

The existence of cognitive plateaus in Alzheimer's disease

Andrea C. Bozoki^{a,*}, Hyonggin An^b, Eva S. Bozoki^c, Roderick J. Little^d

^aDepartment of Neurology, Michigan State University, East Lansing, MI, USA

^bDepartment of Biostatistics, College of Medicine, Korea University, Seoul, Korea

^cBrookhaven National Laboratories, Upton, NY, USA (retired)

^dDepartment of Biostatistics, University of Michigan, Ann Arbor, MI, USA

Abstract

Background: The objective of this study was to evaluate the existence of cognitive plateaus in some individuals during the course of Alzheimer's disease (AD).

Methods: Data came from the historical patient group collected via the Consortium to Establish a Registry for Alzheimer's Disease (CERAD, Duke University, 1988–1996). Data reduction was performed by using principal components analysis to derive a single cognitive measure (F1), followed by application of a novel plateau-searching algorithm to individual patient data, looking for stable periods of 3 years or longer. To evaluate the time dependence of F1, we fitted a linear mixed model to the group and to individual data points.

Results: Twenty-two percent of AD subjects (54/243) and 98% of healthy control subjects (253/258) exhibited a plateau. Within the AD plateau group, the most common pattern was a single plateau (mean, 3.6 years; range, 3 to 7 years) that extended for the entire measurement period (28/54 subjects). Briefer plateau durations were seen at the beginning or end of the measurement period. Initial cognitive function (F1) was slightly higher in the plateau group, which was also slightly older and less well-educated. Men and women were equally represented.

Conclusions: In a patient sample predating the widespread use of cholinesterase inhibitors, we found that approximately one fifth of individuals with AD demonstrated periods of prolonged cognitive stability. This significant interindividual variability must be considered when providing prognostic information to families and when assessing individual patient responses to pharmacotherapy. We advise caution when assessing results of potentially disease-modifying agents at the individual patient level.
© 2009 The Alzheimer's Association. All rights reserved.

Keywords:

Alzheimer's disease; Cognition; Dementia; Neuropsychology

1. Introduction

Among dementia researchers, a question of interest has been whether some Alzheimer's disease (AD) patients experience "cognitive plateaus" in the course of the disease, during which they remain cognitively stable [1–3]. This issue has several ramifications from the practical need for counseling in such cases (does the patient still have AD? If so, what does the plateau imply for eventual disease progression?) to the idea that some individuals might have innate resistance to the ravages of the disease. We set out to create a procedure by which

cognitive plateaus could be identified as a first step toward clinical and biologic characterization of such individuals.

Clinically, we believed that a cognitive plateau "could be recognized if we saw it." The difficulty was, first, how to instantiate this intuitive assessment of cognitive stability into a valid, quantifiable, and reliable measurement tool and, second, how to determine whether such plateaus had clinical meaning or resulted from normal intersubject variability.

Several groups have studied the pattern of cognitive decline in AD patients by repeated measurements on a neuropsychological test such as the Folstein Mini-Mental State Examination (MMSE) [4], and results have been analyzed by using regression models with fixed effects [5–7] or linear mixed models [8–10]. However, because these approaches assume that all AD patients have a linear pattern of cognitive

*Corresponding author. Tel.: 517-884-2482; Fax: 517-432-9414.

E-mail address: Andrea.Bozoki@ht.msu.edu

decline, they are not suitable for the detection of plateaus among AD patients, and a single test score does not provide a robust measure of global cognition across the range of decline from mild to severe.

Other clinical studies have classified patients into those with and without plateaus on the basis of subjective visual inspection of each patient's trajectory of repeated measures. For example, the study by Piccini et al [3] subdivided AD patients into plateau and non-plateau groups on the basis of the clinician's direct knowledge and case history data. However, visual inspection lacks objectivity and repeatability, and it is possible that some apparent plateaus are not real, in that they would not exist in a set of repeated measures not subject to measurement error. Thus, a more objective statistical algorithm for quantifying and operationalizing the notion of cognitive plateaus is desirable.

To this end, Haxby et al [1] applied a bilinear model, which fit two lines to each patient trajectory, a plateau and a decline. However, they made the restrictive assumption that the pattern of decline could be described by two straight lines, a plateau with zero slope followed by a linear cognitive decline, and they used visual inspection to determine a change point for each patient. Notably, their technique requires a minimum of three points in each of the two sections, necessitating at least 5 years of data assuming annual follow-up, and their approach excludes plateaus that occur later than the beginning of the observation period. In 1993, Brooks et al [11] proposed a trilinear model that amended the bilinear one by adding a third, terminal component of little to no change, representing a "floor" state. This model, however, had similar limitations to the prior one.

Several more recent approaches to modeling cognitive change over time have been proposed, although these models were not developed to look for plateaus. In 2002, Helmes et al [12] proposed a quadratic model of decline for a single-variable measure of cognition, finding that this model was a better fit to their data (12 autopsy-verified AD patients) than the trilinear model of Brooks et al [11], particularly at change points. As an alternative, an individual growth curve method was used in a recent article that analyzed differences in cognitive decline on the basis of *APOE* status [13]. Unfortunately, neither quadratic models nor growth curves can model an extended period of stability in the process being measured and mathematically cannot distinguish between those who are a "poor fit" because they remain stable over time and those who simply do not conform to the modeling method generally.

A final approach worth mentioning, by Smith et al [14], evaluated the existence of a plateau during the pre-AD phase of memory decline. Rather than examining individual cases, they demonstrated that a bilogistic model that included a term for a plateau was a better fit to their neuropsychometric memory score data than a simple (linear) logistic model. Because they used mean scores for their entire study population, they were able to conclude that, in the aggregate, individuals with pre-AD show a period of leveling off in memory decline.

However, their approach could not address the related question of whether all or only some individuals show this plateau, and whether there are interindividual differences in the duration and outcome of plateau periods.

In response to these limitations, we propose a method that looks for plateaus in individual patients, regardless of their position within the observation period, and that does not presuppose any specific shape (model) for the periods of time before or after the plateau phase. To begin, we identified a number of psychometric variables that collectively represented a fair composite measure of cognitive functioning and used a standard method of data reduction to come up with a single linear combination of the variables. We used a linear mixed model to characterize the behavior of this composite variable, noting differences in both slope and intercept of the AD and normal control (NC) groups, which allowed us to verify the expected differences between groups for our composite variable. The next step, our plateau-searching algorithm, is the main focus of the article. The algorithm was applied directly to our composite variable, not to a model-fitted (ie, linear, multilinear, bilogistic) function. Additional details of this process are available in the Appendix. Finally, we validated our algorithm by using the same linear mixed model approach to characterize the different behavior of the plateau and non-plateau groups, demonstrating that the former has a slightly higher intercept and a markedly different slope.

2. Methods

2.1. Description of the data set

We used the Consortium to Establish a Registry of Alzheimer's Disease (CERAD) data set [15,16] to test our hypothesis that cognitive plateaus lasting at least 3 years exist during the course of this disease in some individuals. Importantly, this group was recruited between 1987 and 1995, with enrollment completed before the introduction of donepezil hydrochloride in 1996. Patients and control subjects were evaluated at entry and annually thereafter with standardized instruments for the neuropsychological assessment of dementia [17]. A total of 1,094 subjects with AD and 463 controls (NC) were initially enrolled; however, only 402 AD and 301 NC subjects had at least three complete sets of observations (neuropsychological testing). NC subjects as a group were matched for age, sex, and educational attainment. All enrollment, inclusion, and exclusion criteria for the group are set out in the initial publication of the CERAD investigator group [18].

For our study, we also excluded subjects from the CERAD data set with an observation period of <3 years because this duration of follow-up was judged inadequate to assess for the presence of a plateau. After both exclusions, our sample contained 243 subjects with AD and 258 controls. Demographic information regarding the total CERAD group, the subgroup with at least three complete sets of data, and our final sample are given in Table 1. Our sample does not differ substantively

Table 1

Selected demographic variables and four neuropsychological variables that comprise F1, with associated mean and SD values for both the AD and NC subjects at study entry

	AD						NC					
	Total CERAD group		≥ 3 data sets		≥ 3 data sets covering ≥ 3 years		Total CERAD group		≥ 3 data sets		≥ 3 data sets covering ≥ 3 years	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
% Male	41		42		44*		35		34		34*	
Age at enrollment (y)	73.0	8.0	71.7	7.9	71.7*	7.7*	68.9	8.0	68.7	7.7	68.5*	7.5*
Years of education	12.1	3.8	12.7	3.6	12.5*	3.7*	13.7	3.3	14.2	3.0	14.4*	3.0*
Verbal Fluency	7.3	4.2	8.5	4.2	8.7*	4.3*	17.6	4.9	18.3	4.9	18.6*	4.9*
Boston Naming	10.2	3.6	11.4	3.2	11.5*	3.2*	14.4	1.0	14.7	0.9	14.6*	0.9*
MMSE	17.0	5.9	19.0	4.9	19.2*	4.9*	28.7	1.7	28.9	1.4	29.1*	1.4*
Delayed Recognition	13.7	3.7	14.3	3.3	14.4*	3.3*	19.6	0.9	19.6	0.8	19.6*	0.9*
N	1094		402		243*		463		301		258*	

*Study group. Total CERAD data set and group with adequate neuropsychological data but inadequate duration are included for comparison.

from previously published longitudinal studies of this group (eg, Morris et al [19]), which note that a significant percentage of enrollees failed to return for follow-up. Average duration of follow-up in our analyzed sample was 4.3 years, with 4.6 neuropsychological evaluations per subject.

2.2. Data reduction by using principal components analysis

Of the CERAD variables available, we chose those reflecting the widest variety of cognitive tasks, although we were constrained by several practical issues: (1) the more tests included, the more extensive the missing data problem; (2) some variables had more missing data than others; and (3) certain variables included in the data set were not independent.

Our final selection contained four variables: (1) total correct on the 15-item Boston Naming Test (BNTot); (2) score on the MMSE; (3) total correct items on the category-naming test (VFTot); and (4) total correct on the delayed verbal recognition test (Recog). Recognition rather than free recall was selected as a result of a significant floor effect of the latter in the AD subjects. Table 1 provides mean and standard deviation (SD) of each variable for both AD and NC groups. For comparison, the means and SDs of the entire CERAD data set and the group that had at least three complete sets of data are also given.

Principal components analysis (PCA), a common method of data reduction, was used to consolidate the four variables into a single aggregate variable. PCA was performed for all subjects (AD and NC) and all time points together, ensuring that the scales of the calculated principal factors were the same for both groups and allowing us to compare their behavior directly.

We found that 87% of the total variance could be accounted for by the first principal factor, F1, calculated as the following:

$$F1 = .04514 \text{ VFTot} + 0.02131 \text{ BNTot} + 0.05029 \text{ MMSE} + 0.02780 \text{ Recog} - 2.35 \quad (\text{equation 1}).$$

The minimum, mean, and maximum values of F1 were calculated as -2.35 , 0 , and 2.1 , respectively. Furthermore, the distribution of F1 was close to a normal distribution for both NC and AD subjects. The means and SDs of the two groups were NC, $.79 \pm .24$ and AD, $-.79 \pm .82$. Overall mean and SD is 0 ± 1 , as expected from the calculation method.

2.3. Linear mixed model fit of the F1 aggregate score

We fitted a linear mixed model to the repeated F1 values. This model assumes that the F1 values for subject- i are generated as the following:

$$F1_i(t) = G(t) + S_i(t) + \epsilon_{it} \quad (\text{equation 2}),$$

where $G = b_0 + b_1 t$ is the subject-independent “fixed” part, characterizing average behavior of the whole group; $S_i = \beta_{0i} + \beta_{1i} t$, is the subject-dependent “random” part, characterizing individual subject variation; and ϵ_{it} is the measurement error of F1 observed at time t for the i^{th} subject, assumed to be normally distributed with mean 0 and variance σ^2 . Thus, time dependence of individual $F1_i$ points are described by a linear equation, with $(b_0 + \beta_{0i})$ as intercept and $(b_1 + \beta_{1i})$ as slope.

Distributions of the β_{1i} individual slopes for both NC and AD groups were expected to be normal, with means equal to zero. However, if a group of subjects within the total AD group had plateaus, we predicted that their corresponding β_{1i} values would have a distribution around a positive mean (because the fixed portion of the linear model has a negative slope, but subjects with plateaus would be described by a nearly flat line), whereas the non-plateau subgroup would have a distribution of β_{1i} values with a negative mean to keep the total group mean at zero. The $b_1 + \text{mean}(\beta_{1i})$ total individual slope for the plateau subgroup was not expected to be zero as a result of observation points outside the plateau region for individual subjects.

2.4. The plateau-searching algorithm

Independently of the linear model fitting, we examined the set of observed F1 values for each individual subject to

determine whether some exhibited plateau characteristics. Our algorithm operationalized a clinical definition of plateau on the basis of seeking periods of at least 3 years where F1 leveled off. A good definition of a plateau in this sense would be “a stable period, buttressed by steeper rates of decline.” However, (1) the initial decline might have been before the start of the study; (2) the declining period after a stable period might be missing if the subject dropped out from the study; and (3) the average follow-up duration of our sample was 4.3 years. Because we required a plateau to last at least 3 years, we could not be assured of finding the flanking declining portions of the plateau. Consequently, we defined plateau as “a stable period of 3 years or longer,” without the requirement for either initial or terminal decline.

We chose 3 years as a minimum duration of the stable period to be accepted as a plateau to reduce the likelihood of a type I error. We required the length of a cognitive plateau to be greater than the minimum period of time necessary for a reliable and replicable drop in each of the neuropsychological test scores to become manifest, ie, when test-retest variability in a given subject was no longer expected to be smaller than the SD for that test. For example, the SD of the initial CERAD group of AD subjects on the MMSE was 4.5, whereas the annualized rate of change (on the basis of 4 years of data) was 3.9 [16]. It would therefore take 2 years before the drop in test score would exceed the SD. This was also true for the Boston Naming Test and Verbal Recognition subtests. For Verbal Fluency, somewhat more than 2 years was necessary (annualized drop was 1.9, but SD was 4.1). Because we used a principal component that combines all four of these variables, we had to choose a duration that accommodated even Verbal Fluency, thus the 3-year requirement.

The algorithm searches for stable periods in which the underlying cognitive score is unchanged within measurement error. A patient is considered to have a plateau if there is a stable period of 3 years or longer that lies above a minimum floor value. The algorithm also searches for cases in which

there is fluctuation around a central value over time, because some subjects had considerable test-retest variation but no overall tendency toward decline. Otherwise, the patient was categorized into the non-plateau group (for details see Appendix).

The algorithm requires the determination of a “corridor” around the prospective plateau value, within which F1 has to be confined to be considered as exhibiting a plateau. We chose the $SD(F1)$ for the NC group, 0.24, as the lower bound for c and $3*SD(F1)$, or 0.72, as the upper bound of c . Therefore, successive F1 values that were between $F1-0.24$ and $F1+0.72$ were accepted as being within the corridor and therefore part of the plateau. We allowed more leeway for cognitive improvement than for decline, because improvement cannot be attributed to progressing dementia. It also requires the choice of a floor value. We defined this as -1.6 , the value of F1 when each of the four variables comprising it is 2 SDs below the mean of the AD group. All F1 values were between -2.35 and 2.1 .

3. Results

3.1. Linear mixed model fit of F1: AD versus NC groups

We performed a set of linear regression analyses in the form $F1_i = (b_0 + b_1t) + (\beta_{0i} + \beta_{1i}t) + \epsilon_{it}$ (see equation 2), thus obtaining the value of b_0 and b_1 parameters, common to all subjects, as well as the value of the β_{0i} , β_{1i} , and ϵ_{it} parameters for each of the $i = 1, N$ subjects individually. This procedure was applied for NC and AD subjects separately. The scatter diagrams of F1 versus time are shown in Figure 1 for all the AD and NC subjects.

Although healthy subjects did not change over time, AD subjects deteriorated significantly.

As expected, AD subjects started at a considerably lower F1 value (-0.31 versus $+0.80$). Also not surprisingly, the slope of AD subjects was $-0.23/\text{year}$, indicating a decline in cognition over time, as opposed to the ~ 0 slope of the NC

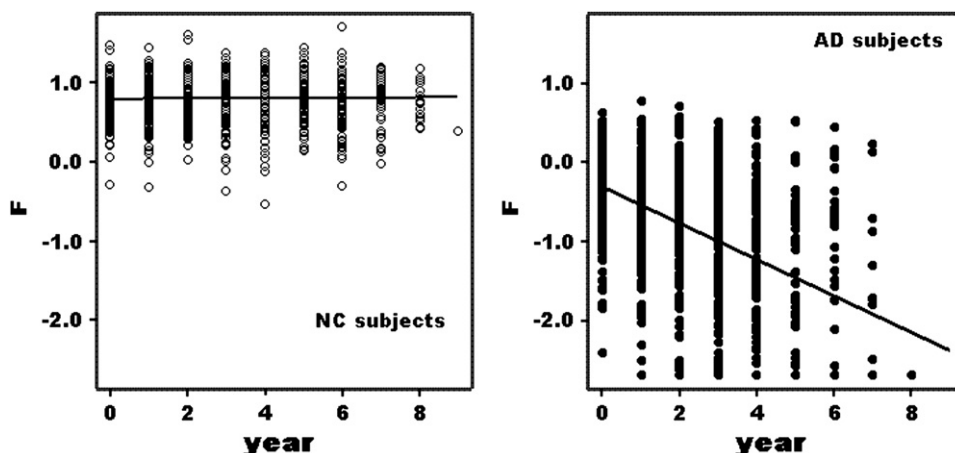


Fig. 1. F1 vs time for NC and AD subjects. Solid lines represent the corresponding $G = b_0 + b_1t$ regression lines. $G = (.80 \pm .013) + (.001 \pm .003)*\text{Year}$ for NC subjects, and $G = (-.31 \pm .03) + (-.23 \pm .01)*\text{Year}$ for AD subjects.

group. Spread of individual data was also much larger for AD than for NC subjects, resulting in larger SDs. The annualized rate of change of F1 for AD subjects was $-.23 \pm .01$; thus, the drop during a 3-year period was predicted to be -0.69 .

Distribution of individual “random” slopes (β_{1i}) for both NC and AD subjects was found to be normal (means, 0; SDs, .02 and .14, respectively). Spread for AD subjects was an order of magnitude larger than for NC subjects. Distribution of ϵ residuals was also normal (means, 0; SDs, .15 and .33, respectively).

3.2. Searching for plateaus in the CERAD data

Twenty-two percent of AD subjects (54/243) and 98% of NC subjects (253/258) exhibited plateaus (standard error, 3% for AD group and 1% for NC). In the plateau group, average plateau length was 3.6 years (range, 3 to 7 years). Figure 2 illustrates typical subject data. Among the plateau group, several different patterns could be seen. The most common, occurring in 28 of 54 subjects, was that the cognitive plateau began at the beginning of the measurement period, and F1 remained near its high end for the entire or almost entire measurement period (eg, subjects 230002, 230007). The remaining plateau subjects’ patterns were fairly equally distributed among initial decliners (8/54; 3 of these subjects also had a declining phase at the end, an orthodox plateau, eg, subject 150001) and terminal decliners (14/54; eg, 230046), with four subjects showing a pattern of two plateaus (each 36 months or longer) separated by a drop in the middle.

For the non-plateau group, cognitive ability declined strongly over time, in most cases dropping at least a full F1 point, although in a few cases a 1- or 2-year stable period could not be excluded.

3.3. Linear mixed model fit of F1: Plateau versus non-plateau subgroups

Initial F1 values were considerably less different between these groups than for the AD versus NC comparison, although still significant (-0.04 ± 0.57 for the plateau group and -0.29 ± 0.41 for the non-plateau; $t = 3.02$; $P = .003$) (Figure 3). However, decline for non-plateau subjects was five times more rapid than for plateau subjects ($-.30/\text{year}$ versus $-.06/\text{year}$). Overall decline in F1 of the plateau group did not contradict the existence of plateaus; it is caused by observation points outside the plateau region for individual subjects.

Distributions of the individual β_{1i} values for AD subjects with plateaus and for the rest of the AD subjects are shown in Figure 4. The two normal distributions are distinctly shifted from each other, with a positive mean and a much narrower distribution of values for the plateau subgroup.

3.4. Age, gender, and education differences

There was a small but significant difference in mean age at enrollment between groups (74 ± 6.6 years for

plateau cases and 71 ± 7.8 years for non-plateau cases; $t = 2.62$; $P = .009$).

The percentage of men did not differ between groups (44%, plateau versus 45%, non-plateau). There was no effect of gender on β_{1i} for plateau or non-plateau cases, whereas there was a marginal effect on β_{0i} ; men had slightly higher values in the non-plateau group ($t = 2.37$; $P = .02$) but not in the plateau group ($t = 1.89$; $P = .06$).

There was a small but significant difference in the average years of education between groups; the plateau group had a slightly lower attainment than the non-plateau group (11.5 ± 4.3 years versus 12.8 ± 3.4 years; $t = 2.33$; $P = .021$). Age at enrollment was independent of education for plateau cases ($r_2 = 0.002$; $F = 0.119$; not significant), whereas it decreased with increasing educational attainment in the non-plateau group ($r_2 = 0.049$; $F = 9.548$; $P = .002$).

3.5. Differences by disease severity

Because Clinical Dementia Rating (CDR) information is also available in the CERAD database, we were able to examine the effects of disease severity on plateaus. Overall, mean CDR \pm SD at first visit was 1.17 ± 0.46 in the plateau group and 1.38 ± 0.6 in the non-plateau group ($t = 2.39$; $P = .018$). The actual number of subjects and the percentage they represent by CDR score for the plateau and the non-plateau groups are shown in Table 2.

4. Discussion

Our findings demonstrate that 22%, or approximately one fifth of individuals with AD in this sample, exhibited a cognitive plateau as defined by our criteria. In about half, the plateau lasted as long as that patient was followed through the CERAD study. Individuals in the plateau group were slightly older and less well-educated, suggesting that the presence of a plateau is not merely an effect of younger age or higher education. Initial F1 values of the plateau group were slightly higher than those of the non-plateau group, suggesting that the former started with slightly milder disease; but this difference (0.25F1) seems too small to account for the plateauing.

The relationship between disease severity and plateaus was further examined by use of CDR scores. These results corroborated our findings with initial F1 values, because the average initial CDR score was slightly higher (worse) in the non-plateau group, although, again, the difference was slight. More useful information about the influence of severity on plateau occurrence was obtained by examining the percentage of plateaus by stage. We found that plateaus occur at all stages of disease (short of severe, which we could not examine because of lack of inclusion of that group at enrollment), although the percentage of cases with plateaus appears to be highest in the mild (CDR = 1) stage. The relatively small numbers of subjects in all but the mild stage make it difficult to draw strong conclusions regarding plateau distribution.

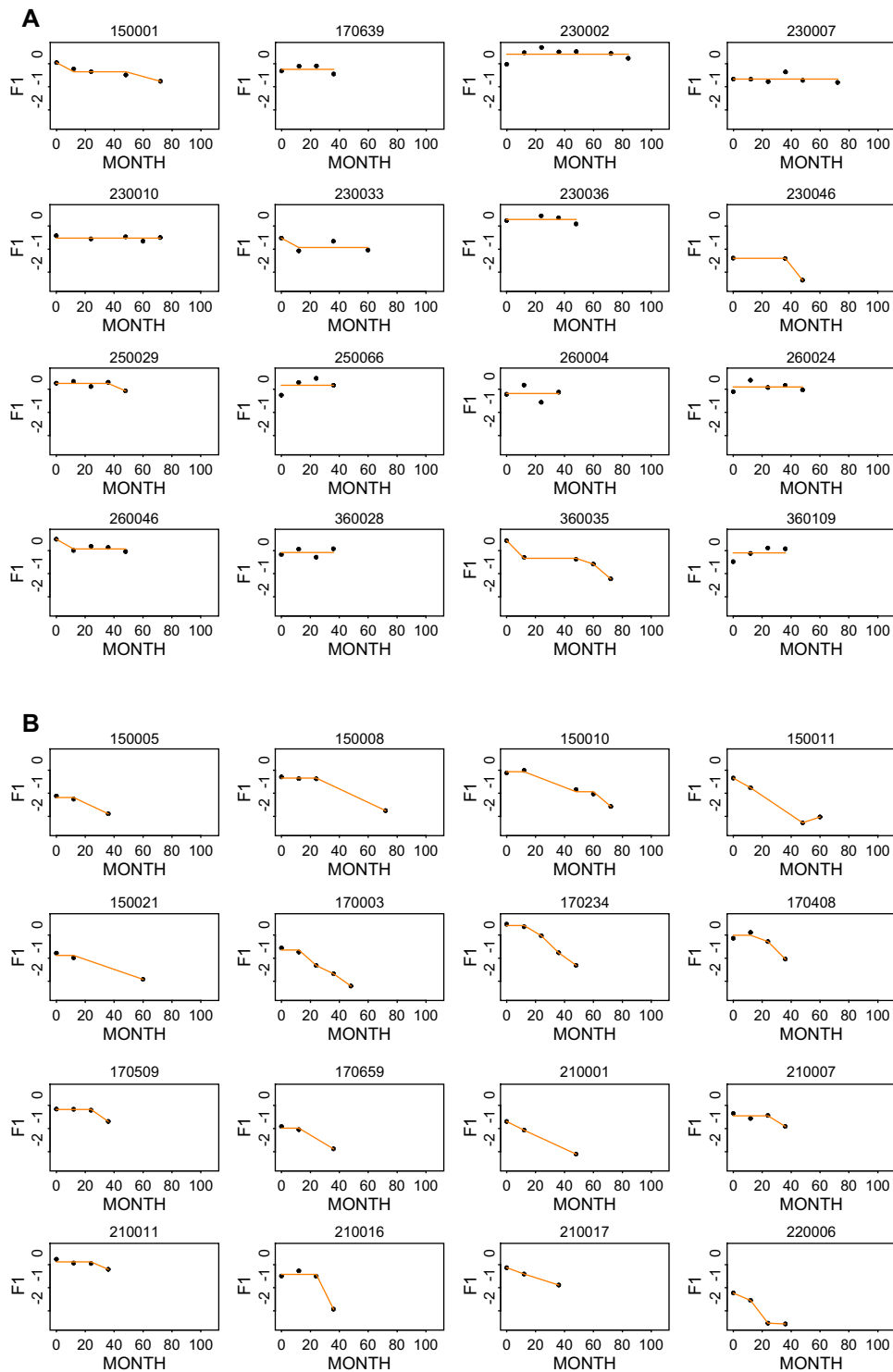


Fig. 2. Some characteristic F1 vs time scatter plots for the (A) plateau group and for the (B) non-plateau group subjects. The first 16 consecutive cases from the plateau and non-plateau groups are shown.

In most subjects, the plateau occurred at the beginning of the observation period. Because all subjects had to have an initial decline from normalcy to have received an AD diagnosis, a period of declining cognition preceded the observation period. This suggests an initial clinically evident manifestation of disease followed by a secondary (typically temporary)

halt in clinical disease progression. This is also the main finding of the recent article by Smith et al [14], although they looked at individuals in the pre-AD phase of memory decline.

Men and women were equally distributed among plateau and non-plateau groups. Men in the non-plateau group had slightly higher starting F1 scores but only a trend toward

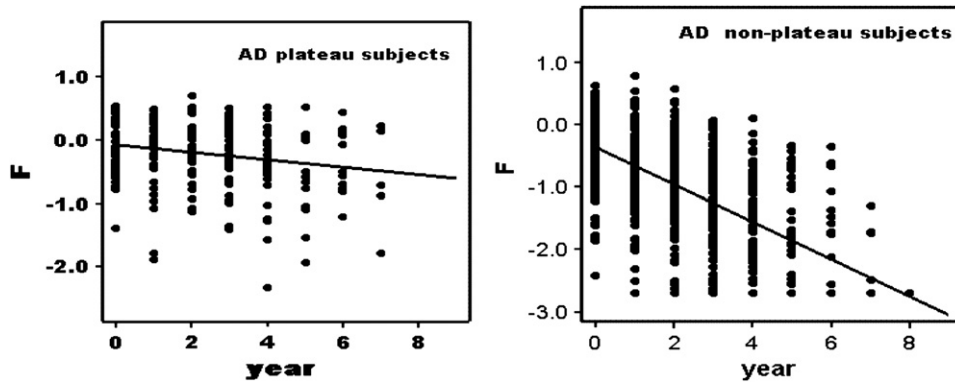


Fig. 3. F1 vs time for AD subjects with plateaus and without plateaus. Solid lines represent the corresponding $G = b_0 + b_1t$ regression lines. $G = (-.08 \pm .07) + (-.06 \pm .02) \cdot \text{Year}$ for AD subjects with plateaus, and $G = (-.36 \pm .06) + (-.30 \pm .02) \cdot \text{Year}$ for AD subjects without plateaus.

this in the plateau group, likely as a result of smaller sample size. The observation that men have milder disease at diagnosis is not new [16], suggesting a bias in detecting disease symptoms among family and physicians. Consistent with other observational studies of gender effects in progression of AD [20], we found no correlation with disease progression in the non-plateau group.

Educational attainment was significantly lower in the plateau than the non-plateau group. A recent study [21] similarly found that in individuals already diagnosed with AD, the annual rate of decline was greater in those with more education. The authors attributed this phenomenon to the comparatively greater disease burden reached before diagnosis. Our results support this finding because those in the CERAD data set with less education were more likely to present with plateaus during their course. We also note that this observation extends the more widely acknowledged finding that higher levels of education are a protective factor against developing AD [22,23].

The percentage of cases characterized as having plateaus is strongly influenced by the minimum length of time chosen to define a plateau, by the choice of corridor values, and by the floor value for F1. The longer the time period and the smaller the value of the corridor, the fewer individuals will be determined as exhibiting plateaus. By constraining the duration of a plateau to exceed the longest period of time required for a 1 SD drop in any of the component measures of F1 and then forcing the corridor to be less than 1 SD(F1), our choice of these parameters is conservative, possibly excluding some cases that clinicians would label as a plateau. There is not a single correct way to address this issue, and different methods will have different operating characteristics.

Implications of our findings include the need to counsel families not only that some patients decline more rapidly than others (as most physicians already do), but that some fortunate individuals might appear to stop progressing altogether for at least 3 years. Although this study did not find any obvious predictors of who would experience a plateau and who would not, at one fifth of the sample, this is not a neg-

ligible number of patients. In a disease that is widely considered uniformly relentless and unstoppable, this message of hope might be appreciated. As a corollary, there needs to be greater recognition of the interindividual variability that occurs in this disease, which is obscured by tracking only group change in research studies, as is currently done. Methods for measuring change at the individual level should be added to the arsenal of techniques used in both observational and experimental (ie, clinical trials) studies that determine the effectiveness and/or efficacy of preventative and therapeutic compounds.

We acknowledge several limitations.

1. Our data do not cover any patient's entire clinical course. Data collection does not cover the pre-AD period, and subjects entered the study at diverse disease points. Similarly, some subjects dropped out before death. Heterogeneity in data collection prevented us

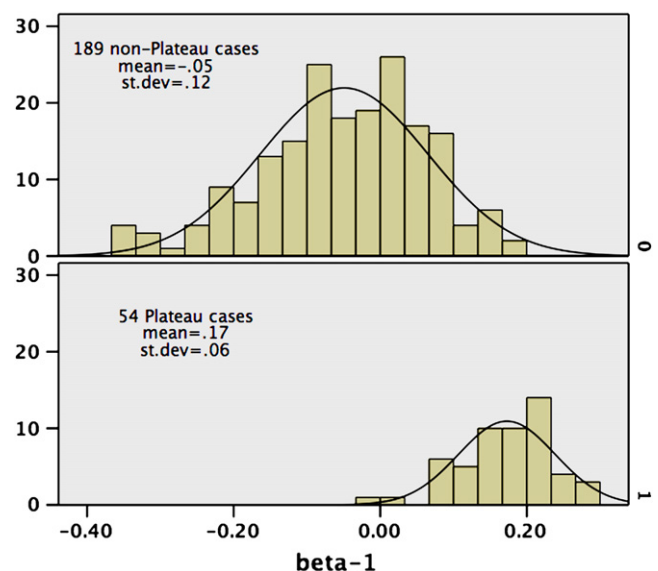


Fig. 4. Distribution of individual β_1 values for AD subjects with and without plateaus.

Table 2
Number of subjects (rows 1 and 2) and distribution of plateaus (row 3) by CDR stage

	0.5	1	2	3
Plateau group	2 (3.7%)	43 (79.6%)	8 (14.8%)	1 (1.9%)
Non-plateau group	8 (4.3%)	112 (60.2%)	58 (31.2%)	8 (4.3%)
Percentage of plateaus by CDR stage	20%	28%	12%	11%

from using a “true” plateau model, in which a period of stability is buttressed by steeper rates of decline.

2. PCA extracts the variability that is common to all cases; it does not account for the variability between cases. However, despite including all time points for both normal and AD subjects in our PCA, the first factor explained greater than 80% of the variance, suggesting that, in fact, this was a reasonable tool for data compression. We note that the assumptions of the PCA linear model might have contributed to plateaus by not accounting for covariance that is subsumed into the error terms of the PCA model.
3. Because our sample lacked pathologic confirmation, it is possible that the subgroup of individuals with plateaus represents a group of individuals without AD. In the most recent published pathologic confirmation of the CERAD data set [24], AD was confirmed in 176 of 201 (87.6%) subjects.
4. Our choice of cognitive variables was based primarily on available data, rather than representing a systematic inclusion of all cognitive domains. A data set with more variables might result in different findings.
5. A limitation of the original CERAD database was the high dropout rate, combined with many AD subjects' inability to cooperate with neuropsychological testing at later visits, which resulted in both a large attrition rate overall and selectively greater losses for the AD (78%) than the NC group (42%). Indeed, all four of the F1 component variables are slightly higher in our analyzed sample than in the total enrolled CERAD data set. This bias in data collection might have skewed our results more heavily toward AD subjects who were stable over time and artificially inflated the number with plateaus.

5. Implications for future work

Cholinesterase inhibitors and N-methyl-D-aspartate antagonists are known to produce only symptomatic improvements in cognition and behavior, although a recent study by Atri et al [25] demonstrated that combined therapy with both agents during a longer period of time might reduce the rate of cognitive and functional decline. Current clinical trial designs have been developed to detect disease-modifying effects in ways that would allow trial completion in a relatively short time (ie, 1 to 2 years) to avoid the burdensome costs and higher dropout rates of more prolonged studies. Our findings

suggest that identifying cognitive plateaus during such a short time span is misleading; more than 44% of the CERAD AD group would have met criteria for cognitive stability during a period of 2 years.

We chose to examine a historical patient population who demonstrated the natural history of AD. However, such populations are now difficult to identify, because many practitioners prescribe cholinesterase inhibitors for mild-moderate AD patients. If we were to apply a plateau-identifying algorithm to current populations, it would be important to learn how much stability can be expected from the use of symptomatic agents alone, with or without the inclusion of an individual who has a predisposition to cognitive stability.

One of the hardest issues for clinicians is knowing how to translate the results of clinical trials and epidemiologic studies to individual patients. Does a drug improve symptoms by 10% in every individual, or does it improve symptoms by 40% in some individuals and not at all in others? The results of the recent phase II study of bapineuzumab, in which *APOE* ϵ 4 positive individuals did not respond as vigorously as *APOE* ϵ 4 negative ones, highlight this phenomenon: an understanding of the natural history of a disease and its modifiers is essential to practicing evidence-based medicine.

When cognition in AD subjects is looked at in the aggregate, a steady decline over time is seen. However, individual patient trajectories vary considerably; some show a devastatingly rapid rate of decline throughout their course; others demonstrate long periods of stability. Such intersubject variation has long been ignored in the face of a relentlessly progressive disease, but we propose that such individual differences might be a much needed clue into the pathophysiology of underlying symptom progression. Subjects with cognitive plateaus might represent a group with a natural capacity for disease resistance and are worthy of further study.

References

- [1] Haxby VJ, Raffaele K, Gillette J, Schapiro BM, Rapoport IS. Individual trajectories of cognitive decline in patients with dementia of the Alzheimer type. *J Clin Exp Neuropsychol* 1992;14:575–92.
- [2] Helmes E, Merskey H, Fox H, Fry RN, Bowler JV, Hachinski VC. Patterns of deterioration in senile dementia of the Alzheimer type. *Arch Neurol* 1995;52:306–10.
- [3] Piccini C, Bracco L, Falcini M, Pracucci G, Amaducci L. Natural history of Alzheimer's disease: prognostic value of plateaus. *Neurol Sci* 1995;131:177–82.
- [4] Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1976;12:189–98.
- [5] Kazman R, Brown T, Thal LJ, Fuld PA, Aronson M, Butters N, et al. Comparison of rate of annual change of mental status score in four independent studies of patients with Alzheimer's disease. *Ann Neurol* 1988;24:384–9.
- [6] Teri L, McCurry SM, Edland SD, Kukull WA, Larson EB. Cognitive decline in Alzheimer's disease: a longitudinal investigation of risk factors for accelerated decline. *J Gerontol* 1995;50(A):49–55.
- [7] Yesavage JA, Poulsen SL, Sheikh J, Tanke E. Rates of change of common measures of impairment in senile dementia of the Alzheimer type. *Psychopharmacol Bull* 1988;24:531–4.

- [8] Han L, Cole M, Bellavance F, McCusker J, Primeau F. Tracking cognitive decline in Alzheimer's disease using the Mini-Mental State Examination: a meta-analysis. *Int Psychogeriatr* 2000;12:231–47.
- [9] Laird NM, Ware JH. Random effects models for longitudinal data. *Biometrics* 1982;38:963–74.
- [10] Jacqmin-Gadda H, Fabrigoule C, Commenges D, Dartigues JF. A 5-year longitudinal study of the Mini-Mental State Examination in Normal Aging. *Am J Epidemiol* 1997;145:498–506.
- [11] Brooks JO, Kraemer HC, Tanke ED, Yesavage JA. The methodology of studying decline in Alzheimer's disease. *J Am Geriatr Soc* 1993;41:623–8.
- [12] Helmes E, Merskey H, Fox H, Bowler JV, Hachinski VC. Profiles of cognitive decline in Alzheimer disease. *Alzheimer Dis Assoc Disord* 2002;16:123–7.
- [13] Hoyt BD, Massman PJ, Schatschneider C, Cooke N, Doody RS. Individual growth curve analysis of ApoE $\epsilon 4$ -associated cognitive decline in Alzheimer disease. *Arch Neurol* 2005;62:454–9.
- [14] Smith GE, Pankratz VS, Negash S, Machulda MM, Petersen RC, Boeve BF, et al. A plateau in pre-Alzheimer memory decline. *Neurology* 2007;69:133–9.
- [15] Welsh KA, Butters N, Mohs RC, Beekly D, Edland S, Fillenbaum G, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): part V—a normative study of the neuropsychological battery. *Neurology* 1994;44:609–14.
- [16] Heyman A, Peterson B, Fillenbaum G, Pieper C. The consortium to establish a registry for Alzheimer's disease (CERAD): part XIV—demographic and clinical predictors of survival in patients with Alzheimer's disease. *Neurology* 1996;46:656–60.
- [17] Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, et al. The Consortium to establish a registry for Alzheimer's disease (CERAD): clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 1989;39:1159–65.
- [18] Cattel C, Gambassi G, Sgadari A, Zuccalà G, Carbonin P, Bernabei R. Correlates of delayed referral for the diagnosis of dementia in an outpatient population. *J Gerontol A Biol Sci Med Sci* 2000;55:M98–102.
- [19] Morris JC, Edland S, Clark C, Galasko D, Koss E, Mohs R, et al. The consortium to establish a registry for Alzheimer's disease (CERAD): part IV—rates of cognitive change in the longitudinal assessment of probable Alzheimer's disease. *Neurology* 1993;43:2457–65.
- [20] Suh GH, Ju YS, Yeon BK, Shah A. A longitudinal study of Alzheimer's disease: rates of cognitive and functional decline. *Int J Geriatr Psychiatry* 2004;19:817–24.
- [21] Hall CB, Derby C, LeValley A, Katz MJ, Verghese J, Lipton RB. Education delays accelerated decline on a memory test in persons who develop dementia. *Neurology* 2007;69:1657–64.
- [22] Ott A, Breteler MM, van Harskamp F, Claus JJ, van der Cammen TJ, Grobbee DE, et al. Prevalence of Alzheimer's disease and vascular dementia: association with education: the Rotterdam study. *Br Med J* 1995;310:970–3.
- [23] Cobb JL, Wolf PA, Au R, White R, D'Agostino RB. The effect of education on the incidence of dementia and Alzheimer's disease in the Framingham Study. *Neurology* 1995;45:1707–12.
- [24] Mirra SS. Neuropathological assessment of Alzheimer's disease: the experience of the Consortium to Establish a Registry for Alzheimer's Disease. *Int Psychogeriatr* 1997;9(Suppl 1):263–8.
- [25] Atri A, Shaughnessy LW, Locascio JJ, Growdon JH. Long-term course and effectiveness of combination therapy in Alzheimer disease. *Alzheimer Dis Assoc Disord* 2008;22:209–21.

Appendix. Plateau-searching algorithm

Let N be the total number of patients, n_i the number of observations for the i^{th} patient, and Y_{it_j} be the F score of i^{th}

patient at time t_j $j = (1, 2, \dots, n_i)$, where the t_j 's are ordered so that $t_1 < t_2 < \dots < t_j < \dots < t_{n_i}$. The algorithm requires the choice of a corridor, c , around the prospective plateau value of the cognitive measure. The algorithm also requires the choice of ceiling value Y_{ceiling} and floor value Y_{floor} of the measure, within which a stable period must lie to potentially qualify as an apparent plateau.

The algorithm searches for a set of consecutive observations that constitutes a stable period (P_i) for the i^{th} subject, with an associated running mean equal to the average of the Y values within each set. The algorithm for the i^{th} patient is as follows:

1. Assume that the first observation, Y_{it_1} , is the starting point of a possible stable period P_i . Initialize the running mean as $\bar{Y}_i = Y_{it_1}$.
2. Calculate the difference between the next observation, Y_{it_2} , and the current running mean: $\Delta = Y_{it_2} - \bar{Y}_i$. In addition, keep track of the running sum of the Δ differences, Σ_i . (At the last point, the value of Σ_i will be the difference between the last and the first observation.)
3. Examine the Δ difference.
 - (a) If $-c \leq \Delta \leq 3c$, then accept Y_{it_2} in the stable period and update the associated running mean \bar{Y}_i as the average of Y_{it_2} and the observations previously included in P_i .
 - (b) Otherwise, assume that Y_{it_2} is the starting point of a possible stable period P_i , and initialize the running mean as $\bar{Y}_i = Y_{it_2}$.
4. Repeat steps 2 to 3 for all Y_{it_j} observation points.
5. If no stable period is found by the above criteria or if not all observations fall into the stable period but $\Sigma_i \leq c$, then accept all Y_{it_j} observation points into the P_i stable period, with the mean of all observations as the \bar{Y}_i plateau value.
6. Calculate the length of the stable period. If it lasts 36 months or longer and has an associated running mean between Y_{floor} and Y_{ceiling} , then that patient is categorized into the plateau group. Otherwise, the patient is categorized into the non-plateau group.
7. Let the p_i plateau-index be defined as the following:

$$p_i = \begin{cases} 1 & \text{if plateaus are detected in } i^{\text{th}} \text{ patient} \\ 0 & \text{otherwise} \end{cases}$$

8. Repeat steps 1 to 7 for all subjects, $i = 1, N$.