The Course of Cognitive Impairment in Preclinical Alzheimer Disease

Three- and 6-Year Follow-up of a Population-Based Sample

Brent J. Small, PhD; Laura Fratiglioni, MD, PhD; Matti Viitanen, MD, PhD; Bengt Winblad, MD, PhD; Lars Bäckman, PhD

Objectives: To examine the ability of the total score and individual items from the Mini-Mental State Examination in predicting the development of Alzheimer disease (AD) across a 3- and 6-year period in a population-based sample, and to describe the longitudinal changes in these measures across the same follow-up periods.

Design: Prospective follow-up of a community-based cohort, with 3 times of testing across a 6-year period. At each time of measurement, participants were clinically examined by physicians to identify demented and nondemented participants according to Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition, criteria.

Participants: The study population consisted of all participants who were nondemented at the first follow-up and participated in the second follow-up examination. Among those, 459 remained nondemented and 73 developed AD during the second follow-up period.

Results: Baseline differences in the total Mini-Mental State Examination score and the delayed memory item were seen 6 years before eventual dementia diagnosis (P < .01). Analysis of the longitudinal changes showed no differences in the rate of decline for the incident AD or nondemented group between time 1 and time 2 (P > .10). However, the incident AD group exhibited precipitous declines in 8 of the 10 subscales between time 2 and time 3, the point at which they were clinically diagnosed (P < .01). Logistic regression analyses showed that only the delayed memory item was a significant predictor of who would develop AD, independent of age, sex, and years of education, at both of the first 2 times of measurement (P < .001).

Conclusions: The diagnosis of AD is preceded by a long preclinical phase in which deficits in memory performance are most common. These deficits remain relatively stable up until the time that a dementia diagnosis can be rendered.

Arch Neurol. 2000;57:839-844

Identifying individuals who are at increased risk of developing Alzheimer disease (AD) is currently a topic of great theoretical interest and practical importance. As more effective pharmacological therapies become available, the administration of these agents to individuals who are subtly impaired may render the treatments more effective.

One issue that has garnered attention is the length of time during which cognitive deficits among those who will develop AD are apparent. Most studies have documented the presence of preclinical deficits across follow-up periods between 2 and 3 years, and some research has examined follow-up intervals longer than 5 years and found significant decrements in cognitive performance among individuals who will eventually develop dementia. However, a potential problem with studies that have examined individuals over longer periods is the variability in the length of follow-up. Combining individuals with diverse follow-up periods has the potential to obscure effects related to the presence of early cognitive deficits. In the present study, we observed a group of people over standardized follow-up intervals and examined the development of AD within this cohort.

Another unresolved issue deals with the specificity of preclinical cognitive deficits in terms of whether particular domains of functioning are more affected than others. It is becoming clear that measures of memory performance are the first indicators to demonstrate preclinical deficits in AD. For example, using the Mini-Mental State Examination (MMSE) item scores, we found that the delayed recall and orientation to time items were the only significant predictors of incident AD across a 3-year follow-up period. In the present study, we extend these findings by using the MMSE item scores to determine whether potential preclinical
SUBJECTS AND METHODS

SUBJECTS

The study population consisted of all subjects who were nondemented after the first follow-up and participated in the second follow-up examination in the Kungsholmen Project (n = 369). The Kungsholmen Project has been approved by the ethics committee of the Karolinska Institute in Stockholm, Sweden, and written informed consent was obtained from all participants after details of the procedure had been fully explained. This population was derived from the original cohort according to the scheme reported in Table 1. The original population included 1810 inhabitants aged 75 years and older in the Kungsholmen parish of Stockholm. Of those, 1475 participants were diagnosed as nondemented at the initial time of measurement.15 During the first follow-up (between time 1 and time 2 examinations; mean, 3.43±0.53 years), 318 died, 168 moved or refused participation, 199 developed dementia, and 790 remained nondemented.16 Across the 3-year interval (mean, 3.27±0.48 years) between time 2 and time 3, 73 people developed possible or probable AD, 37 were diagnosed as having dementia of another type, 177 died, 44 moved or refused participation, and 459 remained nondemented. The persons who developed AD between time 2 and time 3, as well as those who remained nondemented, constitute the groups of interest for the present study. Because of missing data on 4 participants, the final sample consisted of 457 nondemented adults and 71 incident cases of AD.

At each time of examination, the same physicians (L.F. and M.V.) rendered the final diagnosis of dementia and dementia type, which were based on the agreement between 2 independently derived preliminary diagnoses according to Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition,11 criteria, with some modifications.16,67 Brain imaging was not a component of the diagnostic process.

Baseline (time 1) demographic characteristics of the nondemented participants and those who were diagnosed as having AD at second follow-up are presented in Table 2. Univariate analyses indicated that the nondemented group was younger than the incident AD group but had similar sex distribution and years of education.

MEASURES

The Swedish version of the MMSE10 was administered according to standardized procedures, and the total score is out of a maximum of 30. In addition to the total score, 10 individual item scores were examined. They were orientation to time, orientation to place, word recall—immediate, word recall—delayed, naming, repetition, following commands, reading, writing, and design copy. In the present study, sustained attention was not examined at the individual item level, because a great deal of data were missing on 1 of the 2 items used to index this ability. Although administering only 1 of the 2 items to derive the total score is acceptable practice,10 the demands of the tasks are sufficiently different18,39 as to discourage us from combining across the 2 measures in examining the individual item scores. As a result, sustained attention was not examined at the item level. However, the lack of complete data on these 2 questions did not vary as a function of eventual diagnostic status (P > .10). This indicates that the delay between the immediate and delayed recall items was comparable for both groups.

DATA ANALYSIS

The data analysis consisted of 2 sections. First, mean-level analyses were conducted separately on the total score and individual MMSE items to assess whether cognitive deficits were apparent either 6 years or 3 years before diagnosis. In addition, longitudinal analyses were performed to examine changes in performance from time 1 to time 2 and from time 2 to time 3. Finally, the presence of differential change as a function of impending dementia status was inferred by examining the results of the time × diagnostic group interaction. Analyses were conducted with a multivariate analysis of covariance with dementia group (nondemented and incident AD) as the between-subjects factor, time of testing (time 1, time 2, and time 3) as the within-subjects factor, and age as a covariate. To adjust for type I error rate, a modified Bonferroni procedure20 with familywise α at .01 was used for the univariate tests of significance.

The second set of analyses used logistic regression analyses to assess the power of the MMSE items in predicting the incidence of AD. In this case, diagnostic category at time 3 was used as the outcome variable. Age, sex, and years of education at time 1 were entered in the first step and treated as covariates. The relative importance of the MMSE items was assessed with stepwise forward logistic regression procedures with α set at .01. To facilitate comparisons of the MMSE items in predicting the incidence of AD, the raw data were transformed into z scores. For the logistic regression analyses, 3 separate analyses were conducted in which the predictor variables varied. In the first analysis, the baseline (time 1) scores were used as predictors. In the next, the time 2 scores were used as predictors. In the final analysis, the predictive power of the difference scores (time 2–time 1) was examined.

Data are given as mean±SD.

deficits observed 6 years before the diagnosis of AD are specific to memory functioning or involve other ability domains.

A final issue concerns the use of change scores as predictors of impending AD. Charting of preclinical cognitive deficits in AD has come almost exclusively from cross-sectional comparisons.2,6,5 This is unfortunate, as a key characteristic in the diagnostic criteria for AD is a requirement that individuals have experienced changes from previous levels of cognitive functioning.11,12 However, there are some exceptions to this rule. Two studies13,14 reported statistically significant decline in cognitive performance preceding the diagnosis of AD. However, in both cases, this exacerbated decline was seen relatively close to the time of diagnosis; hence, it is unclear whether the differential change was present many years before eventual diagnosis. In the present study, we charted the longitudinal changes in MMSE item score performance across a 6-year follow-up period for groups of persons who did or did not develop AD.

(Reprinted) Arch Neurol/Vol 57, June 2000 www.archneurol.com

©2000 American Medical Association. All rights reserved.
RESULTS

MEAN-LEVEL DIFFERENCES AND LONGITUDINAL CHANGES

Table 3 presents the means and SDs for the MMSE items as a function of dementia status and time of measurement. Analysis of the total score revealed significant effects of diagnostic group (P < .001), time of measurement (P < .001), and a time x diagnostic group interaction (P < .01). For MMSE total score, there were significant group differences, in favor of the nondemented group, across all 3 times of measurement, with the magnitude of these decrements increasing at each successive measurement point. The longitudinal changes in performance are shown in Table 4. Between time 1 and time 2, the incident AD group exhibited significant decline in the total score (P < .001), but the nondemented group did not. Between time 2 and time 3, both groups exhibited significant longitudinal decline on the MMSE total score (P < .001), but the incident AD group declined by almost 3 SD, whereas the nondemented group declined by less than half an SD unit (Table 4).

In the analysis of the individual item scores (Table 3), the delayed memory item exhibited significant group differences at both the first and second times of measurement (P < .001). At the second time of testing, orientation to time also exhibited significant group differences in performance (P < .01). In all cases, the incident AD group performed more poorly than the nondemented group. At time 3, there were significant differences in performance, favoring the nondemented group (P < .01), for all items, with the exception of naming and repetition.

The analysis of longitudinal changes showed that the incident AD group exhibited significant 3-year decrements between time 1 and time 2 on 2 of the 10 MMSE items, repetition and design copy (P < .01; Table 4). The nondemented participants also exhibited significant negative changes on these 2 items (P < .001), as well as orientation to time (P < .01), and actually demonstrated a slight improvement in delayed word recall between time 1 and time 2 (P < .001). The absence of a significant time x diagnostic group interaction for any of the individual items indicated that both groups exhibited similar trajectories of change across the 3-year follow-up period.

Between time 2 and time 3, the incident AD group exhibited reliable longitudinal decrements on 8 of the 10 MMSE item scores (P < .01). By contrast, the nondemented group exhibited significant changes on only 6 of the items (P < .01), and the magnitude of these changes was considerably smaller than that of the incident AD group. This pattern is exemplified by the fact that the time x diagnostic group interaction was statistically significant for all items except repetition. In all cases, the incident AD group exhibited more precipitous decline from time 2 to time 3 than did persons who would remain nondemented.

PREDICTION OF INCIDENT CASES OF AD

Among the covariates, only age was a significant predictor. In this case, increased age was associated with a higher risk of developing AD (odds ratio, 1.14; 95% confidence interval, 1.08-1.21). Neither sex nor years of education was a significant predictor in the model. Among the MMSE items measured at baseline (time 1), only delayed recall added significantly to the model. For this task, individuals with higher initial performance were less likely to develop AD after 6 years (adjusted odds ratio, 0.62; 95% confidence interval, 0.47-0.82).

A similar set of predictive relationships was observed when the time 2 MMSE items were used as predictors. Only delayed recall was a significant predictor of who would eventually develop AD after 3 years (adjusted odds ratio, 0.48; 95% confidence interval, 0.36-0.63). The only difference between the two sets of analyses was the magnitude of the observed relationships. As expected, the risk ratios were higher for the time 2 scores than for the time 1 scores.

A final logistic regression analysis was conducted in which the time 2–time 1 difference scores served as predictor variables. However, this analysis indicated that none of the individual items was a significant predictor of the development of AD.

The results of the baseline mean-level analyses indicated significant differences between the two diagnostic
groups almost 7 years (mean, 6.71 ± 0.49 years) before the clinical diagnosis of AD was rendered. Specifically, the incident AD group performed more poorly on the delayed word recall measure than the group that remained free of dementia. Furthermore, significant diagnostic group differences in delayed word recall and orientation to time were observed at the second time of measurement, approximately 3 years before clinical diagnosis.

These results contribute to the growing body of literature indicating that cognitive deficits appear many years before the clinical diagnosis of AD. The time frame examined herein is consistent with that in other studies that have reported preclinical deficits appearing more than 5 years before diagnosis. However, a unique aspect of this study is that standard follow-up intervals were used for all participants. Unlike other studies whose average follow-up interval is rather long, although some patients may be followed up for only 1 or 2 years, we used a consistent time interval for all participants. This is evidenced by the nonsignificant correlation between length of follow-up period and diagnostic group (Spearman r = -0.02). This method is more desirable, as an overrepresentation of individuals who are followed up in proximity to clinical diagnosis may overestimate the magnitude of the cognitive deficits associated with preclinical AD.

The results also indicated that measures with some type of memory referent exhibited preclinical cognitive deficits and were predictive of incident AD. These results are consistent with studies that have used more comprehensive cognitive assessment batteries. Similarly, the results correspond well to our previous report that examined the predictive utility of MMSE items across a 3-year follow-up period. Taken together, our results converge on the common theme that memory deficits are the first indicators of AD. This locus of impairment is consistent with both histopathological and morphologic evidence, indicating that the earliest changes in the brains of persons who will develop AD occur in the hippocampus and neighboring regions. Lesion and imaging studies demonstrate the pivotal role of the hippocampal formation in acquiring new memories.

The final set of results demonstrated an absence of differential longitudinal effects between time 1 and time 2, both in terms of differences in the magnitude of longitudinal changes between the incident AD and nonde-

<table>
<thead>
<tr>
<th>Time</th>
<th>IAD</th>
<th>ND</th>
<th>IAD</th>
<th>ND</th>
<th>IAD</th>
<th>ND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total score</td>
<td>27.11</td>
<td>27.64†</td>
<td>26.27</td>
<td>27.57‡</td>
<td>19.56</td>
<td>26.40‡</td>
</tr>
<tr>
<td>Orientation to time</td>
<td>1.42</td>
<td>1.52</td>
<td>1.72</td>
<td>1.78</td>
<td>2.77</td>
<td>2.73</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>4.84</td>
<td>4.90</td>
<td>4.66</td>
<td>4.83‡</td>
<td>2.55</td>
<td>4.71‡</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>0.40</td>
<td>0.33</td>
<td>0.67</td>
<td>0.41</td>
<td>1.57</td>
<td>0.63</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>4.90</td>
<td>4.91</td>
<td>4.87</td>
<td>4.92</td>
<td>4.00</td>
<td>4.80‡</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>0.30</td>
<td>0.29</td>
<td>0.34</td>
<td>0.28</td>
<td>1.24</td>
<td>0.46</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>0.94</td>
<td>1.43‡</td>
<td>1.07</td>
<td>1.83‡</td>
<td>0.32</td>
<td>1.50‡</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>0.88</td>
<td>0.95</td>
<td>0.95</td>
<td>0.97</td>
<td>0.60</td>
<td>1.00</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>2.00</td>
<td>1.99</td>
<td>2.00</td>
<td>1.99</td>
<td>1.86</td>
<td>1.97</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.13</td>
<td>0.49</td>
<td>0.23</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>0.97</td>
<td>0.98</td>
<td>0.82</td>
<td>0.91</td>
<td>0.58</td>
<td>0.76</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>0.17</td>
<td>0.12</td>
<td>0.39</td>
<td>0.29</td>
<td>0.50</td>
<td>0.43</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>2.90</td>
<td>2.88</td>
<td>2.73</td>
<td>2.85</td>
<td>2.38</td>
<td>2.75‡</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>0.30</td>
<td>0.33</td>
<td>0.48</td>
<td>0.38</td>
<td>0.88</td>
<td>0.54</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.73</td>
<td>0.89‡</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>0.00</td>
<td>0.05</td>
<td>0.12</td>
<td>0.08</td>
<td>0.45</td>
<td>0.32</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>0.96</td>
<td>0.93</td>
<td>0.83</td>
<td>0.89</td>
<td>0.54</td>
<td>0.85‡</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>0.20</td>
<td>0.26</td>
<td>0.38</td>
<td>0.31</td>
<td>0.50</td>
<td>0.36</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>0.76</td>
<td>0.82</td>
<td>0.54</td>
<td>0.66</td>
<td>0.24</td>
<td>0.61‡</td>
</tr>
<tr>
<td>Orientation to place</td>
<td>0.43</td>
<td>0.38</td>
<td>0.50</td>
<td>0.47</td>
<td>0.43</td>
<td>0.49</td>
</tr>
</tbody>
</table>

* MMSE indicates Mini-Mental State Examination; IAD, incident Alzheimer disease group (n = 71); and ND, nondemented group (n = 457).
† P < .01, within-time comparison of IAD and ND.
‡ P < .001, within-time comparison of IAD and ND.
mented group, and in the inability of the difference scores to predict AD in the logistic regression analysis. At first glance, these results may seem at odds with those from other groups\textsuperscript{13,14} that report differential cognitive decline among persons who do or do not develop AD. However, in these studies, the final time of measurement that was used to index cognitive functioning was the one immediately preceding the change in diagnostic status. By contrast, the changes between time 1 and time 2 preceded the diagnosis of AD by at least 3 years. If we did use the change in diagnostic status as the final measurement point, indexing change between time 2 and time 3, our results would be entirely consistent with previous reports.

A possible reason for the lack of differential change between time 1 and time 2 may be related to the advanced age of the participants. Comparatively little is known about patterns of cognitive performance\textsuperscript{26} and development of AD\textsuperscript{46} among very old adults. However, had a younger cohort of adults been examined, one where the presence and magnitude of normal age-related cognitive deficits across this follow-up interval would have been substantially reduced,\textsuperscript{27,28} a different pattern of results may have emerged. Specifically, we might have seen differential change between the incident AD group and the nondemented group mainly because of the preservation of performance among persons free of dementia. Taken together, we found evidence of cognitive impairment, especially in the domain of memory functioning, many years before the diagnosis of AD. However, we also found evidence that the magnitude of these impairments is relatively stable, up until the point at which the diagnosis of AD is rendered.

The presence of cognitive deficits almost 7 years before the diagnosis of AD suggests the question of just how long before the clinical presence of AD cognitive deficits are apparent. Several lines of evidence suggest that these deficits may appear several decades before the eventual diagnosis of AD. La Rue and Jarvik\textsuperscript{29} noted deficits on multiple cognitive measures for individuals who would be diagnosed as having dementia 20 years later, compared with persons who would not. Similarly, Snowdon and colleagues\textsuperscript{30} reported that impoverished linguistic ability, derived from personal autobiographies written when the participants were in their 20s, was associated with the clinical expression of AD almost 60 years later. Thus, our results and those of others suggest that, although the diagnosis of AD is preceded by an exceptionally long period in which cognitive differences are present, the magnitude of these impairments may remain relatively stable up until shortly before clinical diagnosis.

Although the results of the present study are informative, there are several limitations that must be acknowledged. First, our use of the MMSE as the primary criterion measure may be criticized because of the lack of sensitivity in this instrument. In our view, the results we observe do not imply that the MMSE can be used for diagnostic purposes many years before clinical expression. Rather, we are buoyed that this instrument was able to detect prodromal deficits many years before diagnosis and believe that more sensitive, comprehensive neuropsychological batteries should be applied to detect the subtle cognitive deficits associated with preclinical AD.

A second limitation has to do with the fact that we were forced to remove the sustained attention measures from consideration. Although this was unfortunate, we have reason to believe that this decision had little effect on our ability to predict incident cases of AD. In a previous analysis of the MMSE items and 3-year incidence of AD,\textsuperscript{4} we reported that the measure of serial 7s was not a significant predictor in the logistic regression analyses. Similarly, Galasko and colleagues\textsuperscript{19} found that the measures of sustained attention did not significantly add to the predictive model for AD.

\textsuperscript{Accepted for publication January 24, 2000.}

\textsuperscript{This research was supported by grants from the Swedish Council for Research in the Humanities and the Social Sciences (Dr Bäckman), from the Swedish Council for Social Research (Drs Winblad and Bäckman), and from the Swedish Medical Research Council (Dr Fratiglioni), Stockholm, Sweden.}

\textsuperscript{We thank all members of the Kungsholmen Project Study Group for collaboration and data collection.}
Reprints: Brent J. Small, PhD, Department of Gerontology, Mailbox SOC 107, University of South Florida, 4202 E Fowler Ave, Tampa, FL 33620 (e-mail: bsmall@luna.cas.usf.edu).

REFERENCES