The Fate of the 0.5s: Predictors of 2-Year Outcome in Mild Cognitive Impairment

Eleni Aretouli, Ozioma C. Okonkwo, Jaclyn Samek, AND Jason Brandt

1Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University School of Medicine, Baltimore, Maryland
2Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, Maryland
3The Copper Ridge Institute, Sykesville, Maryland

(RECEIVED July 15, 2010; FINAL REVISION November 23, 2010; ACCEPTED November 23, 2010)

Abstract
Impairments in executive cognition (EC) may be predictive of incident dementia in patients with mild cognitive impairment (MCI). The present study examined whether specific EC tests could predict which MCI individuals progress from a Clinical Dementia Rating (CDR) score of 0.5 to a score ≥1 over a 2-year period. Eighteen clinical and experimental EC measures were administered at baseline to 104 MCI patients (amnestic and non-amnestic, single- and multiple-domain) recruited from clinical and research settings. Demographic characteristics, screening cognitive measures and measures of everyday functioning at baseline were also considered as potential predictors. Over the 2-year period, 18% of the MCI individuals progressed to CDR ≥1, 73.1% remained stable (CDR = 0.5), and 4.5% reverted to normal (CDR = 0). Multiple-domain MCI participants had higher rates of progression to dementia than single-domain, but amnestic and non-amnestic MCIs had similar rates of conversion. Only three EC measures were predictive of subsequent cognitive and functional decline at the univariate level, but they failed to independently predict progression to dementia after adjusting for demographic, other cognitive characteristics, and measures of everyday functioning. Decline over 2 years was best predicted by informant ratings of subtle functional impairments and lower baseline scores on memory, category fluency, and constructional praxis. (JINS, 2011, 17, 277–288)

Keywords: Mild cognitive impairment, Dementia, Predictors of decline, Executive cognition, Clinical Dementia Rating scale, MCI outcome

INTRODUCTION
Older adults with mild cognitive impairment (MCI) are at substantially higher risk of dementia than cognitively normal elderly (5–10% for MCI vs. 1–2% for normal elderly per year) (Mitchell & Shiri-Feshki, 2009; Petersen et al., 1999). Distinct clinical subtypes of MCI have been identified, the most extensively studied of which is amnestic MCI (aMCI). These are persons with memory complaints and psychometric evidence of memory decline, but intact overall cognitive function and generally preserved activities of daily living (Winblad et al., 2004). The latter is typically defined by a score no higher than 0.5 on the Clinical Dementia Rating (CDR) (Hughes, Berg, Danziger, Coben, & Martin, 1982). It has been suggested that aMCI has a less favorable prognosis than do the non-amnestic subtypes (naMCI) (Fischer et al., 2007; Morris & Cummings, 2005; Ravaglia et al., 2008). However, this is not a consistent finding (Mitchell, Arnold, Dawson, Nestor, & Hodges, 2009; Nordlund et al., 2010; Rozzini et al., 2007). In fact, many have reported that it is the presence of deficits in multiple cognitive domains [multiple-domain MCI (mdMCI)] rather than any specific domain that confers increased risk for dementia (Alexopoulos, Grimmer, Perneczky, Domes, & Kurz, 2006b; Baars et al., 2009; Manly et al., 2008; Rasquin, Lodder, Visser, Lousberg, & Verhey, 2005).

Although research has focused on the progression from MCI to dementia, reversion from MCI to normal cognition has also been documented, with rates ranging from 4% to 53% (Fisk, Merry, & Rockwood, 2003; Kryscio, Schmitt, Salazar, Mendiondo, & Marksbery, 2006; Larrieu et al., 2002; Ravaglia et al., 2006). Finally, a considerable number of individuals with MCI (11–21%) remain stable when followed for up to 10 years (Fisk & Rockwood, 2005; Ganguli, Dodge, Shen, & DeKosky, 2004; Mitchell & Shiri-Feshki, 2009; Visser, Kester, Jolles, & Verhey, 2006). Thus, the outcome of MCI is varied.

Distinguishing the MCI individuals who are on a trajectory toward dementia from those with temporary or non-progressive
cognitive difficulties is of increasing importance. There is some evidence that pharmacological intervention in MCI could delay the onset of dementia, or slow its progression (Aisen, 2008; Doody et al., 2009; Petersen et al., 2005; Yesavage et al., 2007). An accurate prognosis can help patients and their family plan for the care of the patient and make long-term financial decisions. Finally, identifying the “preclinical” stages of the disease could enhance our understanding of its early evolution, enabling us to study the pathological changes that occur before unambiguous symptoms emerge.

There are multiple approaches to the preclinical detection of a neurodegenerative disease, including use of brain imaging (Jack et al., 2008), blood and cerebrospinal fluid biomarkers (Mattsson et al., 2009), and genetic risk assessment (Caselli et al., 2009). The present study considers neurocognitive characteristics, since they are so easily attainable, widely available, and usually part of the routine clinical assessment of people at risk for dementia.

There are several reasons to believe that deficits in executive cognition (EC) are particularly predictive of the development of dementia. First, they are relatively common among individuals with MCI (Chen et al., 2000; Daly et al., 2000; Griffith et al., 2003) and present at the early stages of many diseases that cause dementia, such as Alzheimer’s disease, Parkinson’s disease, or Huntington’s disease (Jacobs, Marder, et al., 1995; Mahieux et al., 1998; Woods & Troster, 2003). Second, many of the cognitive tests that others have reported to have prognostic value for dementia have substantial executive control requirements (Bondi et al., 1994; Elias et al., 2000; Jacobs, Sano, et al., 1995; Rapp & Reischies, 2005). Finally, several longitudinal studies suggest that EC measures are associated with subsequent decline (Albert, Moss, Tanzi, & Jones, 2001; Chen et al., 2000; Crowell, Luis, Vanderploeg, Schinka, & Mullan, 2002; Daly et al., 2000).

Executive cognition encompasses a wide range of abilities and processes (Miyake et al., 2000). An important limitation of previous studies investigating the ability of EC to predict incident dementia is that the sampling of EC was limited, typically to two or three tests (Albert et al., 2001; Amieva et al., 2004; Chen et al., 2000; Crowell et al., 2002; Daly et al., 2000; Tierney et al., 1996). Although including multiple EC measures as potential predictor variables may be statistically problematic, it is conceptually necessary when exploring the predictive utility of specific varieties of EC. In fact, patients with MCI have been found to be impaired on only specific aspects of EC (Brandt et al., 2009; Zhang, Han, Verhaeghen, & Nilsson, 2007).

Previous studies had other methodological limitations, as well. The participants’ outcome status was usually defined based on the presence or absence of a clinical diagnosis of dementia. Differences in diagnostic procedures or limiting the outcome to patients receiving the diagnosis of probable Alzheimer’s disease (i.e., excluding those with non-AD dementias) may have contributed to discrepant findings or may have limited the generalizability of previous findings. We circumvented this limitation by assessing cognitive and functional decline sufficient for a dementia diagnosis, but operationalized as a score ≥ 1 on the CDR. Furthermore, some previous studies have not sufficiently considered the role of demographic characteristics, such as age and education. Finally, it is unclear whether executive functioning predicts development of dementia above and beyond other factors, such as abilities to perform daily tasks at baseline or other aspects of cognition.

The purpose of the present study was to determine whether specific aspects of EC contribute uniquely to the detection of individuals who progress to dementia (CDR scores ≥ 1) and whether they do so beyond and above other demographic, other cognitive and clinical characteristics. We examined the outcomes of patients with four subtypes of MCI over a 2-year period and their association with baseline performances on 18 clinical and experimental EC tests. In addition, we examined whether specific MCI subtypes were associated with higher rates of cognitive and functional decline. Multiple-domain MCI patients were predicted to be at higher risk of decline than the other MCI subtypes.

METHOD

Participants

MCI groups

A total of 141 persons with MCI participated in the study. Most (70%) were recruited from the Johns Hopkins Alzheimer’s Disease Research Center (ADRC) and from other research studies. A smaller number of participants (30%) were referred from University clinics and physicians in the community. Ninety percent of the participants were Caucasian, and 10% were African American.

Participants were diagnosed with MCI according to Petersen criteria (Petersen, 2004; Winblad et al., 2004). Inclusion criteria included a Mini-Mental State Exam (MMSE) score in the normal range (i.e., at or above the 20th percentile for age and education; range = 26–30 in our sample) (Bravo & Hebert, 1997) and a global score of 0.5 on the CDR (Hughes et al., 1982). In addition, participants were required to perform at or below 1.5 standard deviations below the mean for age and education (i.e., 6.7th percentile), according to published norms, on one or more of the following screening tests: Logical Memory (story A) of the Wechsler Memory Scale-Revised (WMS-R) (Wechsler, 1987), the 30-item version of the Boston Naming Test (Brandt et al., 1989; Goodglass & Kaplan, 1983), word list generation (for the letters FAS [Phonemic Fluency] and the semantic categories animals and vegetables [Category Fluency]) (Rascovsky, Salmon, Hansen, Thal, & Galasko, 2007), and clock drawing to request (Rouleau, Salmon, Butters, Kennedy, & McGuire, 1992). Finally, each participant was required to have a study partner who could provide information about his/her functional abilities.

Exclusion criteria were any history of major mental illness, central nervous system disorder, or active systemic illness (e.g., cancer). Volunteers with past or present depression were not excluded since depression is very common in MCI and may be related to outcome (Jorm, 2001; Lyketsos et al., 2002).
A total of 141 participants received baseline neuropsychological evaluations, fulfilled our criteria for MCI, and were subsequently assigned to one of four MCI subgroups: amnestic single domain (AS; n = 36), amnestic multiple domain (AM; n = 45), non-amnestic single domain (NAS; n = 26), and non-amnestic multiple domain (NAM; n = 17). The criteria for subgroup membership were based on performance on the four cognitive screening tests described above; see Brandt et al. (2009) for details. Of the 141 participants with a baseline assessment, 104 had a follow-up examination after 2 years. Five MCI participants were deceased by the second evaluation, 15 refused participation, and 9 were not reachable (attrition rate per year = 10.9%). Eight participants had not yet received follow-up evaluations at the time of the analysis and were, therefore, excluded. The final sample included 30 AS, 38 AM, 23 NAS, and 13 NAM patients. Additional details regarding the study participants have been published previously (Brandt et al., 2009).

Group outcome

Participant status 2 years after baseline assessment was assessed with a repeat CDR interview with the study partner. Each participant had the same study partner for both CDR interviews, except 9 participants, whose initial informants were not available for the follow-up interview. For the purposes of this study, subjects are described as having progressed to dementia if they had a CDR global score ≥ 1 at follow-up. If they obtained a CDR global score of 0, they are described as having reverted to normal. If they obtained a 0.5 at follow-up, they are described as still having MCI.

Table 1. Summary of tests of executive function (Brandt et al., 2009)

<table>
<thead>
<tr>
<th>Proposed domain</th>
<th>Test</th>
<th>Measures</th>
<th>Scores range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spontaneous Flexibility and Generativity</td>
<td>Alternate Uses Test (Guilford, Christensen, Merryfield, &amp; Wilson, 1978)</td>
<td>Total raw score</td>
<td>0–36</td>
</tr>
<tr>
<td></td>
<td>Random Number Generation (Brugger, Monsch, Salmon, &amp; Butters, 1996)</td>
<td>Sum of written trial RNG and oral trial RNG</td>
<td>0–2</td>
</tr>
<tr>
<td></td>
<td>Tinker Toy Test (Koss, Patterson, Mack, Smyth, &amp; Whitehouse, 1998; Lezak, 1982)</td>
<td>Total raw score</td>
<td>0–13</td>
</tr>
<tr>
<td>Inhibition of Prepotent Responses</td>
<td>D-KEFS Stroop Test (Delis, Kaplan, &amp; Kramer, 2001)</td>
<td>Inhibition trial scaled score</td>
<td>0–19</td>
</tr>
<tr>
<td></td>
<td>Hayling Test (Burgess &amp; Shallice, 1997)</td>
<td>Total scaled score</td>
<td>0–10</td>
</tr>
<tr>
<td></td>
<td>Completions &amp; Corrections Test (Manning &amp; Brandt, 2006, February)</td>
<td>Total correct score</td>
<td>0–12</td>
</tr>
<tr>
<td>Planning and Sequencing</td>
<td>Porteus Maze Test (Porteus, 1965)</td>
<td>Test Age score</td>
<td>0–17</td>
</tr>
<tr>
<td></td>
<td>D–KEFS Tower Test (Delis et al., 2001)</td>
<td>Total achievement scaled score</td>
<td>0–19</td>
</tr>
<tr>
<td></td>
<td>Tic–Tac–Toe (Brandt et al., 2009; Crowley &amp; Siegler, 1993)</td>
<td>Total correct score</td>
<td>(−16)–16</td>
</tr>
<tr>
<td>Concept/Rule Learning and Set Shifting</td>
<td>D-KEFS Sorting Test (Delis et al., 2001)</td>
<td>Confirmed sorts scaled score</td>
<td>0–19</td>
</tr>
<tr>
<td></td>
<td>Brixton Spatial Anticipation Test (Burgess &amp; Shallice, 1997)</td>
<td>Total scaled score</td>
<td>0–10</td>
</tr>
<tr>
<td></td>
<td>Verbal Concept Attainment Test (Bornstein, 1982)</td>
<td>Total correct raw score</td>
<td>0–24</td>
</tr>
<tr>
<td>Decision-Making and Judgment</td>
<td>Stanford Binet Absurdities Test (Thordike, Hagen, &amp; Sattler, 1986)</td>
<td>Total raw score</td>
<td>0–32</td>
</tr>
<tr>
<td></td>
<td>Iowa Gambling Test (Bechara, Damasio, Tranel, &amp; Anderson, 1998)</td>
<td>Advantageous selections (C+D deck responses) on block 5 minus block 1</td>
<td>_</td>
</tr>
<tr>
<td></td>
<td>Experimental Judgment Test</td>
<td>Mean percent deviation</td>
<td>_</td>
</tr>
<tr>
<td>Working Memory and Resource-Sharing</td>
<td>Trail Making Test (Reitan, 1958)</td>
<td>Time on Part B minus time on Part A</td>
<td>_</td>
</tr>
<tr>
<td></td>
<td>Brief Test of Attention (Schretlen, Bobholz, &amp; Brandt, 1996)</td>
<td>Total correct raw score</td>
<td>0–20</td>
</tr>
<tr>
<td></td>
<td>TEA; Telephone Search While Counting (Robertson, Ward, Ridgeway, &amp; Nimmo-Smith, 1994)</td>
<td>Dual task decrement score</td>
<td>_</td>
</tr>
</tbody>
</table>

Note. D-KEFS = Delis-Kaplan Executive Function System; TEA = Test of Everyday Attention.
tic-tac-toe (Brandt et al., 2009; Crowley & Siegler, 1993). A standardized version of this well-known paper-and-pencil game was developed specifically for the present study. Sixteen trials were played, with the examiner and participant alternating who moves first. On half the trials on which the examiner started, s/he purposely made a sub-optimal initial move (i.e., one other than a corner), thereby allowing the patient an advantage. The examinee is credited on every trial s/he wins, and is debited one point for every loss; ties result in no change in score.

Experimental judgment test (Brandt et al., 2009). This task requires that the subject estimate three attributes of the examiner, specifically his/her age, height, and weight. Score on the test is the mean percent deviation from actual values on the three items.

Measures of functional status

Three informant-reported ratings of daily functioning were included in the present study, since even mild restriction in instrumental activities of daily living (IADL) has been found to be associated with a much higher risk of progression to dementia and a much lower chance of reversibility of cognitive deficit (Peres et al., 2006). The study partners’ responses on questionnaires assessing functional status were not considered in the calculation of the CDR scores.

The Activities of Daily Living-Prevention Instrument developed by the Alzheimer’s Disease Cooperative Study Committee (ADCS ADL-PI) (Ferris et al., 2006; Galasko et al., 2006) consists of 15 items assessing performance on complex activities of daily living rated on a 3-point scale (0 = “no difficulty” to 2 = “a lot of difficulty”). The sum of ratings on the 15 functional items constitutes the score.

The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (Jorm & Jacomb, 1989) assesses changes in an elderly subject’s everyday cognitive abilities as manifested in daily life. Twenty-six cognitive activities of daily life are described and rated on a 5-point scale compared to 10 years previously (1 = “much better” to 5 = “much worse”). The mean of the 26 ratings was used in our analyses.

The Dysexecutive Questionnaire (DEX) (Wilson, Alderman, Burgess, Emslie, & Evans, 1996) measures behavioral changes evident in daily functioning that result from a dysexecutive syndrome. Twenty statements describing common problems of everyday life are rated on a 5-point scale according to their frequency (0 = “never” to 4 = “very often”).

The Johns Hopkins University Institutional Review Board fully reviewed and approved the study protocol. Written informed consent was obtained from all participants and their study partners.

Statistical Analysis

Statistical analyses were performed using PASW 17. One-way analyses of variance (ANOVAs) and chi-square ($\chi^2$) were used to compare the four groups on demographic and clinical characteristics at baseline. The $\chi^2$ was further used to test the association of participants’ group and subgroup membership (aMCI vs. nMCI; single domain MCI [sdMCI] vs. multiple domain MCI [mdMCI]) at baseline with the frequency of development of dementia at the follow-up visit. In addition, we compared the baseline characteristics and performances of those who progressed to a CDR global score of 1 or more with the performances of those who remained at 0.5.

The ability of demographic and clinical characteristics, the cognitive screening tests and the EC measures to predict progression from CDR = 0.5 to CDR \geq 1 was assessed with univariate logistic regression analyses. Stable MCI was the reference group in all analyses. Variables significantly predicting the MCI outcome in the univariate regression models were then entered in a hierarchical logistic regression model together with demographic characteristics. The hierarchical regression model was developed to determine whether EC predicts outcome in MCI patients above and beyond demographic, clinical, and other cognitive factors. Therefore, in the first block, demographic variables (age, education, sex) were forced to enter; in the second block, measures of everyday functioning that were significantly associated with cognitive status at 2-year follow-up could enter; in the third block, the screening battery tests that predicted progression to dementia at the univariate level could enter; finally, in the fourth block, the EF variables that were significant predictors at the univariate level were allowed to enter. Variables in each block were entered in a forward stepwise manner, using the likelihood ratio criterion.

A receiver operating characteristic (ROC) curve was plotted to illustrate the sensitivity and specificity of the predictive model that resulted from the hierarchical regression analysis.

An omnibus forward stepwise logistic regression model, using the likelihood ratio criterion, was computed to test the predictive ability of the variables that were significant predictors of progression at the univariate level when entered simultaneously. The purpose of this omnibus regression analysis was to test the prognostic value of EC in the context of all other clinical and cognitive measures.

Due to sample size constraints, all the regression analyses were performed within the pooled MCI sample, rather than within the subgroups separately. In the present paper, although we describe all three possible outcomes of MCI patients (reversion to normal, stable MCI and progression to dementia), we modeled only progression to dementia due to the small number of the MCI individuals who reverted to normal (i.e., only nine persons scored 0 on the CDR at the 2-year follow-up).

RESULTS

Demographic and Clinical Baseline Characteristics

The demographic and clinical characteristics of the participants at baseline are summarized in Table 2. The groups did not differ in age ($F(3,104) = 2.06; p > .05$) or education ($F(3,104) = 2.03; p > .05$), but had different sex distributions ($\chi^2 = 10.06; df = 3; p = .018$); women predominated in the NAM group. The groups had equivalent MMSE scores...
Table 2. Demographic and baseline clinical characteristics of the participants who had 2-year follow-ups

<table>
<thead>
<tr>
<th></th>
<th>Amnestic MCI</th>
<th>Non-amnestic MCI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single domain</td>
<td>Multiple domain</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>Age, years</td>
<td>74.97 (SD = 5.82)</td>
<td>78.32 (SD = 7.25)</td>
</tr>
<tr>
<td>Education, highest grade completed</td>
<td>16.17 (SD = 2.17)</td>
<td>16.53 (SD = 2.42)</td>
</tr>
<tr>
<td>Sex, male : female</td>
<td>22 : 8</td>
<td>24 : 14</td>
</tr>
<tr>
<td>Mini-Mental State Exam, score</td>
<td>28.24 (0.21)</td>
<td>28.29 (0.12)</td>
</tr>
<tr>
<td>Geriatric Depression Scale, score</td>
<td>2.31 (0.42)</td>
<td>2.79 (0.37)</td>
</tr>
</tbody>
</table>

Note. All participants had Clinical Depression Rating (CDR) scale global scores of 0.5 at first visit. Means ± standard errors, unless otherwise noted. For continuously distributed variables, effect sizes are $\eta^2$ and $p$ values are based on one-way analysis of variance (ANOVA). For frequency counts (sex), effect size is Cramer’s $V$ and $p$ value is based on Pearson’s chi-squared test. MCI = mild cognitive impairment.

(F(3,104) = .91; $p > .05$), and self-reported symptoms of depression ($F(3,104) = .90; p > .05$), but their CDR sum of boxes (CDR-SB) differed ($F(3,104) = 5.12; p = .002$), with AS patients having lower scores (less impairment) than the three other groups.

Group Outcomes

From the 104 MCI patients at baseline, 76 remained stable (73.1%), 19 progressed to CDR $\geq 1$ (18%, or 9% per year), and 9 reverted to CDR = 0 (9%, or 4.5% per year) at 2-year follow-up (see Figure 1). The presence of impairment in memory (i.e., aMCI vs. naMCI) was not associated with progression to dementia ($\chi^2 = 1.97; df = 2; p > .05$), whereas the presence of impairments in multiple domains (i.e., sdMCI vs. mdMCI), regardless of the specific domains involved, was so associated ($\chi^2 = 18.34; df = 2; p < .001$). More specifically, of the 51 mdMCI individuals at baseline, 16 (31%) progressed to dementia over a 2-year period, compared with only 3 of 53 (6%) sdMCI patients. None of the mdMCI patients reverted to normal, whereas 9 (17%) of the sdMCI did.

Detailed information on the baseline and follow-up neuropsychological test scores and clinical characteristics of the participants by MCI subgroup is presented in Table 3.

Baseline Demographic, Clinical and Cognitive Characteristics of Groups Defined by Outcome Status

To further explore the outcome of the MCI patient groups, we compared the baseline characteristics and performances on the screening measures of those who converted to dementia ($n = 19$) with those who remained MCI ($n = 76$). The groups did not differ on age ($F(1,92) = 2.14; p = .147$), years of education ($F(1,92) = 0.02; p = .890$), sex distribution ($\chi^2 = 0.28; df = 1; p = .595$), MMSE scores ($F(1,92) = 3.50; p = .065$) or ratings on the CDR-SB ($F(1,92) = 2.83; p = .096$). However, those who developed dementia had lower initial scores on the Category Fluency Test ($F(1,92) = 8.59; p = .004$), the Clock Drawing Test ($F(1,92) = 4.82; p = .031$), and the delayed recall of Logical Memory of the WMS-R ($F(1,92) = 7.95; p = .006$) than those who remained MCI after 2 years.

Predictors of Progression to Dementia in MCI

Univariate logistic regression models

Univariate logistic regression models were used to examine the predictive value for dementia of each measure separately. Demographic characteristics in our sample were not associated with the CDR score after 2 years. However, ratings on specific measures of everyday functioning, in particular the ADCS ADL-PI and the IQCODE, predicted progression to CDR $\geq 1$ at the follow-up visit. Higher scores on both measures (indicating worse functioning) were associated with increase in the likelihood of developing dementia at 2 years. In addition, lower scores on specific cognitive screening tests at baseline—the Clock Drawing Test, the delayed recall of the WMS-R Logical Memory, and the Category Fluency—were associated with higher likelihood of having dementia at 2 years. From the 18 executive function measures, only three were predictive of decline over the 2-year period. Lower scores on the Hayling Test, the Alternate Uses Test, and the Verbal

Fig. 1. Clinical status at 2-year follow-up as a function of group at entry. F/U = follow-up; CDR = Clinical Dementia Rating scale.
### Table 3. Baseline and follow-up scores and clinical characteristics of study participants by subgroup over the 2-year interval

| Subgroup | Logical Memory WMS-R | Boston Naming Test | Phonemic Fluency | Category Fluency | Clock Drawing Test | Alternate Uses Test | Random Number Generation | Tinker Toy Test | D-KEFS Stroop Test | Hayling Test | Completions & Corrections Test | Porteus Maze Test | D-KEFS Tower Test | Tic-Tac-Toe | D-KEFS Sorting Test | Brixton Spatial Anticipation Test | Verbal Concept Attainment Test | Stanford Binet Absurdities Test | Iowa Gambling Testing | Experimental Judgment Test | Trail Making Test | Brief Test of Attention | TEA Telephone Search While Counting | Geriatric Depression Scale | ADCS ADL-PI | IQCODE | DEX |
|----------|----------------------|--------------------|-----------------|-----------------|-------------------|-------------------|----------------------|-------------------|-------------------|-------------|----------------------|-------------------|-------------------|----------|----------------------|---------------------------------|-------------------|-------------------|-------------|----------------------|-------------------------|-------------------|-------------|--------|--------|
| Single Domain | 4.78 (0.51) | 28.04 (0.76) | 40.59 (2.19) | 22.73 (1.30) | 9.57 (0.25) | 8.43 (0.38) | 0.68 (0.04) | 8.43 (0.38) | 10.19 (0.68) | 4.75 (0.30) | 9.85 (0.39) | 14.98 (0.49) | 9.79 (0.52) | -4.71 (0.42) | 11.86 (0.54) | 4.68 (0.41) | 16.89 (0.70) | 29.29 (0.59) | 4.37 (1.26) | 9.74 (0.97) | 69.43 (9.44) | 14.46 (0.69) | 1.45 (0.56) | 2.33 (0.44) | 3.00 (0.79) | 3.24 (0.07) | 13.44 (2.23) |
| Multiple Domain | 2.71 (0.46) | 24.11 (0.67) | 36.41 (1.95) | 22.31 (1.12) | 7.67 (0.88) | 7.75 (0.35) | 0.63 (0.04) | 7.75 (0.35) | 9.35 (0.61) | 3.84 (0.28) | 9.23 (0.37) | 13.69 (0.46) | 9.25 (0.49) | -4.45 (0.40) | 11.22 (0.51) | 4.22 (0.38) | 15.03 (0.66) | 26.58 (0.54) | 3.10 (1.19) | 9.75 (0.89) | 66.68 (8.97) | 12.19 (0.65) | 3.57 (0.53) | 2.86 (0.39) | 4.41 (0.70) | 3.49 (0.06) | 15.84 (2.00) |

**Baseline scores**

**Follow-up scores**

| Subgroup | Logical Memory WMS-R | Boston Naming Test | Phonemic Fluency | Category Fluency | Clock Drawing Test | Alternate Uses Test | Random Number Generation | Tinker Toy Test | D-KEFS Stroop Test | Hayling Test | Completions & Corrections Test | Porteus Maze Test | D-KEFS Tower Test | Tic-Tac-Toe | D-KEFS Sorting Test | Brixton Spatial Anticipation Test | Verbal Concept Attainment Test | Stanford Binet Absurdities Test | Iowa Gambling Testing | Experimental Judgment Test | Trail Making Test | Brief Test of Attention | TEA Telephone Search While Counting | Geriatric Depression Scale | ADCS ADL-PI | IQCODE | DEX |
|----------|----------------------|--------------------|-----------------|-----------------|-------------------|-------------------|----------------------|-------------------|-------------------|-------------|----------------------|-------------------|-------------------|----------|----------------------|---------------------------------|-------------------|-------------------|-------------|----------------------|-------------------------|-------------------|-------------|--------|--------|
| Single Domain | 10.82 (0.57) | 26.48 (0.84) | 35.35 (2.37) | 27.44 (1.38) | 8.57 (0.27) | 12.04 (1.05) | 0.67 (0.04) | 8.65 (0.42) | 9.39 (0.74) | 4.26 (0.33) | 9.68 (0.43) | 14.08 (0.54) | 9.43 (0.58) | -5.04 (0.46) | 11.04 (0.60) | 4.61 (0.45) | 16.57 (0.77) | 27.70 (0.65) | 2.74 (1.36) | 7.16 (0.10) | 87.48 (10.42) | 13.39 (0.77) | 1.58 (0.61) | 1.91 (0.49) | 2.63 (0.99) | 3.35 (0.09) | 17.19 (2.78) |
| Multiple Domain | 9.92 (0.74) | 23.69 (1.12) | 33.00 (3.15) | 23.77 (1.84) | 7.46 (0.36) | 7.75 (1.45) | 0.70 (0.06) | 6.58 (0.58) | 8.92 (1.02) | 3.92 (0.45) | 9.45 (0.61) | 10.67 (0.75) | 8.25 (0.80) | -5.55 (0.67) | 10.25 (0.83) | 3.00 (0.62) | 15.83 (1.07) | 24.33 (0.90) | 2.20 (2.06) | 6.47 (1.46) | 90.42 (14.42) | 10.83 (1.06) | 2.77 (0.88) | 2.54 (0.64) | 6.70 (1.25) | 3.62 (0.12) | 26.60 (3.52) |

**Note.** Means (±SE). MCI = mild cognitive impairment; WMS-R = Wechsler Memory Scale-Revised; D-KEFS = Delis-Kaplan Executive Function System; TEA = Test of Everyday Attention; ADCS ADL-PI = Activities of Daily Living-Prevention Instrument developed by the Alzheimer’s Disease Cooperative Study Committee; IQCODE = Informant Questionnaire on Cognitive Decline in the Elderly; DEX = Dysexecutive Questionnaire.
Concept Attainment Test were associated with higher likelihood of having a CDR score ≥ 1. Table 4 presents the results of the univariate logistic regression analyses for the variables that significantly predicted CDR score at the follow-up visit.

Hierarchical logistic regression model

Demographic characteristics (Block 1) were forced to enter into the model first. Among the measures of everyday functioning (Block 2), only ADCS ADL-PI score entered, accounting for 12.3% of the variance in the outcome. From the screening battery (Block 3), WMS-R Logical Memory (delayed recall) entered at step 1, Category Fluency at step 2 and Clock Drawing Test at step 3, explaining an additional 24.9% of variance. Beyond this, none of the EC measures contributed independently to the prediction of progression from MCI to dementia. Together, the four test variables in the model accounted for 41.3% of the variance in the outcome. Overall classification accuracy of the model was 84.7% (see Table 5).

Omnibus logistic regression model

The logistic regression analysis with entry of all variables that were significant in the univariate analysis together in a stepwise manner resulted in findings similar to those of the hierarchical regression model. Specifically, the ADCS ADL-PI entered the model first [accounting for 12.8% of the variance of the outcome; odds ratio (OR) = 1.179; 95% confidence interval (CI); 1.043–1.333; p = 0.009], followed by the WMS-R Logical Memory (11.6% variance explained; OR = 0.820; 95% CI: 0.697–0.963; p = 0.016), then the Category Fluency (12.2% variance explained; OR = 0.880; 95% CI: 0.790–0.981; p = 0.021), and the Clock Drawing Test (6.7% variance explained; OR = 0.623; 95% CI: 0.399–0.972; p = 0.037). No demographic characteristic and no EC measure entered the model. These predictors explained a total of 40.3% of the variance of the outcome.

Sensitivity and specificity of the prediction models

ROC analysis of the “probability of progression to CDR ≥ 1” score that results from the combined predictors of the hierarchical regression model indicated high discriminative accuracy (see Figure 2; AUC = .847; CIs = .742–.953). Using a probability for dementia cutoff value of 0.146, the model yielded a sensitivity of 0.94 and a specificity of 0.71. Thus, 94% of the MCI patients who progressed to dementia had a probability score higher than 0.146, whereas 71% of stable MCI patients had a probability score of less than 0.146.

| Table 5. Hierarchical logistic regression predicting functional/cognitive status (MCI or dementia) at 2 year follow-up as a function of demographic characteristics, and baseline clinical and cognitive factors |
|---|---|---|---|---|---|
| Block | Step | Factor | β | p | Odds ratio (95% CI) |
| Block 1 | | Age | .046 | .250 | 1.047 (.968–1.133) |
| | | Education | −.012 | .920 | .988 (.787–1.241) |
| | | Sex | .525 | .351 | 1.690 (.561–5.095) |
| Block 2 | Step 1 | ADCS ADL-PI | .168 | .010 | 1.183 (.104–1.344) |
| Block 3 | Step 1 | Logical Memory WMS-R (delayed recall) | −.215 | .015 | .806 (.678–.958) |
| | Step 2 | Word List Generation (semantic categories) | −.126 | .031 | .881 (.786–.988) |
| | Step 3 | Clock Drawing Test | −.447 | .054 | .639 (.406–1.008) |
| Block 4 | | None | | | |

Note: MCI = mild cognitive impairment; CI = confidence interval; ΔR² = Nagelkerke R Square change; ADCS ADL-PI = Activities of Daily Living—Prevention Instrument developed by the Alzheimer’s Disease Cooperative Study Committee; WMS-R = Wechsler Memory Scale-Revised.
DISCUSSION

The present study is, to our knowledge, the first systematic and comprehensive study of the predictive value of EC for the outcome of MCI patients. During a 2-year period, 18% of the MCI individuals progressed to dementia, corresponding to an annual rate of 9%. The majority of the patients remained stable over time, whereas 9% returned to normal cognitive status. We found that MCI individuals with impairments in multiple domains at baseline were at higher risk of decline over the 2-year period than those with impairment in a single domain (Alexopoulos et al., 2006b; Loewenstein et al., 2009; Manly et al., 2008; Mitchell et al., 2009; Nordlund et al., 2010). One plausible explanation is that MCI might represent a more advanced stage of the disease process than SDMCI (Alexopoulos, Grimmer, Perneczky, Domes, & Kurz, 2006a). It is also possible that some SDMCI individuals might have always performed poorly in a specific cognitive domain (Petersen et al., 1999) or they did so temporarily in response to one or more stressful life events, such as illness or death of a family member. Regardless of the explanation, and considering the fact that only SDMCI reverted to normal, it appears that an isolated cognitive impairment in an elderly person may be relatively benign.

Whether the presence of memory impairments increases the risk for subsequent dementia among persons with MCI is controversial. As in several previous studies, we found that memory impairment was not specifically associated with an increased rate of progression to dementia (Mitchell et al., 2009; Nordlund et al., 2010). However, others concluded that MCI individuals with isolated memory impairment were more likely to progress to dementia than those with a single non-memory impairment or even those with impairments in multiple domains (Ravaglia et al., 2006; Yaffe, Petersen, Lindquist, Kramer, & Miller, 2006). Manly et al. (2008) found that MCI individuals with memory impairments are at highest risk, especially when other cognitive deficits are present. Of note, a large community-based study concluded that, although a high-risk group for dementia, aMCI is unstable and heterogeneous; significant proportions remain stable or improve on long-term follow-ups (Ganguli et al., 2004).

A major finding of the present study is the limited predictive power of EC measures for the development of dementia. Only 3 of 18 EC measures—the Alternate Uses Test, the Hayling Test, and the Verbal Concept Attainment Test—were significantly associated with incident dementia at the univariate level. These tests assess spontaneous flexibility and generativity, inhibition of prepotent responses and concept rule learning, respectively. However, all three tests also rely on semantic memory. Both Hayling and Alternate Uses Test require the participant to inhibit a semantically constrained response and the Verbal Concept Attainment Test is very highly correlated with verbal tests such as the Wechsler Adult Intelligence Scale (WAIS) Vocabulary (Belleville, Rouleau, & Van der Linden, 2006; Bornstein, 1982). Thus, their dependence on semantic memory might also contribute to their sensitivity to predict cognitive decline and dementia.

Although these three tests were found to be predictive at the univariate level, none of them added to the predictive ability of a model that first considered demographic characteristics, cognitive screening tests, and everyday functioning. Our observed lack of predictive value of EC over other measures is in contrast to the reported findings of several longitudinal studies. Specifically, measures of set-shifting and sequencing (Albert et al., 2001; Chen et al., 2000), attention (Amieva et al., 2004; Tierney et al., 1996), and abstract reasoning (Elias et al., 2000) have been identified as predictors of incident dementia in MCI patients. However, in none of these studies was the prognostic value of EC examined after controlling for other factors (i.e., demographic characteristics, non-executive cognition, and everyday functioning). Manly et al. (2008) reported that MCI patients with isolated executive impairment were less likely to have underlying AD neuropathology and progress to dementia than those with an isolated memory or isolated language impairment. Similarly, Farias, Mungas, Reed, Harvey, and DeCarli (2009) indicated that executive dysfunction was not associated with an increased risk of progression to dementia over other factors such as age, education, clinic recruitment source, and functional status.

There are many confounding factors when one is trying to assess the predictive value of neurocognitive measures at baseline for incident dementia. One of the parameters that has not been adequately emphasized is that the predictive value of a measure is conditional on the other measures in the model. Differences in the variables included in the predictive model might explain why, for example, the Trail Making Test has been found to be the strongest predictor in several studies (Albert et al., 2001; Chen et al., 2000; Crowell et al., 2002; Daly et al., 2000), but not in others (Grober et al., 2008; Mitchell et al., 2009). Our results suggest that the predictive utility of EC over other cognitive measures might have been
Outcome in mild cognitive impairment

overestimated. An alternative interpretation is that most cognitive measures, especially those of any complexity, have substantial executive requirements (Salthouse, Atkinson, & Berish, 2003), and they consequently restrict the unique variance that “pure” EC measures can account for. Indeed, both the Clock Drawing Test and the Category Fluency Test rely on executive skills such as planning, search strategies, self-monitoring, and organization (Henry & Crawford, 2004; Lewis & Miller, 2007; Lowery et al., 2003). Finally, other factors, such as the interval between the baseline and the follow-up assessment, might affect the sensitivity of a test to detect decline, since it can be related to the stage of the potential preclinical course of the MCI patient (Grober et al., 2008).

Among all the cognitive measures, the delayed recall of the Logical Memory, the Category Fluency and the Clock Drawing Test were the most sensitive predictors of MCI outcome. Impairment in episodic memory and semantic fluency have been shown to be among the earliest cognitive changes that are consistent over very long follow-ups in contrast to deficits in other cognitive domains, which might appear at early stages but tend to be unstable over time (Hodges, Erzinclioglu, & Patterson, 2006). The Category Fluency Test and the Clock Drawing Test implicate executive processes but are also dependent on intact semantic memory (Henry & Crawford, 2004; Hodges, Salmon, & Butters, 1992; Lowery et al., 2003). There is also considerable evidence that verbal generativity and “idea density” is significantly compromised several years before dementia is diagnosed (Oulhaj, Wilcock, Smith, & de Jager, 2009; Riley, Snowden, Desrosiers, & Markesbery, 2005; Snowden et al., 1996). In addition, semantic fluency appears dependent on the integrity of the temporal neocortex, a brain region proportionately affected very early in the Alzheimer’s disease (Arnold, Hyman, Florý, Damasio, & Van Hoesen, 1991; Baldo, Schwartz, Wilkins, & Dronkers, 2006; Henry & Crawford, 2004).

Changes in daily functioning measured with the ADCS ADL-PI, and the IQCODE at baseline were also found to have significant prognostic value for the MCI outcome. ADCS ADL-PI measures minor changes in functional status that occur in the transition from cognitively normal to MCI or early dementia and IQCODE is considered “a measure of the disability caused by cognitive decline” (Jorm et al., 1996, p. 137). Unlike the cognitive tests administered at baseline, these are measures of change or deterioration. Thus, these findings may simply indicate that prior change predicts future change. Of note, both measures had only weak correlations with the EC measures (ranging from r = −.009 to r = −.232) and their exclusion from the regression model did not change the contribution of EC measures for the prediction of MCI outcome. Several other studies have documented that worse ratings on functional measures at baseline are strongly related to a diagnosis of dementia at follow-up (Daly et al., 2000; Dickerson, Sperling, Hyman, Albert, & Blacker, 2007; Morris & Cummings, 2005; Peres et al., 2006; Rozzini et al., 2007). Finally, the increased risk of dementia in clinic-versus population-based studies appears to be explained by the greater baseline functional impairment in the MCI patients recruited from clinics (Farias et al., 2009). Our results underline further the necessity to carefully assess everyday functioning in persons at risk for dementia.

The present study has certain limitations. First, although our MCI sample is relatively large, it is probably not adequate for the number of predictors we tested. We recognize that our analyses are rather exploratory, but the consistency of our findings using different regression models supports the robustness of our conclusions. Second, our criteria for defining MCI may be questioned, since there is no consensus on how the Petersen/Mayo criteria should be operationalized. However, we applied both clinical criteria (interview with an informant, yielding a CDR score of 0.5), as well as psychometric criteria (based on well-recognized neuropsychological procedures). Similar methodology has been implemented in several other studies (Alexopoulos et al., 2006b; Dickerson et al., 2007; Grundman et al., 2004; Storandt, Grant, Miller, & Morris, 2006) and conforms to the current standard for the diagnosis of MCI (Petersen, 2004). Third, we operationalized the development of dementia as progression from CDR = 0.5 to CDR ≥ 1. Participants were not formally evaluated by a clinician in this study (although many were so-evaluated and diagnosed in other research studies or clinics) and we cannot exclude the possibility that a participant could be assigned a global CDR score ≥ 1 for other reasons (i.e., caregiver distress). However, CDR has been used in numerous studies to identify both MCI and dementia cases (Daly et al., 2000; DeCarli et al., 2004; Storandt et al., 2006) and neuropathological findings have validated the utility of CDR scores to detect the presence of AD (Berg, McKeel, Miller, Baty, & Morris, 1993; Morris, 1997).

In summary, our 2-year follow-up of MCI individuals confirmed previous findings that persons with multiple cognitive impairments are at higher risk for functional decline than those with isolated deficits. EC has limited incremental validity as a predictor of which MCI individuals progress from CDR = 0.5 to CDR ≥ 1. However, other cognitive measures (Logical Memory of the WMS-R, Category Fluency and the Clock Drawing Test) and measures of everyday functioning (ADCS ADL-PI) can predict the fate of the 0.5s with reasonably high accuracy. It still remains unclear whether EC can predict which MCI individuals revert to normal cognitive status, when followed over time.

ACKNOWLEDGMENTS

The authors thank Laura Wulff, Ph.D., Chiadi Onyike, M.D., Marilyn Albert, Ph.D., and the staff and participants of the Johns Hopkins Alzheimer’s Disease Research Center. Eleanor Neijstrom, M.S., Jaclyn Samek, B.S., and Kevin Manning, M.S., collected the data. This study was supported by the National Institute on Aging (J.B., grant AG-005146). No conflicts of interest exist.

REFERENCES


Outcome in mild cognitive impairment


