., Almkvist, O., Shigeta, M., Julin, P., Nordberg, A., et al. (2000). Quantitative electroencephalography in mild cognitive impairment: Longitudinal changes and possible prediction of Alzheimer's disease. *Neurobiology of Aging*, 21, 533–540.
- John, E. R., Easton, P., Prichep, L. S., & Friedman, J. (1993). Standardized varimax descriptors of event-related potentials: Basic considerations. *Brain Topography*, 6, 143–162.
- Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis* (5th ed.). New York: Prentice-Hall.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3), 187–200.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141–151.
- Kaplan, E. F., Goodglass, H., & Weintraub, S. (1978). *The Boston Naming Test*. Philadelphia: Lea & Febiger.
- Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D. P., et al. (2004). Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Annals of Neurology*, 55, 306–319.
- Lachenbruch, P. A. (1975). *Discriminant analysis*. New York: Hafner Press.
- Lekun, F., Magis, D., Marique, P., Delbeuck, X., Bechet, S., Guillaume, B., et al. (2010). The California Verbal Learning Test and other standard clinical neuropsychological tests to predict conversion from mild memory impairment to dementia. *Journal of Clinical and Experimental Neuropsychology*, 32(2), 164–173. doi:10.1080/13803390902889606
- Mack, W. J., Freed, D. M., Williams, B. W., & Henderson, V. W. (1992). Boston Naming Test: Shortened versions for use in Alzheimer's disease. *Journal of Gerontology: Psychological Sciences*, 47, 154–158.
- Maioli, F., Coveri, M., Pagni, P., Chiandetti, C., Marchetti, C., Ciarrocchi, R., et al. (2007). Conversion of mild cognitive impairment to dementia in elderly subjects: A preliminary study in a memory and cognitive disorder unit. *Archives of Gerontology and Geriatrics Supplement*, 1, 233–241.
- Mapstone, M., Steffenella, T., & Duffy, C. (2003). A visuospatial variant of mild cognitive impairment: Getting lost between aging and Alzheimer's disease. *Neurology*, 60, 802–808.
- Marcos, A., Gil, P., Barabash, A., Rodriguez, R., Encinas, M., Fernández, C., et al. (2006). Neuropsychological markers of progression from mild cognitive impairment to Alzheimer's disease. *American Journal of Alzheimer's Disease & Other Dementias*, 21(3), 189–196.
- McKhann, G., Drachman, D., Folstein, M. F., Katzman, R., Price, D., & Stadlan, E. M. (1984). 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*, 34, 939–944.
- Money, J. A. (1976). *A Standardized Road Map of Directional Sense, manual*. San Rafael, CA: Academic Therapy Publications.
- Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., et al. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, 39(9), 1159–1165.
- Osterrieth, P. A. (1944). Le test de copie d'une figure complexe [The test for copying a complex figure]. *Archives de Psychologie*, 30, 206–356.
- Papaliagkas, V. T., Anogianakis, G., Tsolaki, M. N., Koliakos, G., & Kimiskidis, V. K. (2009). Prediction of conversion from mild cognitive impairment to Alzheimer's disease by CSF cytochrome c levels and N200 latency. *Current Alzheimer Research*, 6, 279–284.
- Perri, R., Serra, L., Carlesimo, G. A., & Caltagirone, C. (2007). Amnesic mild cognitive impairment: Difference of memory profile in subjects who converted or did not convert to Alzheimer's disease. *Neuropsychology*, 21, 549–558.
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, 256, 183–194.
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, 56, 303–308.
- Petersen, R. C., Stevens, J. C., Ganguli, M., Tangalos, E. G., Cummings, J. L., & DeKosky, S. T. (2001). Practice parameters: Early detection of dementia: Mild cognitive impairment (an evidence-based review). Report on the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*, 56, 1133–1142.
- Reitan, R. M. (1958). Validity of the Trail Making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, 271–276.
- Rey, A. (1941). L'examen Psychologique dans les cas d'encéphalopathie traumatique [Psychological testing in the case of traumatic encephalopathy]. *Archives de Psychologie*, 28, 286–340.
- Rey, A. (1964). *L'Examen Clinique en Psychologie* [Clinical testing in psychology]. Paris: Presses Universitaires de France.
- Rozzini, L., Chilovi Vicini, B., Bertolotti, E., Conti, M., Delrio, I., Trabucchi, M., et al. (2008). The importance of Alzheimer Disease Assessment Scale–Cognitive Part in predicting progress for amnesic mild cognitive impairment to Alzheimer disease. *Journal of Geriatric Psychiatry and Neurology*, 21(4), 261–267.
- SAS Institute Inc. (2002). SAS OnlineDoc 9.1.3. Retrieved January 1, 2007, from <http://support.sas.com/onlinedoc/913/docMainpage.jsp>
- Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Reviews Psychology*, 46, 561–584.
- Simonsen, A. H., McGuire, J., Hansson, O., Zetterberg, H., Podust, V. N., Davies, H. A., et al. (2007). Novel panel of cerebrospinal fluid biomarkers for the prediction of progression to Alzheimer dementia in patients with mild cognitive impairment. *Archives of Neurology*, 64, 366–370.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn and Bacon.
- Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., et al. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Archives of General Psychiatry*, 23, 916–924.
- Taylor, E. M. (1959). *Psychological appraisal of children with cerebral deficits*. Cambridge, MA: Harvard University Press.
- Tierney, M. C., Yao, C., Kiss, A., & McDowell, I. (2005). Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. *Neurology*, 64, 1853–1859.
- Tuokko, H., Hadjistavropoulos, T., Miller, J. A., & Beattie, B. L. (1992). The clock test: A sensitive measure to differentiate normal elderly from those with Alzheimer's disease. *Journal of the American Geriatrics Society*, 40, 579–584.
- Wechsler, D. (1945). A standardized memory scale for clinical use. *The Journal of Psychology*, 19, 87–95.
- Wechsler, D. (1987). *Wechsler Memory Scale–Revised manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale–III (WAIS–III) manual*. New York: The Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale III: Administration and scoring manual*. Orlando, FL: Psychological Corporation.
- Yesevage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17(1), 37–49.