Neuropsychological Measures in Normal Individuals That Predict Subsequent Cognitive Decline

Deborah Blacker, MD, ScD; Hang Lee, PhD; Alona Muzikansky, MS; Emily C. Martin, PhD; Rudolph Tanzi, PhD; John J. McArdle, PhD; Mark Moss, PhD; Marilyn Albert, PhD

Objective: To examine neuropsychological measures among normal individuals that predict time to subsequent cognitive decline.

Design: Cognitive performance, as measured by 6 neuropsychological tests, was examined at baseline. Participants were followed up for approximately 5 years. Cox proportional hazards models were used to evaluate the neuropsychological measures at baseline that predicted time to progression from normal cognition to mild impairment. Comparable data also examined time to progression from mild impairment to a diagnosis of Alzheimer disease.

Setting: Community volunteer-based sample examined at a medical institution.

Participants: One hundred and seven individuals who were cognitively normal and 235 individuals with mild cognitive impairment at baseline.

Main Outcome Measures: Time to progression from normal cognition to mild impairment and time to progression from mild impairment to a diagnosis of Alzheimer disease.

Results: The risk of progressing from normal to mild impairment was considerably greater among those with lower scores on tests of episodic memory (eg, hazard ratio for a 1-SD decrease in the California Verbal Learning Test, 0.55; P<.001). Normal individuals who carried at least 1 copy of the apolipoprotein E ε2 allele were less likely to develop cognitive impairments over time than individuals with no ε2 allele (hazard ratio for presence of allele, 0.13; P = .006). Measures of both episodic memory and executive function were significant predictors of time to progression from mild impairment to a clinical diagnosis of Alzheimer disease (eg, hazard ratio for a 1-SD decrease in California Verbal Learning Test score, 0.67; P = .005; hazard ratio for a 1-SD increase in the time to complete part B of the Trail Making test, 1.40; P = .007). Among individuals with mild impairments, the apolipoprotein E ε4 allele increased risk for Alzheimer disease in a dose-dependent manner; however, this effect was not significant within the context of multivariable models.

Conclusions: Episodic memory performance among normal individuals predicts time to progression to mild impairment while apolipoprotein E ε2 status is associated with lower risk of cognitive decline among normal individuals. Tests of both episodic memory and executive function are predictors of time to progression from mild impairment to a clinical diagnosis of Alzheimer disease.

Arch Neurol. 2007;64:862-871

There is increasing evidence that the pathology of Alzheimer disease (AD) can take many years, if not decades, to evolve. This suggests that some individuals who appear normal (ie, functionally asymptomatic) have gradually accumulating pathology. Recent reports based on autopsies of well-characterized normal individuals corroborate this hypothesis. For example, in one study of 60 normal individuals who had come to autopsy, approximately 60% had a low “likelihood” of AD based on criteria from the National Institute on Aging–Reagan Institute while 40% had an intermediate or high likelihood of AD. These findings suggest that it may be possible to identify characteristic changes using existing methodologies among normal individuals with a high likelihood of substantial underlying AD pathology during life. One potential approach is to determine whether neuropsychological performance among normal individuals can predict the time to develop mild degrees of cognitive impairment that are suggestive of the incipient stages of AD.

There are numerous studies of neuropsychological performance among individuals in the prodromal phase of AD, commonly referred to as mild cognitive impairment (MCI). Most have focused on which cognitive tests or domains are pre-
predictive of subsequent diagnosis of AD and have found that the cognitive alterations are selective. There is widespread consensus that performance on tests of episodic memory are significantly lower among mildly impaired but nondemented individuals who are destined to meet criteria for AD over time. 5-19 There is also consensus that at least 1 other domain may be altered among such individuals. Many studies have reported that executive function performance is altered during this prodromal phase of AD, 15,16,20 but other neuropsychological alterations have been reported as well. 21 If cognitive changes exist among normal individuals that predict time to develop mild impairment, it seems reasonable to hypothesize that they would be in the area of episodic memory because changes in this domain are clearly evident by the time individuals show mild deficits in daily life.

We therefore chose to examine neuropsychological performance among a group of 107 normal individuals who were followed up annually for an average of 4.7 years. Approximately half of them (n=54) developed mild impairments in daily life while the rest (n=53) remained normal. To determine whether the cognitive deficits observed among normal controls who subsequently develop mild impairment are likely to be part of the evolution of AD, we examined a second group of individuals who had been followed up for a comparable period of time. This group consisted of 235 mildly impaired but nondemented individuals who had been followed up for an average of 5.6 years using similar procedures. Of these, 69 subsequently progressed to a diagnosis of AD.

Cox proportional hazards models were used to evaluate time to progress from one level of severity to another as a function of neuropsychological test performance. The impact of the apolipoprotein E (APOE) gene on time to progression was also examined.

SELECTION OF PARTICIPANTS

A total of 342 individuals were included in the analyses. All of them were participants in an ongoing longitudinal study of the evolution of AD. They had been recruited through the print media (rather than from a clinic or other medical referral source) with advertisements indicating that a research study was seeking individuals both with and without memory difficulty.

Volunteers underwent a multistage screening procedure. The details of the screening procedures have been described elsewhere. 19 Briefly, to be included in the study, participants had to be aged 65 years and older (with the exception of 17 individuals aged 57-64 y); have an informant who could provide information about their daily function; be free of significant underlying medical, neurologic, or psychiatric illness; and be willing to participate in the study procedures. In addition, individuals with evidence of major vascular risk factors (eg, atrial fibrillation, insulin-dependent diabetes mellitus, cerebral infarcts, etc) were excluded. All subjects were required to be either cognitively normal or mildly impaired but nondemented, ie, to have a Clinical Dementia Rating (CDR) 22 of either CDR 0 or CDR 0.5.

At baseline, the study procedures included a medical evaluation (consisting of a physical examination and medical history, electrocardiogram, and standard laboratory tests), a semistructured interview, neuropsychological testing, a magnetic resonance imaging scan, a single-photon emission computed tomography scan, and blood withdrawn for genetic analysis. Only the semistructured interview was repeated annually; the remaining study procedures were repeated in subsets of participants. All subjects provided informed consent prior to the initiation of the study in accordance with the guidelines of Massachusetts General Hospital, Boston.

ASSESSMENT OF CLINICAL SEVERITY

The degree of clinical severity of the subjects was evaluated by the annual semistructured interview. This interview generates both an overall CDR rating and a measure known as the CDR Sum of Boxes (CDR-SB). 23 The interview is based on the initial subject protocol that was used in the development of the CDR scale. 22 It includes a set of questions regarding cognitive and functional status asked of the subject and an informant (eg, family member, friend) and a standardized neurologic, psychiatric, and mental status evaluation of the subject. The mental status evaluation included the Blessed Memory and Orientation Test, 24 which assesses episodic memory, working memory, and orientation; a set of similarities and differences, which assessed executive function; calculations that assessed arithmetic skill and general knowledge; and a standardized language evaluation, including naming, repetition, and comprehension. To be sensitive to clinical impairments at the mildest end of the spectrum, a special set of questions were added to the interview, and the reliability and validity of the revised interview was examined. 23 The mean interrater reliability of the overall CDR rating was high (r=0.99, P<.001) as was the interrater reliability of the 6 CDR subcategories (r=.90) that are used to generate the overall CDR rating. The CDR-SB represents the sum of the ratings in each of the 6 CDR subcategories.

In the current study, each interview was administered annually by a masters’ or doctoral-level clinician (eg, psychiatrist, neuropsychologist, or physician’s assistant) and took approximately 1 to 2 hours to complete. The interviews were performed without knowledge of the other study procedures, including the neuropsychological testing. A consensus review of each case was conducted annually by 2 or more members of the research group (which included the interviewers mentioned here).

GROUP CHARACTERISTICS

AT BASELINE AND AT FOLLOW-UP

Baseline

Based on their initial CDR interview, subjects were divided into 2 groups (Table 1). The normal group consisted of 107 subjects with normal cognition (CDR 0); their mean (SD) CDR-SB score was 0.014 (0.08). (The nonzero mean CDR-SB reflects a handful of individuals with slight difficulties in a domain other than memory; a score of 0.5 in the memory domain is required for an overall rating of 0.5.) The MCI group consisted of 235 nondemented individuals with a CDR rating of 0.5; their mean (SD) CDR-SB was 1.34 (0.77). The 2 groups were similar in age at study entry, although the normal group was somewhat younger (71.4 vs 72.9 years in the MCI group; P=.009). Other demographic variables were not significantly different across the 2 groups, as follows: the educational level (means of 15.6 years vs 15.4 years), the sex distribution (59.8% vs 56.6% female), and the racial distribution (4.7% vs 8.1% nonwhite). The mean Mini-Mental State Exam scores at baseline were similar (29.5 vs 29.1; range, 27-30 vs 24-30, respectively) but were nevertheless significantly different from one another (P<.04). Thus, both groups were largely white and well educated and had high scores on the Mini-Mental State Exam.
The distribution of CDR-SB scores among the mildly impaired subjects was broad (Table 1). At the mild end of the spectrum (ie, CDR-SB 0.5-1.5), many subjects would not meet psychiatric cut-offs commonly used to select MCI subjects in epidemiological studies and clinical trials.22-26 The subjects at the more impaired end of the spectrum (ie, CDR-SB≥2) are comparable with MCI subjects recruited from these settings, based on likelihood of progression to a diagnosis of AD.25 We use MCI here to refer to the entire group of mildly impaired subjects. A retrospective review of the cases indicated that approximately two thirds would fall into the category of amnestic MCI while approximately one third would be considered approximately two thirds would fall into the category of amnestic MCI while approximately one third would be considered amnestic MCI cases, based on the revised criteria for MCI.29

Current Status of Normal Group

Of the 107 individuals who were normal at baseline, a total of 54 were categorized as impaired over time. Of these 54 individuals, 46 were diagnosed as MCI at the time data were analyzed for the present article. Approximately 74% of these had a CDR-SB of 0.5 to 1.5, and approximately 26% (n=12) had a CDR-SB of 2 or higher. Eight of the normal subjects who progressed now carry a diagnosis of dementia (7 of the 8 have a clinical diagnosis of AD); the mean time to progress from normal to AD was 8 years (a rate of approximately 1% per year). The mean (SD) time to progress from normal to MCI was 4 (2.9) years.

Current Status of MCI Group

Of the 235 individuals who were mildly impaired at baseline, 87 individuals were no longer categorized as mildly impaired at the time data were analyzed for the present article. Of these, 69 received a clinical diagnosis of AD, 11 received a clinical diagnosis of a non-AD dementia, and 7 had a high CDR-SB (ie, 4), but the clinical diagnosis was incomplete because of pending medical procedures. Approximately 60% of the remaining subjects (n=84) had a CDR-SB of 2 or higher, and approximately 40% (n=59) had a score of 0.5 to 1.5. Five of the subjects originally categorized as mildly impaired are now categorized as normal (CDR 0). The mean (SD) time to progress from MCI to AD was 4.1 (3.2) years.

Diagnosis of Dementia on Follow-up

As part of the annual review of each case, the consensus diagnostic process determined whether the individual had sufficient impairment for a diagnosis of dementia and, if so, whether the dementia was consistent with research criteria for AD30 or another known diagnostic entity (eg, frontotemporal dementia, vascular dementia).31,32 Diagnoses were based on findings from a combination of clinical history, medical records, laboratory evaluation, and neuroimaging studies.

Autopsy Findings

Over the course of the longitudinal study, 15 subjects have come to autopsy of whom 9 had a clinical diagnosis of probable AD; this diagnosis was confirmed in 6 of the cases based on the criteria for AD from the National Institute on Aging-Reagan Institute. An additional 3 cases showed AD changes insufficient for an autopsy diagnosis of definite AD. These 3 individuals were not considered to have AD in the analyses that follows.

Neuropsychological Measures

All subjects were administered a neuropsychological battery that consisted of 22 tests yielding 23 test scores. The composition of the entire battery has been previously described.16 In the present analyses, we examined 6 of 23 neuropsychological tests measures. These 6 tests included (1) 2 tests of episodic memory; the total recall score of the first 5 trials of the California Verbal Learning Test (CVLT)27 and the free recall score that follows the first 4 trials of the Selective Reminding Test (SRT)31; (2) 3 executive function tests: the time to complete part B of the Trail Making test (Trails B)35 (part A of the Trail Making test was proportional hazards models and facilitate comparisons of effect sizes across tests. Prior to standardization, the raw time to complete Trails B was log-transformed to reduce skewness. Table 2 outlines the distribution of raw test scores and neuropsychological factors for the 2 groups. Mean scores were consistently higher in the normal than in the MCI group. These differences

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Normal at Baseline (n = 107)</th>
<th>MCI at Baseline (n = 235)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, No. (%)</td>
<td>64 (59.8)</td>
<td>133 (56.6)</td>
</tr>
<tr>
<td>Nonwhite, No. (%)</td>
<td>5 (4.7)</td>
<td>8 (8.1)</td>
</tr>
<tr>
<td>Age, y, mean (SD)</td>
<td>71.38 (4.56)</td>
<td>72.90 (5.83)</td>
</tr>
<tr>
<td>&lt;65</td>
<td>4 (3.7)</td>
<td>13 (5.5)</td>
</tr>
<tr>
<td>65-69</td>
<td>34 (31.8)</td>
<td>54 (23.0)</td>
</tr>
<tr>
<td>70-74</td>
<td>50 (46.7)</td>
<td>78 (33.2)</td>
</tr>
<tr>
<td>75-79</td>
<td>13 (12.2)</td>
<td>59 (25.1)</td>
</tr>
<tr>
<td>≥80</td>
<td>6 (5.6)</td>
<td>31 (13.2)</td>
</tr>
<tr>
<td>Education, mean (SD), y</td>
<td>15.64 (2.86)</td>
<td>15.41 (2.91)</td>
</tr>
<tr>
<td>&lt;12</td>
<td>2 (1.9)</td>
<td>6 (2.6)</td>
</tr>
<tr>
<td>12</td>
<td>24 (22.4)</td>
<td>51 (21.7)</td>
</tr>
<tr>
<td>13-15</td>
<td>21 (19.6)</td>
<td>44 (18.7)</td>
</tr>
<tr>
<td>16</td>
<td>15 (14.0)</td>
<td>55 (23.4)</td>
</tr>
<tr>
<td>&gt;16</td>
<td>45 (42.1)</td>
<td>79 (33.6)</td>
</tr>
<tr>
<td>Follow-up time, mean (SD), y</td>
<td>4.70 (3.4)</td>
<td>5.60 (3.8)</td>
</tr>
<tr>
<td>CDR-SB at baseline, mean (SD)</td>
<td>0.01 (0.1)</td>
<td>1.34 (0.8)</td>
</tr>
<tr>
<td>0</td>
<td>104 (97.2)</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>3 (2.8)</td>
<td>67 (28.5)</td>
</tr>
<tr>
<td>1-1.5</td>
<td>NA</td>
<td>104 (44.3)</td>
</tr>
<tr>
<td>2-2.5</td>
<td>NA</td>
<td>51 (21.7)</td>
</tr>
<tr>
<td>3-3.5</td>
<td>NA</td>
<td>13 (5.5)</td>
</tr>
<tr>
<td>MMSE, mean (SD)</td>
<td>29.19 (1.26)</td>
<td>28.45 (0.75)</td>
</tr>
<tr>
<td>APOE carrier status and genotype, No. (%)*</td>
<td>26 (24.5)</td>
<td>71 (30.7)</td>
</tr>
<tr>
<td>APOE ε4 carrier</td>
<td>16 (15.1)</td>
<td>36 (15.6)</td>
</tr>
<tr>
<td>APOE ε2 carrier</td>
<td>13 (12.3)</td>
<td>33 (14.3)</td>
</tr>
<tr>
<td>APOE ε2/3</td>
<td>3 (2.8)</td>
<td>3 (1.3)</td>
</tr>
<tr>
<td>APOE ε3/3</td>
<td>67 (63.2)</td>
<td>127 (55.0)</td>
</tr>
<tr>
<td>APOE ε3/4</td>
<td>21 (19.8)</td>
<td>64 (27.7)</td>
</tr>
<tr>
<td>APOE ε4/ε4</td>
<td>2 (1.9)</td>
<td>4 (1.7)</td>
</tr>
</tbody>
</table>

Abbreviations: APOE, apolipoprotein E; CDR-SB, Clinical Dementia Rating Sum of Boxes; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; NA, not applicable.

*Because of missing APOE data, percentages are based on n = 106 for normal subjects and n = 231 for MCI subjects.
were statistically significant for 5 of 7 test scores (all but letter fluency and the Alpha Span score).

**Genetic Assessment**

The APOE gene was also examined in the participants because the ε4 allele of this gene is overrepresented in patients with AD compared with the general population and is now widely recognized as a risk factor for AD. In addition, APOE ε2 has been associated with decreased risk of AD. APOE genotypes were determined according to previously described methods. APOE genotypes are available on all but 5 subjects in the study sample (98.5%); missing genetic information reduced the effective sample sizes to 106 for those normal at baseline and 231 for those with MCI at baseline (Table 1). Allele frequencies for the normal group (APOE ε2 = 0.08, APOE ε3 = 0.79, and APOE ε4 = 0.13) were similar to the frequencies for the MCI group (APOE ε2 = 0.08, APOE ε3 = 0.76, and APOE ε4 = 0.16). The APOE ε4 carrier frequency was somewhat lower in the normal group (24.5% vs 30.7%), but this difference was not statistically significant. There was no difference in the percentage of APOE ε2 carriers (15.6%) vs 15.1%.

**Statistical Methods**

The time-to-progression data were analyzed using a Cox proportional hazards model as implemented in the PHREG procedure in SAS version 8 (SAS Institute, Cary, NC). These models tested whether specific predictors (ie, neuropsychological test scores and factor scores and, secondarily, demographic factors and APOE genotype) affect the time to develop a specific outcome (ie, increased clinical severity). The hazard ratio indicates the change in risk per 1-unit change in the predictor. For instance, if the hazard ratio for age (measured in years) is 1.06, each year older increases risk by 6%, or if the hazard ratio for the California Verbal Learning Test is 0.55, each 1-SD increase in this test score decreases risk by 45%.

The primary focus of the analyses was time from study entry to the end point of interest. For the subjects normal at baseline, the end point was progression to mild impairment (ie, CDR 0.5). For subjects categorized as MCI at baseline, the end point was progression to a diagnosis of AD. Subjects were considered censored at the time that they died, were lost to follow-up, or developed a dementia other than AD. Subjects were not observed between annual visits, so models with interval censoring were considered, but preliminary analyses revealed that they made little difference, consistent with the very gradual onset of the disease.

Two sets of Cox models were completed. The first set was for the normal subjects who progressed to mild impairment, and the second set was for the subjects with mild impairment who progressed to a clinical diagnosis of AD. A total of 3 models were completed for each of the 2 groups: 2 univariate models and 1 multivariable model. The 2 univariate models first included a univariate model for each of the 6 standardized neuropsychological test scores, as well as APOE ε2 and APOE ε4 carrier status, unadjusted or “crude.” A second univariate model for each of the same 6 neuropsychological test scores, as well as APOE ε2 and APOE ε4 carrier status, was adjusted for age, sex, and educational attainment. The multivariable model was designed to be the “best” multivariable Cox model (given the set of variables). To accomplish this, first, age, sex, and education were “forced” into the model (ie, these variables were retained even if they were not significant); then each of the 6 neuropsychological test scores and APOE ε2 and APOE ε4 carrier status were entered in the model but were only retained if they were significant at the .05 level. Both age at entry and educational level were controlled for linearly because there was no evidence of a nonmonotonic relationship between these variables and the log hazard of progression or conversion.

As noted earlier, the Cox models included 2 ways of examining the APOE effect on risk of cognitive decline, one pertaining to APOE ε4 and one pertaining to APOE ε2. APOE ε2 was coded as to whether or not the subject was an APOE ε2 carrier. We did not examine APOE ε2 “dose” because there were no ε2ε2 subjects. APOE ε4 was coded as to whether or not the subject was an APOE ε4 carrier and also as the linear effect of dose of APOE ε4 allele (ie, 0 vs 1 vs 2 alleles).

The proportional hazards assumption was evaluated descriptively by checking whether the negative log of survival probabilities associated with higher level of each covariate were constant multiples of those of the lower levels across the entire range of event time. This approach was extended to the multivariate Cox models by examining such patterns across the levels of each linear predictor. Model fit was also examined descriptively by checking the distribution of the martingale residuals as well as deviance residuals. Figure 1 and Figure 2, depicting predicted survival curves as a function of variation in the neuropsychological tests and factors (shown at the mean and 1 SD above and below the mean), are based on the adjusted univariate models. They include a single neuropsychological test and the following covariates for adjustment purpose: age (fixed at the mean), sex (for women), and education (fixed at the mean).

**RESULTS**

**RELATIONSHIP OF DEMOGRAPHIC VARIABLES AND CDR-SB TO TIME TO COGNITIVE DECLINE**

In univariate Cox models predicting progression in subjects normal at baseline, no demographic variable was statistically significant either alone (crude hazard ratios or adjusted for the remaining demographic factors. Hazard ratios reported in the text are from the adjusted models (Table 3).
In univariate Cox models predicting progression to a diagnosis of AD in subjects categorized with mild impairment at baseline, age was significant in both crude and adjusted models (HR = 1.06 [1.02-1.11], P = .003). The remaining demographic variables were not significant (Table 3). However, they were included in the adjusted models and forced into the multivariable model to minimize the risk of residual confounding.

RELATIONSHIP OF NEUROPSYCHOLOGICAL VARIABLES TO TIME TO COGNITIVE DECLINE

In the analyses of time to progression among subjects normal at baseline, virtually all of the neuropsychological tests except the Alpha Span were statistically significant in both the crude and adjusted univariate analyses (Table 4). The CVLT total score showed the largest effect per standard deviation across the full sample (HR = 0.55 [0.39-0.79], P = .001). The 2 most significant predictor variables (the CVLT and Trails B) are shown in Figure 1.

In the analyses of time to diagnosis of AD among subjects with mild impairment at baseline, all tests but the Alpha Span and Self Ordering test were significant in univariate analyses (Table 4). There were large effects for the CVLT (HR = 0.53 [0.41-0.68], P < .001), the SRT (HR = 0.48 [0.38-0.60], P < .001), and the time to complete Trails B (HR = 1.76 [1.42-2.18], P < .001); these 3 tests are shown in Figure 2.

EFFECT OF APOE ON TIME TO COGNITIVE DECLINE

Table 5 outlines the results of the univariate analyses of the APOE variables on time to cognitive decline. In the analysis of time to progression to mild impairment among subjects normal at baseline, APOE ε2 was protective (HR = 0.13 [0.03-0.56], P = .006) in the adjusted analyses. There was no APOE ε4 effect.

In the analysis of time to a diagnosis of AD, when modeled as binary APOE ε4 carrier status, there was no significant increase in risk in the adjusted model (HR = 1.50 [0.91-2.43], P = .11), but when modeled as APOE ε4 dose, there was a statistically significant effect of each APOE ε4 allele (HR = 1.64 [1.04-2.57], P = .03).

TIME TO COGNITIVE DECLINE USING MULTIVARIABLE MODELS

In the multivariable model of predictors of time to progression to mild impairment among subjects normal at baseline, only CVLT total score (HR = 0.58 [0.41-0.83], P = .003) and APOE ε2 carrier status (HR = 0.14 [0.03-0.57], P = .006) remained in the model (Table 6). In the multivariable model of progression to a diagnosis of AD among those with mild impairments at baseline, SRT free recall score (HR = 0.63 [0.46-0.87], P = .005) and time to complete Trails B (HR = 1.38 [1.07-1.78], P = .01) remained in the model (Table 6).

Thus, APOE ε4 status did not exhibit an independent effect in the multivariable model, even though it was a predictor of time to diagnosis of AD in the univariate adjusted models. To understand this phenomenon better, we performed linear regressions that included APOE ε4 dose and the neuropsychological tests that were significant in the multivariable model. These regressions exhibited an APOE ε4 dose (modeled as a linear or quadratic term) in the expected direction; that is, there were statistically significant, larger effects for 2 vs 1 ε4 allele for the SRT total score (P = .02) (all P values are from the model with the quadratic term for ε4 dose). We used a quadratic term for APOE ε4 dose, which gives extra weight to those with 2 copies of the ε4 allele (i.e., 0 for no ε4 alleles, 1 for 1 copy, and 2 for 2 copies) because a much larger ε4 effect has previously observed in homozygotes.45

Figure 1. Survival curves for prediction of time to progression to mild impairment among normal controls (based on the adjusted univariate model [shown at the mean and 1 SD above and below the mean]) as a function of variation in (A) the California Verbal Learning Test (total score from trials 1 through 5) and (B) time to complete Trails B (in seconds).
EVALUATION OF STATISTICAL MODELS

We assessed the Cox models to determine whether they met assumptions for proportional hazards. No covariates violated the proportional hazards assumption. In addition, the overall fit was good, and none of the models showed extreme residuals.

COMMENT

PREDICTIVE VALUE
OF NEUROPSYCHOLOGICAL PERFORMANCE

These findings suggest that, among normal individuals, poorer cognitive performance on specific neuropsychological tests predicts the time to develop mild degrees of impairment. In the current analyses, the risk of progressing from normal to MCI was considerably greater among those with lower scores on tests of episodic memory, ie, the CVLT and the SRT. The importance of episodic memory performance as a predictor of time to progression was reflected in the fact that variables representing this cognitive domain were highly significant in both the univariate and multivariable models.

These results also indicated that performance on an executive function test, the Trail Making test, predicts time to progress from normal to MCI. However, statistical significance was only seen in the univariate, but not the multivariable models, making it less convincing that impairment in this type of task is a marker of likelihood of progression from normal to MCI. It may be possible that more sensitive tests of executive function would serve as better markers of incipient impairment.

The neuropsychological variables that predict time to progression from mild impairment to a clinical diagnosis of AD are quite comparable with those that predict progression from normal to MCI. Once again, it was evident that lower performance on tests of episodic memory confers considerably increased risk of cognitive decline as shown by all the statistical models evaluated. In these analyses, poorer performance on an executive function test, the Trail Making test, was also shown to be a predictor of time to progression from mild impairment to a clinical diagnosis of AD (as shown by both the univariate and multivariable model). Thus, the primary difference between the neuropsychological predictors for time to progression among normal subjects and time to a diagnosis of AD among MCI cases was the strength of the finding regarding the Trail Making test. These analyses were repeated using factor scores that represented the domains of episodic memory, executive function, spatial skill, and general knowledge. The findings with respect to the episodic memory factor and the executive function factor confirm the results of the individual test scores presented here (data not shown).

We hypothesize that the episodic memory tasks that were particularly useful for prediction of course in the present study (ie, the CVLT and the SRT) were predictive because not only are they difficult verbal memory tests, but they have the added feature that they can be performed more effectively if the subject takes advan-

Figure 2. Survival curves for prediction of time to progression to a diagnosis of Alzheimer disease among subjects with mild cognitive impairment (MCI) (based on the adjusted univariate model [shown at the mean and 1 SD above and below the mean]) as a function of variation in (A) the total score on the California Verbal Learning Test (total score from trials 1 through 5), the Selective Reminding Test (free recall) (B), and time to complete Trails B (C) (in seconds).
tage of organizational cues inherent in the task itself, thus using both episodic memory and executive function skills. The executive function test that was particularly useful for prediction, Trails B, requires the individual to switch back and forth from one well-learned series to another (ie, numbers and letters). This requires the individual to inhibit the tendency to complete the series in order, thus making switching between these 2 mental sets particularly challenging.

Taken together, these findings are consistent with the recent revision of the MCI criteria\(^29\) and a recent meta-analysis\(^18\) indicating that individuals with MCI can have impairments in cognitive domains above and beyond that of memory. Although we did not conduct analyses that specifically address the order in which cognitive impairments emerge during the evolution of AD, these results suggest that during prodromal AD, most individuals first develop difficulty with episodic memory fol-

---

**Table 3. Analysis of Demographic Factors and Risk of Cognitive Decline**

<table>
<thead>
<tr>
<th>Factor†</th>
<th>Risk of Progression in Subjects Normal at Baseline</th>
<th>Risk of Progression to AD in Subjects With MCI at Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude (95% CI)</td>
<td>Adjusted* (95% CI)</td>
</tr>
<tr>
<td>Age</td>
<td>1.03 (0.97-1.08)</td>
<td>.38</td>
</tr>
<tr>
<td>Sex</td>
<td>0.66 (0.39-1.14)</td>
<td>.14</td>
</tr>
<tr>
<td>Education</td>
<td>1.04 (0.94-1.15)</td>
<td>.47</td>
</tr>
</tbody>
</table>

Abbreviations: AD, Alzheimer disease; CI, confidence interval; MCI, mild cognitive impairment. *Adjusted for the remaining demographic factors. †Hazard ratios given are per year.

**Table 4. Univariate Analyses of Neuropsychological Tests and Risk of Cognitive Decline**

<table>
<thead>
<tr>
<th>Neuropsychological Test*</th>
<th>Risk of Progression in Subjects Normal at Baseline (n = 107)</th>
<th>Risk of Progression to AD in Subjects With MCI at Baseline (n = 235)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude (95% CI)</td>
<td>Adjusted† (95% CI)</td>
</tr>
<tr>
<td>CVLT total score from trials 1-5</td>
<td>0.60 (0.44-0.81)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Selective Reminding Test score</td>
<td>0.66 (0.48-0.91)</td>
<td>.01</td>
</tr>
<tr>
<td>Time to complete Trails B</td>
<td>1.44 (1.04-2.01)</td>
<td>.03</td>
</tr>
<tr>
<td>Letter fluency test (total of F, A, and S)</td>
<td>0.77 (0.58-1.02)</td>
<td>.07</td>
</tr>
<tr>
<td>Alpha Span test total score</td>
<td>0.85 (0.64-1.14)</td>
<td>.28</td>
</tr>
<tr>
<td>Self Ordering test total score</td>
<td>1.06 (1.00-1.13)</td>
<td>.06</td>
</tr>
</tbody>
</table>

Abbreviations: AD, Alzheimer disease; CI, confidence interval; CVLT, California Verbal Learning Test; MCI, mild cognitive impairment; Trails B, Trail Making test part B. *Hazard ratios given are per standard deviation across the entire sample (n = 342). †Adjusted for age, sex, and education.

**Table 5. Analysis of APOE and Risk of Cognitive Decline**

<table>
<thead>
<tr>
<th>APOE Status*</th>
<th>Risk of Progression in Subjects Normal at Baseline</th>
<th>Risk of Progression to AD in Subjects With MCI at Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude (95% CI)</td>
<td>Adjusted† (95% CI)</td>
</tr>
<tr>
<td>APOE ε2 carrier</td>
<td>0.17 (0.04-0.72)</td>
<td>.02</td>
</tr>
<tr>
<td>APOE ε4 carrier</td>
<td>1.01 (0.52-1.96)</td>
<td>.98</td>
</tr>
<tr>
<td>APOE ε4 dose</td>
<td>1.05 (0.57-1.93)</td>
<td>.87</td>
</tr>
</tbody>
</table>

Abbreviations: APOE, apolipoprotein E; AD, Alzheimer disease; CI, confidence interval; MCI, mild cognitive impairment. *Hazard ratios are for presence vs absence of the given allele or per unit of gene dose as linear effect. †Adjusted for age, sex, and education.
Our findings agree with 2 previous reports concerning normal subjects who received detailed neuropathological examinations.46-47 Some of the normal individuals were found to have AD pathology on autopsy, and the investigators reported that episodic memory performance differed between subjects with and without pathological findings. Our results differ, however, from another study that also performed autopsies on a small number of normal subjects who had been carefully examined during life. These investigators31 reported that there was no decline in cognitive performance for normal individuals with pathology (n = 3) on autopsy as compared with those lacking pathological findings (n = 5), based on either individual test scores or a factor score derived from the test battery as a whole. The explanation for this discrepancy may relate to the impact of the small sample size (making it difficult to identify differences between groups) as well as the individual tests examined and the fact that neuropathological testing was administered annually. Although the investigators administered 2 tests of episodic memory, only one (the paired associate test from the Wechsler Memory Scale) examined delayed recall; moreover, a substantial proportion of the paired associates were considered “easy.” Thus, the memory tests may have been insufficiently sensitive to subtle early changes in cognition. In addition, it is likely that annual administration of the tests produced a practice effect that made it difficult to observe subtle changes in performance.32

ROLE OF APOE IN PREDICTION OF TIME TO COGNITIVE DECLINE

In the present analyses, the role of APOE status in prediction of cognitive decline varied depending on whether an individual was normal or had already developed mild impairments in daily life. Normal individuals who carried one or more copy of the APOE e2 allele were less likely to develop cognitive impairments over time than individuals with no copies of the e2 allele. This protective effect of APOE e2 has been previously reported.33 The fact that APOE e2 status was significant in both the multivariate and univariate models indicates that its impact on risk for cognitive decline is not fully accounted for by its influence on memory performance and suggests the possibility that the benefit of e2 carrier status might be greatest at the very earliest stages of disease.
Among individuals with mild impairments, the APOE ε4 allele increased risk for AD in a dose-dependent manner. The linear regression models showed that presence of an APOE ε4 allele was associated with lower memory performance, as our group and others have previously reported. APOE ε4 carrier status was not, however, significant in the multivariable models, suggesting that its influence on cognition is largely accounted for by memory test scores.

STUDY LIMITATIONS

The study must be interpreted in light of its limitations. The subjects are well educated and primarily white, so the results may not generalize to the US population at large. We must also acknowledge that differences in the sample size between the controls and the MCI subjects may have had an impact on the findings. We compared the hazard ratios and confidence intervals for the tests that were significant predictors for the MCI cases who progressed but not for the normal controls and did not see patterns suggesting that limitations in power were responsible for the results. Nevertheless, there is the possibility that larger sample sizes might show additional relationships not found here. In addition, the results may depend on the specific neuropsychological tests involved. While many of the individual tests were significant in the univariate models, only selected ones remained significant in the multivariable model, suggesting that the specific tests identified here may be particularly useful in capturing cognitive impairments predictive of progression.

CONCLUSIONS

These findings indicate that neuropsychological testing may be useful in identifying normal individuals who are likely to progress to mild impairment. One potential application of these findings relates to the design of primary intervention trials seeking to delay the onset of memory problems among individuals who are normal. It might be possible to use the type of episodic memory tests shown to be sensitive in the current analyses to select individuals at high risk for progression, thus increasing the power of such studies to show an effect of the intervention.

Accepted for Publication: November 3, 2006.

Author Affiliations: Departments of Psychiatry (Dr Blacker), Medicine (Dr Lee and Ms Muzikansky), and Neurology (Dr Tanzi), Massachusetts General Hospital, Harvard Medical School, Boston; Department of Biostatistics, Harvard School of Public Health, Boston, Mass (Dr Martin); Department of Psychology, University of Southern California, Los Angeles (Dr McArthur); Department of Anatomy and Neurobiology, Boston University, Boston, Mass (Dr Moss); Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Md (Dr Albert). Dr Martin is currently with Wyeth Research, Cambridge, Mass.

Correspondence: Marilyn Albert, PhD, 1620 McElderry St, Reed Hall East 2, Baltimore, MD 21205 (malbert9@jhmi.edu)

Author Contributions: Study concept and design: Blacker, Tanzi, Moss, and Albert. Acquisition of data: Blacker, Tanzi, and Albert. Analysis and interpretation of data: Blacker, Lee, Muzikansky, Martin, McArthur, and Albert. Drafting of the manuscript: Blacker, Martin, and Albert. Critical revision of the manuscript for important intellectual content: Blacker, Lee, Muzikansky, Tanzi, McArthur, Moss, and Albert. Statistical analysis: Blacker, Lee, Muzikansky, Martin, and McArthur. Obtained funding: Albert. Administrative, technical, and material support: Blacker, Tanzi, Moss, and Albert. Study supervision: Blacker and Albert.

Financial Disclosure: None reported.

Funding/Support: The study was supported by grants P01 AG04953 and P50 AG05146 from the National Institute on Aging.

Acknowledgment: We thank Drs Li-Jen Wei, Lu Tian, Larry Leon, and Rebecca Betensky for contributions to the biostatistical methods; Jeanette Gunther and numerous research assistants for careful attention to our subjects and their data; Dr Mary Hyde for invaluable assistance with data management; and our participants for their generous ongoing involvement in the study.

REFERENCES


52. Lamar M, Resnick S, Zonderman A. Longitudinal changes in verbal memory in older adults: distinguishing the effects of age from repeat testing. *Neurology*. 2003;60:82-86.