Longitudinal Study of the Transition From Healthy Aging to Alzheimer Disease

David K. Johnson, PhD; Martha Storandt, PhD; John C. Morris, MD; James E. Galvin, MD, MPH

**Background:** Detection of the earliest cognitive changes signifying Alzheimer disease is difficult.

**Objective:** To model the cognitive decline in preclinical Alzheimer disease.

**Design:** Longitudinal archival study comparing individuals who became demented during follow-up and people who remained nondemented on each of 4 cognitive factors: global, verbal memory, visuospatial, and working memory.

**Setting:** Alzheimer Disease Research Center, Washington University School of Medicine, St Louis, Missouri.

**Participants:** One hundred thirty-four individuals who became demented during follow-up and 310 who remained nondemented.

**Main Outcome Measures:** Inflection point in longitudinal cognitive performance.

**Results:** The best-fitting model for each of the 4 factors in the stable group was linear, with a very slight downward trend on all but the Visuospatial factor. In contrast, a piecewise model with accelerated slope after a sharp inflection point provided the best fit for the group that progressed. The optimal inflection point for all 4 factors was prior to diagnosis of dementia: Global, 2 years; Verbal and Working Memory, 1 year; and Visuospatial, 3 years. These results were also obtained when data were limited to the subset (n=44) with autopsy-confirmed Alzheimer disease.

**Conclusions:** There is a sharp inflection point followed by accelerating decline in multiple domains of cognition, not just memory, in the preclinical period in Alzheimer disease when there is insufficient cognitive decline to warrant clinical diagnosis using conventional criteria. Early change was seen in tests of visuospatial ability, most of which were speeded. Research into early detection of cognitive disorders using only episodic memory tasks may not be sensitive to all of the early manifestations of disease.

Arch Neurol. 2009;66(10):1254-1259

Recent studies have focused on identifying the beginning of the transition from healthy aging to dementia. As new interventions become available, it will become important to identify the disease as early as possible. A piecewise regression analysis of a measure of episodic memory identified an inflection point 5 years before diagnosis in the Bronx Aging Study. A flat trajectory followed by decline beginning 7 years before diagnosis of dementia was reported for the same measure in the Baltimore Longitudinal Study of Aging, which also found decline in executive function that increased in rate 2 to 3 years before diagnosis. Episodic memory is not the only aspect of cognition that can be affected in preclinical Alzheimer disease (AD). Indeed, mild cognitive impairment, often thought to represent a transitional state between healthy cognitive aging and AD, is defined on the basis of deficits in cognitive domains in addition to memory.

In this article we examine cognitive domains beyond episodic memory and executive function to test hypotheses about the existence of inflection points before clinical diagnosis of dementia. Based on the identification of common factor structures in cognitively healthy individuals and those with AD, we examined global mental ability and 3 specific cognitive domains (verbal memory, working memory, and visuospatial ability) through a long preclinical period as people developed dementia and compared them with those who remained cognitively healthy. Results from longitudinal studies support an early observation that the overall course of healthy aging is relatively stable compared with the cognitive decline, often precipitous, experienced by those who develop dementia. We sought to model
that change to determine when the inflection occurs before dementia detection and the rate of change afterward for different cognitive domains and to validate these clinical observations in an autopsy-confirmed sample.

METHODS

PARTICIPANTS

Longitudinal archival data were examined from 444 volunteers initially aged 60 to 101 years enrolled in the Alzheimer Disease Research Center, Washington University School of Medicine, St Louis, Missouri, between October 1, 1979, and December 31, 2006 (Table 1). All the participants were clinically evaluated to be cognitively healthy (Clinical Dementia Rating 11 [CDR]=0) at the time of their first psychometric assessment and had at least 1 annual additional clinical evaluation through November 30, 2006 (Table 1). The Washington University Human Studies Committee approved all the procedures. Data from these participants have been used in other publications.

CLINICAL EVALUATION

Research-trained clinicians and nurses determined whether the participant was demented (CDR>0) or not demented (CDR=0) based on semistructured interviews with the participant and a knowledgeable collateral source (usually the spouse or an adult child), a health history, medication and depression inventories, an aphasia battery, and a neurologic examination of the participant. The diagnosis of dementia was based on a history of gradual onset and progressive cognitive decline that interfered with the person’s ability to perform accustomed activities. The CDR has high interrater reliability, is sensitive to clinical progression, and is highly predictive (93%) of autopsy-confirmed AD. Participants were seen by different physicians initially older, on average, than the stable group. are included solely for descriptive purposes; recall that they were included Block Design, Digit Symbol, Trailmaking A, and Benton, copy. A prorated factor score was computed if 1 or 2 values were missing for the Global factor and if 1 value was missing for the specific factors; otherwise, the factor value for that assessment for that person was excluded from the analyses.

NEUROPATHOLOGICAL ASSESSMENT

All the brains were examined according to a standard protocol. After fixation in neutral-buffered 10% formalin, tissue blocks were obtained from 30 brain regions. Sections (6 µm) from paraffin-embedded tissue blocks were stained with hematoxylin-eosin, Gallyas and modified Bielschowsky silver stains, and immunohistochemical methods. Histologic criteria for AD were based on the quantification of diffuse and neurtic amyloid deposition in 5 cortical regions with 10-mm2 mi-
inflection point (\(Y^{\hat{T}}_{1}\)) and the next assessment thereafter (\(Y^{\hat{T}}_{2}\)).

To determine the best form of a factor score's trajectory through time for each group (stable and progressed), we used \(\chi^2\) tests for \(-2\) log likelihood ratios (\(-2\)LLs) for nested models of increasing complexity of the slope across time (linear, quadratic, linear piecewise, linear piecewise optimized, and quadratic piecewise); simpler models are nested in more complex models. Model comparisons used \(\chi^2\) tests of deviance scores beginning with the simple linear model. Deviance scores equal the difference between the \(-2\)LL of a simpler model and a more complex one (\(\Delta\chi^2\)).

The linear piecewise change model specified a point of inflection and tested whether rates of cognitive decline differed before and after that point.\(^{22}\) To determine the optimum placement of the change point in the piecewise model, we tested inflection points from 1 to 6 years before the last assessment for the stable group. For those who progressed, we tested for an inflection point at the time of diagnosis of dementia and from 4 years before diagnosis to 2 years after diagnosis (>50 observations at each selected time point). The quadratic piecewise model added a quadratic term for the postinflection time variable using the optimal inflection point.

We conducted additional analyses to determine whether instead of a quadratic function after inflection there was acceleration in the rate of progression (ie, a change in the slope of the slopes across time). Technically, this was not a test of a nested model. First, we used the optimal piecewise model to predict the factor scores at each time of assessment for each person (best linear unbiased predictor). Using these latent values from each time of assessment rather than observed values, we calculated latent difference scores (LDSs) for each person. The LDS equals the difference between the predicted values at 2 adjacent times of assessment (\(Y_{T_{1}}, \ldots, Y_{T_{2}}\), \(Y_{T_{2}}, \ldots, Y_{T_{3}}\), \(Y_{T_{3}}, \ldots, Y_{T_{4}}\), and so forth) beginning with the difference of the predicted value at the optimized inflection point (\(Y_{T_{1}}\)) and the next assessment thereafter (\(Y_{T_{2}}\)).

Then we tested a linear mixed model using the LDS as the dependent variable to determine whether the slope of the LDS values (ie, the slope of the slopes) changed across time. The acceleration coefficient tested herein is a 2-stage regression analogue to other acceleration models derived from structural equation modeling\(^{20}\) and the functional equivalent of acceleration in kinematics (ie, the second derivative of position/intercept). Although this method of measurement may slightly attenuate true acceleration values because of the tendency for latent values to “shrink” in the presence of missing data,\(^{24}\) its computation is robust, is straightforward, and can be estimated without multiple imputation of missing data. It yields results consistent with simultaneous models.\(^{25}\) A simultaneous model that approximates slopes and acceleration was not attempted because we did not know the functional form of the data (the primary aim of this investigation).

After determining the optimal models within each group, slope estimates were based on a mixed model for each factor that included covariates, effects for group (stable vs progressed), time (before and after inflection), and group \(\times\) time interactions. Acceleration coefficients for the group that progressed were estimated from the second stage of the LDS model and were added to slope coefficients after the point of inflection. All analyses were repeated for the subset of the group that progressed that had autopsy confirmation of AD.

**RESULTS**

**SAMPLE CHARACTERISTICS**

All the participants were not demented (CDR=0) at entry and either remained CDR=0 (stable) throughout follow-up (n=310, 37% men) or progressed to CDR>0 (n=134, 34% men) with a clinical diagnosis of uncertain dementia (CDR=0.5) or dementia of the Alzheimer type (CDR≥0.5) by the time of their last evaluation (Table 1). Participants who progressed and whose clinical diagnosis was non-AD dementia (eg, vascular dementia associated with Parkinson disease) were excluded. Individuals who came to autopsy with a clinical diagnosis of dementia of the Alzheimer type but who had another dementia abnormality were also excluded (n=14). Maximum follow-up was 25.7 years (mean [SD], 5.9 [3.3] years). The stable group was slightly more educated (mean, 14.8 years) than the group that progressed (mean, 14.1 years). As might be expected given that AD is age associated, those who progressed were older at entry (mean [SD], 80.4 [8.9] years) than those whose performance remained stable (mean [SD], 74.4 [8.6] years). Apolipoprotein E4 status did not differ between the stable (28% carriers) and progressed (27% carriers) groups. Autopsy confirmation of a diagnosis of AD was available for a subset (n=44, 36% men) of the group that progressed. At the time of progression, they were older (mean age, 90.2 years) than the rest of the group (mean age, 84.1 years), although their mean educational level (13.8 years) was comparable with that of those without autopsy (14.1 years).

**NONDEMENTED AGING**

The linear regression model provided the best fit for each factor in the stable group. The \(-2\)LL values were as follows: Global, 4255.5; Verbal Memory, 3045.3; Visuospatial, 2796.7; and Working Memory, 5912.5. More complex models did not improve fit (\(\Delta\chi^2>14.4\) for all, \(P>0.05\)).

**CHARACTERIZING PRECLINICAL AD:**

**INFLECTION POINTS**

In the progressed group, the linear model provided adequate fit for all 4 factors (\(-2\)LL values) (Table 3). Fit was not improved using a quadratic model for any factor (\(P>0.05\) for all), although it was improved using a piecewise model with an inflection point at the time of diagnosis. The piecewise model fit was improved by moving the inflection point before diagnosis (\(P<0.001\) (Table 3). The optimal inflection point varied for the 4 factors: 2 years before clinical diagnosis for the Global.
factor, 1 year before diagnosis for the Verbal and Working Memory factors, and 3 years before diagnosis for the Visuospatial factor (Figure).

The Verbal Memory factor included episodic and semantic memory measures. An inflection point for a single measure of episodic memory occurred substantially longer than 1 year before diagnosis in previous studies. Therefore, we also examined each of the 2 measures of episodic memory independently. The inflection point occurred 4 years before diagnosis for WMS Associate Learning and 2 years before diagnosis for WMS Logical Memory.

The fit of the piecewise model did not improve with the addition of a quadratic term after the optimal inflection point (P > .05 for all) (Table 3). There was, however, a significant increase in the rate of decline after the inflection point for all 4 factors using the LDS models of acceleration (t > 12.10 for all, P < .001). Thus, the optimal model in the group that progressed was piecewise, with linear slope before inflection and accelerated slope after inflection.

### Table 3. −2LL Values for Each Cognitive Factor for Each Model and Significant Maximum Likelihood Deviance Tests (Δχ²) for the Group That Progressed

<table>
<thead>
<tr>
<th>Model</th>
<th>Global Factor</th>
<th>Verbal Memory Factor</th>
<th>Visuospatial Factor</th>
<th>Working Memory Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>−2LL</td>
<td>Δχ²</td>
<td>−2LL</td>
<td>Δχ²</td>
</tr>
<tr>
<td>Simple linear</td>
<td>2750.8</td>
<td>NA</td>
<td>2777.7</td>
<td>NA</td>
</tr>
<tr>
<td>Quadratic</td>
<td>3019.5</td>
<td>268.7</td>
<td>2277.7</td>
<td>0</td>
</tr>
<tr>
<td>Piecewise linear</td>
<td>2705.1</td>
<td>−45.7</td>
<td>2255.9</td>
<td>−21.8</td>
</tr>
<tr>
<td>At diagnosis</td>
<td>2690.3</td>
<td>−14.8</td>
<td>2243.5</td>
<td>−12.4</td>
</tr>
<tr>
<td>Optimized</td>
<td>2689.5</td>
<td>−0.8</td>
<td>2243.9</td>
<td>0.4</td>
</tr>
<tr>
<td>Piecewise quadratic</td>
<td>2689.5</td>
<td>−0.8</td>
<td>2243.9</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Abbreviations: −2LL, −2 log likelihood ratio (smaller values indicate better fit); NA, not applicable.

1 df for deviance tests.
2 The optimal inflection point was 2 years before diagnosis for the Global factor, 3 years before diagnosis for the Visuospatial factor, and 1 year before diagnosis for the Verbal and Working Memory factors.
3 Quadratic term after the inflection point.

Figure. Longitudinal course of the stable, progressed, and autopsy-confirmed Alzheimer disease (AD) groups before and after diagnosis of AD (DX) on Global factor (A), the Verbal Memory factor (B), the Visuospatial factor (C), and the Working Memory factor (D). All available data were used for the analysis, but the plotted values for the stable and progressed groups include at least 50 observations per time point per group.
CHARACTERIZING PRECLINICAL AD: SLOPE AND ACCELERATION ESTIMATES

Estimates and their standard errors are given in Table 4. Change in performance in all 4 factors is demonstrated in the Figure. The stable and progressed groups shared similar preinflection trajectories; the group × time (before inflection) interactions were not significant for any of the 4 factors (P > .19 for all). There was a significant downward linear trend in global cognitive abilities and in verbal and working memory (P < .01); however, no longitudinal decline was detected in the Visuospatial factor (P = .29) (Figure). Results were similar when data from the group that progressed were limited to individuals with autopsy-confirmed AD.

A different pattern of longitudinal cognitive performance was seen in the group that progressed to dementia. Compared with preinflection slopes, they had steeper downward slopes after inflection (P < .001), and the rate of decline accelerated with time (Table 4 and Figure). The greatest preclinical slope change was in the Working Memory factor (slope = −0.66, acceleration = −0.17) beginning 1 year before dementia diagnosis followed by the Global factor (slope = −0.32, acceleration = −0.06). Postinflection slopes and accelerations were similar for the Verbal Memory and Visuospatial factors. Slopes were steeper when the sample that progressed was restricted to those with autopsy-confirmed AD, but rates of acceleration were comparable with those for the total group that progressed.

We demonstrate models of preclinical decline in a well-characterized longitudinal sample with inflection points in cognitive performance occurring several years before clinical diagnosis of dementia. It is apparent from these models that there is a clear turning point in the transition from normal aging to preclinical AD. A novel finding was that visuospatial abilities demonstrated an inflection point 3 years before clinical diagnosis. This decline on tests that were primarily speeded represented a sharp departure from the previous longitudinal pattern of these initially nondemented individuals, which was similar to that of those who did not become demented. Global cognitive abilities followed decline in visuospatial ability during the next year. Inflection points in the Verbal and Working Memory factors were not seen until 1 year before clinical diagnosis. The delayed inflection point for Verbal Memory probably results from the combination of episodic and semantic memory measures on one factor. If sufficient measures of each type of memory were available to form separate factors, an earlier inflection point for episodic memory would probably emerge based on the results obtained for the 2 individual measures of episodic memory.

The rate of decline accelerated after the downward course began. This was true for all 4 factors but was most apparent for Working Memory. Of course, the estimated rates of decline and acceleration depend on the tests administered, their level of difficulty, and floor and ceiling effects. For example, 3 tests in the battery (the Boston Naming Test, the copy version of the Benton Visual Retention Test, and the WMS Mental Control) have ceiling effects in the preinflection period in nondemented older adults. We were limited to these archival data from a battery that was originally constructed in 1979 for a study of mild dementia, and it does not contain more modern measurements of working memory.

The number of years before diagnosis of dementia that the inflection point occurs in the longitudinal course depends on the method of diagnosis and on the characteristics of the cognitive tests. The period will be longer if one relies on test norms, particularly if the sample is at a high level of function initially, than if one relies on collateral source reports of change from previous levels of function, captured by the CDR. Furthermore, inclusion of people in the preclinical stage in nondemented samples overestimates decline in cognitive ability traditionally attributed solely to age. This makes it more difficult to detect beginning dementia using conventional norms based on these contaminated samples.

A great strength of this study is the replication of the pattern of the longitudinal results observed in the larger sample that progressed in the subset with autopsy-confirmed AD. Although the rates of decline were somewhat steeper in the autopsied subset, the rates of acceleration were the same. The rate of progression in AD is

---

Table 4. Estimated Slopes and Rates of Acceleration for the 4 Cognitive Factors

<table>
<thead>
<tr>
<th>Factor and Group</th>
<th>Before Inflection</th>
<th>After Inflection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>Slope</td>
</tr>
<tr>
<td>Global factor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable</td>
<td>−0.05 (0.01)</td>
<td>NA</td>
</tr>
<tr>
<td>Progressed</td>
<td>−0.07 (0.02)</td>
<td>−0.32 (0.02)</td>
</tr>
<tr>
<td>AD confirmed</td>
<td>−0.10 (0.03)</td>
<td>−0.40 (0.04)</td>
</tr>
<tr>
<td>Verbal Memory factor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable</td>
<td>−0.02 (0.01)</td>
<td>NA</td>
</tr>
<tr>
<td>Progressed</td>
<td>−0.02 (0.01)</td>
<td>−0.15 (0.02)</td>
</tr>
<tr>
<td>AD confirmed</td>
<td>−0.03 (0.02)</td>
<td>−0.19 (0.03)</td>
</tr>
<tr>
<td>Visuospatial factor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable</td>
<td>0.00 (0.01)</td>
<td>NA</td>
</tr>
<tr>
<td>Progressed</td>
<td>−0.02 (0.01)</td>
<td>−0.13 (0.01)</td>
</tr>
<tr>
<td>AD confirmed</td>
<td>−0.02 (0.02)</td>
<td>−0.17 (0.02)</td>
</tr>
<tr>
<td>Working Memory factor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable</td>
<td>−0.08 (0.02)</td>
<td>NA</td>
</tr>
<tr>
<td>Progressed</td>
<td>−0.10 (0.03)</td>
<td>−0.66 (0.04)</td>
</tr>
<tr>
<td>AD confirmed</td>
<td>−0.12 (0.04)</td>
<td>−0.89 (0.06)</td>
</tr>
</tbody>
</table>

**Abbreviations:** AD, Alzheimer disease; NA, not applicable.

* a Data are given as estimate (SE). Slope coefficients were estimated using the piecewise change model of raw data. Acceleration coefficients were estimated using latent difference score regression, in which intercepts were equivalent to piecewise change slope estimates (±0.02 U).

* b The optimal inflection point was 2 years before diagnosis for the Global factor, 1 year before diagnosis for the Verbal and Working Memory factors, and 3 years before diagnosis for the Visuospatial factor.
highly variable\textsuperscript{27}, perhaps those with autopsy-confirmed AD were individuals who progressed more rapidly. Another possibility is that the progressed group may contain individuals who do not have AD and, therefore, do not follow the same pattern of decline.

There are several implications of this study. Some of the earliest signs of preclinical disease may occur on tests of visuospatial and speeded psychomotor skills. Furthermore, the greatest rate of preclinical decline may occur on executive and attention tasks. These findings suggest that research into early detection of cognitive disorders using only episodic memory tasks, such as word lists or paragraph recall, may not be sensitive to either all of the earliest manifestations of disease or the most rapidly changing domain. Furthermore, the preclinical downward course comes after an inflection point. Before that point, the longitudinal course of those who did and did not develop AD was the same. In summary, converging longitudinal evidence suggests that after a sharp departure from the relatively flat course of normal aging there is a preclinical period in AD with insufficient cognitive decline to warrant clinical diagnosis using conventional criteria but that can be seen with longitudinal data from multiple domains of cognition and not just memory.

Accepted for Publication: February 25, 2009.

Correspondence: James E. Galvin, MD, MPH, Alzheimer Disease Research Center, Washington University School of Medicine, 4488 Forest Park, Ste 130, St Louis, MO 63108 (galvinj@neuro.wustl.edu).

Author Contributions: All authors had full access to all the data used in this study. Drs Johnson and Galvin take full responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Johnson, Storandt, and Galvin. Acquisition of data: Storandt, Morris, and Galvin. Analysis and interpretation of data: Johnson, Storandt, and Galvin. Drafting of the manuscript: Johnson, Storandt, and Galvin. Critical revision of the manuscript for important intellectual content: Johnson, Storandt, Morris, and Galvin. Statistical analysis: Johnson and Storandt. Obtained funding: Storandt and Morris. Administrative, technical, and material support: Galvin. Study supervision: Morris and Galvin.

Financial Disclosure: None reported.

Funding/Support: This study was supported by grants P01 AG03991, P50 AG05681, P01 AG026276 (Dr Morris), and K08 AG20764 (Dr Galvin) from the National Institute on Aging, National Institutes of Health.

Additional Contributions: We thank the Clinical and Neuropathology Cores of the Washington University Alzheimer Disease Research Center for the clinical, cognitive, and postmortem assessments and the Genetics Core for the apolipoprotein E genotype data.

REFERENCES


18. Armitage SG. An analysis of certain psychological tests used in the evaluation of brain injury. Psychol Monogr. 1946;60(1-48).


