Executive Dysfunction in Alzheimer Disease

Margaret M. Swanberg, DO; Rochelle E. Tractenberg, PhD, MPH; Richard Mohs, PhD; Leon J. Thal, MD; Jeffrey L. Cummings, MD

Background: Executive dysfunction (EDF) is common in Alzheimer disease (AD); however, its relationship to other symptoms is difficult to assess in patients with AD.

Objectives: To determine the prevalence of EDF and study its relationship to cognitive, functional, and neuropsychiatric symptoms in patients with AD.

Design, Setting, and Patients: A retrospective analysis of data from participants in the English Instruments Protocol of the Alzheimer’s Disease Cooperative Study. Subjects were drawn from a sample of patients evaluated at tertiary referral centers.

Results: A total of 64% of AD patients were classified as having EDF. Patients with EDF performed worse on tests of cognition ($P<.001$), dementia severity ($P<.001$), and activities of daily living ($P=.01$) and had more frequent symptoms of psychosis ($P=.03$) with greater emergence during the 12-month interval ($P=.03$) compared with patients with normal executive function. Less than 30% of the variance in executive function performance was explained by cognitive measures.

Conclusion: These findings support the assessment of executive function in persons with AD and the importance of frontal lobe dysfunction in AD.

Arch Neurol. 2004;61:556-560

ALZHEIMER DISEASE (AD) affects an estimated 4 million people in the United States and is the leading cause of late-onset dementia worldwide. Its core features include impairments in memory, visuospatial functions, language, executive functions (EFs), and neuropsychiatric symptoms. Defined as the ability to abstract, plan, organize, set, and adapt current and past knowledge to future behavior, EF occurs in AD, but its prevalence and relationship to other clinical and demographic features of the disease are unknown. Recent studies suggest that executive dysfunction (EDF) is a common manifestation of AD and occurs in all stages of the illness, although it is more mild than in the frontotemporal lobar degenerations. There also is evidence to support a frontal variant of AD. Johnson et al conducted a clinicopathologic study on a sample of 16 patients with AD and found that there was a subset that had early and prominent impairment on tests of EF; other test scores were similar across the groups. Autopsy studies of the frontal variant patients revealed significant increases in the number of neurofibrillary tangles in the frontal cortex compared with patients with more typical AD. Several studies using standard neuropsychological tests of EF have demonstrated links among EDF and functional decline measured by poor performance on activities of daily living (ADL) scales. Using functional imaging and neuropathologic data, a link has been found between frontal involvement and psychosis in AD and between agitation and frontal dysfunction, suggesting that EDF might be linked to behavioral disorders.

In routine clinical settings, standardized tests of EF that were not specifically designed to test EF in patients with AD, such as the Stroop color word interference test, the Wisconsin Card Sorting Test, and the Controlled Oral Word Association Test, result in floor effects for many AD patients, suggesting that better measures of assessing EF in patients with dementia are needed. The original Alzheimer’s Disease Assessment Scale—Cognitive portion (ADAS-Cog) was recently modified to include 2 tests of EF—letter cancellation and mazes—that are easy to administer and can be performed by AD patients in mild and moderate stages of the disease. The prevalence of abnor-
malities of EF as measured by these tests and the association between the executive measures behavior and function have not been systematically evaluated.

This project sought to define EDF using ADAS-Cog measures and to estimate the prevalence of EDF in a sample of AD patients drawn from specialty referral centers. We hypothesized that patients with EDF as identified by ADAS-Cog tasks would have (1) worse general cognitive function, (2) greater progression of cognitive impairment measure 12 months after baseline, (3) more impairment of ADL, and (4) more abnormal behaviors.

METHODS

SUBJECTS

We retrospectively analyzed responses from cognitively normal, elderly controls (NECs) and a subset of patients who had participated in the English Instruments Protocol of the Alzheimer’s Disease Cooperative Study but were not participants in the treatment of agitation protocol. The English Instruments Protocol was specifically designed to test new measures that might be used in clinical trials and to determine their psychometric properties. Patients did not receive experimental treatment in the case of this study. A total of 137 subjects (62% female) with probable AD diagnosed using the National Institutes of Neurological and Communicative Disorders and Stroke-Alzheimer’s Disease and Related Disorders Association criteria and with Mini-Mental State Examination (MMSE) scores higher than 10 were identified for inclusion in these analyses. A control sample of 64 cognitively normal volunteers (57% female) was included for comparison. All subjects were older than 45 years (mean ± SD age, 70.3 ± 8.8 years for controls and 72.9 ± 8.4 years for AD patients), with a minimum of 6 years of education (mean ± SD years of education, 13.8 ± 2.9 for controls and 13.3 ± 2.8 for AD patients), and had reliable caregivers. Subjects were free of preexisting psychiatric illness, including schizophrenia, recurrent depression, and substance abuse. Those requiring use of psychoactive agents, with a history of significant head trauma, or with a history of significant medical problems, or with a history of significant head trauma were excluded. A complete description of inclusion and exclusion criteria used in this protocol is found elsewhere.

PROCEDURES

All subjects were evaluated at baseline and again at scheduled intervals during the next 12 months. Information and test results obtained at baseline and 12 months were included in this study. Six participants in the AD cohort did not have valid EF test data at baseline and were therefore not included, so our sample size was 131.

EXECUTIVE FUNCTION

Executive function was tested using the most complex of a series of 6 cancellation tasks and the time to complete the first 3 mazes from the expanded ADAS-Cog. The letter cancellation task tests the subject’s ability to concentrate and use appropriate search strategies. During this “either of 2 numbers” task, the patient is asked to cross out either of 2 numbers (eg, 3 or 7) that were randomly mixed in with other numbers on a sheet of paper. The score is obtained by subtracting the number of incorrectly crossed off items and the number of reminders given from the number of correctly crossed off items, with a minimum score imposed as 0. A maximum score is 40; lower scores indicate worse performance.

The maze task assesses impulse resistance, planning, reasoning, and foresight. During the maze task, subjects are given a series (up to 7) of increasingly difficult mazes to complete as quickly as possible. Mohs et al recommended that the first 3 mazes be used in future studies because these could be completed by most AD patients, thus limiting floor effects. One incorrect decision is allowed, after the second “dead end” the maze was discontinued, and the maximum time per maze was assigned. The maximum time allowed to complete each maze is 240 seconds, yielding a total maximum time of 720 seconds. Higher scores indicate worse performance.

We defined EDF as scores more than 1.5 SDs below the mean scores obtained by the NECs. To establish the validity of this approach, we defined cutoff values for both the cancellation task and the maze times based on NEC performance at both the baseline and 12-month visits.

OTHER MEASURES

The cognitive function of patients was measured by the MMSE, a 30-point test that assesses the domains of attention, orientation, calculation, memory, language, and visuospatial functioning, and with the Clinical Dementia Rating Scale sum of boxes (CDR SB). The CDR SB has a range of 0 to 18, representing the sum of 6 individual domains in the instrument; higher scores indicate worse dementia. Behavioral disturbances were assessed using the Cohen-Mansfield Agitation Inventory (CMAI), a 36-item, informant-based scale that rates behaviors observed in the past 2 weeks, and the Behavior Rating Scale for Dementia (BRSD), a 48-item, informant-based scale that rates behaviors that have occurred in the last month. The BRSD has been subjected to a factor analysis, with 6 factors being identified: behavioral dysregulation, depression, inertia, irritability/aggression, psychosis, and vegetative symptoms. The Alzheimer’s Disease Cooperative Study Activities of Daily Living Inventory, which assesses both basic ADLs, such as bathing, grooming, walking, and dressing, and instrumental ADLs, such as handling mail, discussing current events, and using household appliances, was administered to measure functional impairment. The range of scores is 0 to 78, with higher score indicating better function.

STATISTICAL ANALYSIS

The distribution of the control scores on each task at baseline and 12 months was evaluated to determine if establishing cutoff values based on the mean and standard deviation was appropriate. We used Shapiro-Wilk tests of normality and examined skew, kurtosis, and quantile-quantile plots of the control scores. Means and standard deviations for the 2 EF measures in NECs were calculated at baseline and 12 months so that 4 cutoffs were established to evaluate the validity of classifications of AD patients as EDF. However, only cutoffs based on normally distributed control scores at baseline were used in the exploration of the relationships between EF and cognitive, functional, and behavioral variables.

After determining the most appropriate way of identifying AD patients as having EDF or normal EF (NEF), unpaired t tests were used to explore the association between EDF and the CDR SB, MMSE, ADL Inventory, CMAI, and BRSD at baseline and 12 months. Holm adjustment for multiple comparisons was used, and adjusted P values less than .05 were considered statistically significant.

Emergence of BRSD subscores was calculated as the number of items in that subscore that emerged divided by the number of items in that subscore that were eligible to emerge. Eligibility for emergence was defined as the proportion of BRSD items given a frequency rating of 0 or 1 (symptoms present 2 or fewer days in the previous month) at the baseline visit and
Association Between EDF and NEF and Other Domains in 137 Patients With AD*

<table>
<thead>
<tr>
<th>Variable</th>
<th>AD Patients With EDF</th>
<th>AD Patients With NEF</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education, y</td>
<td>12.9 ± 2.7</td>
<td>14.2 ± 3.0</td>
<td>.14</td>
</tr>
<tr>
<td>Duration of AD, y</td>
<td>4.8 ± 3.3</td>
<td>4.5 ± 1.9</td>
<td>.52</td>
</tr>
<tr>
<td>Age, y</td>
<td>72.4 ± 8.4</td>
<td>74.0 ± 8.5</td>
<td>.62</td>
</tr>
<tr>
<td>CDR SB, baseline</td>
<td>8.5 ± 2.6</td>
<td>6.8 ± 2.3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>CDR SB, change†</td>
<td>-2.9 ± 2.4</td>
<td>-1.4 ± 2.0</td>
<td>.01</td>
</tr>
<tr>
<td>MMSE, baseline</td>
<td>17.2 ± 4.2</td>
<td>21.4 ± 4.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>MMSE, change†</td>
<td>-5.5 ± 4.4</td>
<td>-2.6 ± 4.6</td>
<td>.02</td>
</tr>
<tr>
<td>Cancellation task, baseline, No. of hits†</td>
<td>8.2 ± 4.7</td>
<td>22.4 ± 5.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>3 Maze times, baseline, s†</td>
<td>243.0 ± 157.5</td>
<td>100.3 ± 106.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>ADL, baseline†</td>
<td>48.3 ± 14.3</td>
<td>57.1 ± 14.2</td>
<td>.01</td>
</tr>
<tr>
<td>CMAI, baseline</td>
<td>22.9 ± 19.3</td>
<td>17.4 ± 13.8</td>
<td>.54</td>
</tr>
<tr>
<td>BRSD-dep, baseline</td>
<td>26.3 ± 16.5</td>
<td>20.7 ± 15.0</td>
<td>.46</td>
</tr>
<tr>
<td>BRSD-psych, baseline</td>
<td>4.8 ± 5.0</td>
<td>3.0 ± 3.3</td>
<td>.14</td>
</tr>
<tr>
<td>BRSD-psych, baseline</td>
<td>1.8 ± 3.1</td>
<td>0.5 ± 1.3</td>
<td>.01</td>
</tr>
</tbody>
</table>

*Data are presented as mean ± SD.
†Change during 12 months.

that were given a frequency rating higher than 2 (symptoms present at least 3 days in the month before the 12-month visit) at the 12-month visit. Nonparametric t tests were used to compare emergence rates for patients characterized as EDF or NEF.

Finally, the 2 cognitive severity scores were univariately regressed on the baseline score on which the EDF classification was based. Using this approach, we estimated the proportion of variance in EDF task performance that could be explained by cognitive functioning.

**RESULTS**

**DEFINITION OF EDF BASED ON NEC SCORES**

Distribution of NEC scores for maze times was not normally distributed, but distribution of NEC scores for letter cancellation was, making cancellation task scores more appropriate for defining EDF, estimating prevalence, and exploring relationships between EDF and other domains. The cutoff scores for the cancellation tasks were 37.2/40 for the baseline visit and 37/40 for the 12-month visit. Any person with scores at or better than these levels was classified as having NEF for that visit. For the maze times, the cutoff scores were 88.2/720 seconds for the baseline visit and 69.4/720 seconds for the 12-month visit. Any person with a sum of maze times at or faster than this was classified as having NEF at that visit.

Using the cutoff scores at baseline for letter cancellation, 6% of NECs and 64% of AD patients were classified as having EDF. Based on maze times at baseline, 2% of NECs and 58% of AD patients were classified as having EDF. Neither floor nor ceiling effects were present.

Misclassification, defined as an EDF label at baseline and an NEF label at 12 months, was minimal using the cancellation task: 3 of the 64 controls were labeled as having EDF at baseline; only 1 was misclassified (scored as having NEF) based on 12-month results. Two (3%) of 64 AD patients labeled as having EDF at baseline were labeled as having NEF at 12 months. However, using the maze times criteria, 13 (22%) of 58 AD patients labeled as having EDF at baseline were labeled as having NEF at 12 months; no controls were misclassified based on maze times.

**ASSOCIATION BETWEEN EDF AND OTHER DOMAINS**

The cancellation task at baseline was used to classify the AD cohort as having EDF or NEF because of the low degree of misclassification and its normal distribution. We therefore explored the association between EDF or NEF status and scores from tests of other domains in this cohort.

The Table presents the means and standard deviations of the scores across the AD cohort grouped as EDF or NEF, as well as the results of independent-sample t tests performed to compare scores on the CDR SB and MMSE at baseline and change during the 12-month study and baseline scores of the ADL, CMAI, and BRSD total and subscores. Age, years of education, and duration of dementia were not different between subjects in the EDF and NEF categories.

Patients classified as having EDF had significantly more severe dementia (based on CDR SB), worse cognitive functioning (MMSE score), poorer ADL scores, and more frequent symptoms of psychosis at baseline. These individuals also demonstrated significantly greater worsening in terms of dementia and cognitive functioning 12 months after baseline. Similar results were seen when EDF and NEF were based on the sum of maze times.

The emergence of psychosis during 12 months in those patients with EDF at both baseline and 12 months (n=61) measured by cancellation task scores was 5 times (7.2%) the rate observed in those with NEF at both baseline and 12 months (P=.03). Only 12 of the 34 patients classified as having NEF at baseline continued to have NEF at the follow-up visit. Psychosis emergence in those with EDF at baseline was nearly double the rate observed in those with NEF at baseline; however, this was not statistically significant (P=.20).
Univariate regression for each of the instrument scores on the baseline cancellation task scores for the AD cohort revealed significant association with baseline values on CDR SB ($F_{1,129}=28.1, P<.001$), MMSE ($F_{1,129}=48.8, P<.001$), and ADL ($F_{1,129}=26.7, P<.001$). These analyses showed that with no other variables in the model, 28% of the variance in cancellation task performance at baseline was explained by baseline MMSE score and 18% was explained by baseline CDR SB score in the AD patients.

Executive dysfunction was common in this sample of AD patients, with an estimated prevalence of 64%. Executive dysfunction was significantly associated with MMSE, CDR SB, and ADL scores and frequency of psychosis at baseline and with worsening demonstrated by change in MMSE and CDR SB scores during the 12 months. Psychosis symptom emergence also was significantly associated with the presence of EDF at the baseline and 12-month visits. The association with CDR SB and ADL scores indicates that the sample with EDF had poorer everyday and community function.

The finding that less than 30% of the variance in the EF task performance could be attributed to overall dementia severity or global cognitive status suggests that the cancellation task recommended by Mohs et al.20 is an appropriate measure of EF in this particular patient population and is not simply a measure of general cognitive decline. These data, derived from a subset of patients tested by Mohs et al.,20 suggest that the letter cancellation task may have predictive value, with poorer performance indicating greater neuropsychiatric impairment at baseline and in the future.

Several reports in the literature have examined the association between ADL performance and cognitive status. A few have specifically evaluated the association between EF and ADL. Tekin et al.6 assessed the relationship of ADL with neuropsychiatric symptoms, cognition, and medical illness burden. They found a correlation between instrumental ADL performance measured using the Functional Activities Questionnaire33 and MMSE score. However, they found that neuropsychiatric symptoms were more strongly correlated with Functional Activities Questionnaire score than MMSE performance. Other investigators have identified correlations between impaired functional abilities and EDF.32,33 When studies examined ADL performance and cognition, they showed ADL performance declined with increasing dementia severity; however, when studies examined individual cognitive domains, executive abilities accounted for most of the variance.34 Similarly, we found a strong correlation between ADL performance and EF. Patients who are unable to perform many of their hobbies, chores at home, and personal grooming (items rated in both the CDR and the ADL Inventory) because of EDF would have worse scores for these items compared with AD patients without EDF.

We found a relationship between psychotic symptoms and executive impairment in the present study, and this has been reported previously.16,35-38 In the few studies32,39 that looked specifically at the executive domain, it was found that patients with psychosis perform more poorly on executive tasks compared with their performance on tests of other cognitive functions. Our results show that patients with EDF at baseline have a greater risk of developing psychosis at the 12-month visit. This suggests that EDF may have predictive value. Functional imaging data also suggest that there is a greater degree of frontal hypometabolism and perfusion in patients with psychotic symptoms.8,14

Limitations of this study should be considered when interpreting its results. Our sample population was drawn from specialty referral centers, and findings may not be generalizable to a broader community sample of patients with AD. Sampling patients who came into referral centers may have introduced bias (ie, patients with greater EDF and more functional impairment may have been less willing to participate in this study). This bias would lead us to underestimate the prevalence of EDF in the general community. Alternatively, it is possible that patients with more impairment were recruited because physicians are more likely to refer problematic and challenging patients to referral centers. Our measures of EF are not standard measures given in neuropsychological test batteries. These measures were, however, adapted from paradigms used in cognitive and clinical neuropsychology33,35 to test EF. No concurrent validity assessment was performed, and we cannot draw conclusions about the prevalence or correlates of EDF as measured by other EF tests. Finally, some aspects of behavior mediated by the frontal lobes such as apathy are not well assessed by the BRSD or CMAI. Therefore, our behavioral measures may not have been sufficiently sensitive to the symptoms most likely to be associated with EDF and/or frontal lobe impairment.

In summary, this study has demonstrated that EDF is present in approximately 60% of a community-dwelling cohort of AD patients and is associated with greater dementia severity, worse overall cognitive status and functional impairment, and more frequent and higher emergence rates for symptoms of psychosis. Although the cancellation task was better for establishing a cutoff at which EDF could be defined, both the mazes and the cancellation task may be valuable in the clinical assessment of patients because poor performance correlates with functional decline and neuropsychiatric symptoms. These 2 brief measures of EF may be helpful to clinicians when more extensive standardized tests of EF are too difficult to be performed by AD patients. The association of EDF and psychosis emergence may help physicians and caregivers in the monitoring and treatment of these symptoms. Finally, the ADAS-Cog with the addition of letter cancellation and mazes is increasingly used in drug trials to monitor clinical response. The association we found between EDF as measured by these tests and performance on the ADL may provide additional means of assessing relationships among cognitive, functional, and behavioral changes in response to therapy.

Accepted for publication October 29, 2003.

Author contributions: Study concept and design (Drs Swanberg, Tractenberg, Mohs, Thal, and Cummings); ac-
quisition of data (Dr Mohs and Thal); analysis and interpretation of data (Drs Swanberg, Tractenberg, Mohs, and Cummings); drafting of the manuscript (Drs Swanberg and Tractenberg); critical revision of the manuscript for important intellectual content (Drs Swaberg, Thal, and Cummings); statistical expertise (Dr Tractenberg); administrative, technical, and material support (Drs Tractenberg and Mohs); study supervision (Dr Thal and Cummings).

This program was supported by the Alzheimer’s Disease Cooperative Study (National Institute on Aging grant 1-U01 AG-10483), a National Institute on Aging Alzheimer’s Disease Research Center grant (AG 16570), an Alzheimer’s Disease Research Center grant of California grant, and the Sidell-Kagan Foundation, Los Angeles.

Corresponding author and reprints: Jeffrey L. Cummings, MD, UCLA Alzheimer’s Disease Center, Department of Neurology, David Geffen School of Medicine at UCLA, 710 Westwood Plaza, Los Angeles, CA 90095-1769 (e-mail: cummings@ucla.edu).

REFERENCES