The CERAD Neuropsychological Battery Total Score and the Progression of Alzheimer's Disease

Heidi C. Rossetti, M.S.[1], **C Munro Cullum, PhD**[1,3], **Linda S. Hynan, PhD**[2,1], and **Laura Lacritz, PhD**[1]

[1]Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas

[2]Department of Clinical Sciences (Biostatistics), University of Texas Southwestern Medical Center at Dallas

[3]Department of Neurology, University of Texas Southwestern Medical Center at Dallas

Abstract

Objective—To establish the CERAD neuropsychological battery as a valid measure of cognitive progression in Alzheimer's disease (AD) by deriving annualized CERAD Total Change Scores and corresponding confidence intervals in AD and controls from which to define clinically meaningful change.

Method—Subjects included 383 Normal Control (NC) and 655 AD subjects with serial data from the CERAD registry database. Annualized CERAD Total Change Scores were derived and Reliable Change Indexes (RCIs) calculated to establish statistically reliable change values. CERAD Change Scores were compared to annualized change scores from the MMSE, Clinical Dementia Rating Scale (CDR) Sum of Boxes, and Blessed Dementia Rating Scale (BDRS).

Results—For the CERAD Total Score, the AD sample showed significantly greater decline than the NC sample over the four year interval, with AD subjects declining an average of 22.2 points compared to the NCs' improving an average 2.8 points from baseline to last visit [Group $\times$ Time interaction ($F(4,1031) = 246.08$, $p < .001$)]. By Visit 3, the majority of AD subjects (65.2%) showed a degree of cognitive decline that fell outside the RCI. CERAD Change Scores significantly correlated ($p<0.001$) with MMSE ($r = -.66$), CDR ($r = -.42$), and BDRS ($r = -.38$) change scores.

Conclusion—Results support the utility of the CERAD Total Score as a measure of AD progression and provide comparative data for annualized change in CERAD Total Score and other summary measures.

Introduction

Alzheimer's disease (AD) is a chronic, debilitating illness with a hallmark of progressive cognitive decline. The assessment of cognitive change over time is integral for the diagnosis and management of AD. In order to examine AD progression, clinicians and researchers require a standardized, brief, and reliable means of evaluating the presence and degree of cognitive impairment, with an emphasis on repeated assessment. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuropsychological battery was

Correspondence: Heidi Rossetti, 6363 Forest Park Rd, Suite 265, Dallas, TX 75235-8846, Heidi.Rossetti@utsouthwestern.edu, Phone: 214-577-5953, Fax: 214-648-4660.

developed to meet these needs by providing uniform longitudinal data on a large sample of AD patients.[1, 2] The CERAD project amassed a large database of rich clinical information with which to study the progression of AD.

Multiple studies have supported the reliability and validity of the CERAD neuropsychological battery in assessment of cognitive decline; however, until recently progression studies used the individual subtests, not a summary score.[3, 4, 5] A method for tabulating a total score for the CERAD neuropsychological battery was recently developed, providing a simple summary reference point for global cognitive performance.[6] The resulting CERAD Total Score demonstrated good test-retest reliability and differentiated between normal controls (NC) and AD subjects as well as the MMSE, with fewer false negatives.[6]

The CERAD Total Score enhances the value and utility of the CERAD neuropsychological battery but has yet to be examined in terms of AD progression. To this end, we examined the utility of the CERAD neuropsychological battery as a measure of progression and comparing annualized CERAD Change Scores to the annual rates of change in other established measures such as the MMSE.

Methods

The current study was based on the 1146 AD subjects and 465 NC in the CERAD registry database (CERAD Archive Rev 2.0 Version 8). Subjects with any missing data were excluded, and only subjects with a baseline (entry) assessment and at least one annual follow-up were included. Although previous studies excluded minorities to provide more uniform samples,[6, 8] problems of ethnic bias may be attenuated in progression studies since subjects serve as their own control.[9] This study included both Caucasian and African-American subjects; however, Hispanic subjects were excluded due to limited sample size and lack of follow-up data (N = 2 NC, 21 AD). CERAD registry controls with a Clinical Dementia Rating Scale (CDR) score ≥ 0.5 were excluded (N=27). The resulting sample size was 383 NC and 655 AD subjects.

Data through the fourth annual visit were included for a total of 5 years of testing data (baseline visit + 4 annual visits). There were 342 NC and 594 AD subjects with Visit 1 data (1-year follow-up), 259 NC and 375 AD subjects at Visit 2, 214 NC and 204 AD subjects at Visit 3, and 178 NC and 98 AD subjects at Visit 4. Of the 5 total possible visits, there were 95 NC and 275 AD subjects with 2 visits, 84 NC and 218 AD subjects with 3 visits, 86 NC and 103 AD subjects with 4 visits, and 118 NC and 64 subjects with 5 visits.

Controls consisted of spouses of CERAD AD subjects and community volunteers and the NC group remained normal throughout the study (CDR=0). AD subjects were diagnosed with probable or possible AD using National Institute for Neurological and Communicative Disorders and Stroke - Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) guidelines. [7] All subjects were age 50 or older, English speaking, not institutionalized, and free of co-morbid conditions that could affect cognition. AD subjects had a caregiver or informant who could provide history on the subject. All CERAD registry subjects were administered the CERAD neuropsychological battery, including the CDR. Additionally the Blessed Dementia Rating Scale (BDRS) was administered to all CERAD registry AD subjects. It should be noted that because CDR box scores were not initially included as part of the CERAD battery, the available sample size for the CDR Sum of Boxes was 183 subjects, versus the 655 subjects available for the CERAD, MMSE, and BDRS. Standard administration of the CERAD battery has been described previously.[1]

CERAD Total Scores were calculated according to the method derived by Chandler et al., which excludes the MMSE. [6] Each subject's CERAD Total Score was tabulated at the baseline visit and for each annual follow-up visit. CERAD Change Scores were derived by taking the difference of the Total Score at baseline from the Total Score at the next assessment such that negative numbers represent a decline in performance.

A meaningful degree of cognitive change was established by calculating Reliable Change Index (RCI) confidence intervals from the test-retest data for the NC group according to previously outlined Reliable Change methodology.[10] RCIs take into account measurement error in a test-retest setting and provide the degree of change required to exceed sources of error.[11] CERAD Change Scores exceeding the RCI 90% confidence interval (CI) represent statistically reliable change that occurs only 10% of the time by chance.

To evaluate the change in CERAD Total Score from baseline to the fourth year visit between AD and NC groups, a between-group comparison over time was analyzed using a mixed models analysis of covariance with age at baseline, sex, education and race as possible covariates; covariates remained in the final model if $p < 0.05$. In order to examine the impact of dementia severity on AD progression, annualized change scores for AD subjects with a baseline CDR Stage ≤ 1 were compared to those with a baseline CDR Stage ≥ 2 with an ANCOVA to control for the effects of education. RCIs were calculated to provide a statistical means for determining if AD CERAD Change Scores differed significantly between annual visits.

In order to explore how the annualized CERAD Change Score compared to the annual rate of change in other established cognitive measures, annualized change scores for the CERAD Total Score, MMSE, CDR Sum of Boxes, and BDRS were calculated by taking the difference between each subject's final assessment score and baseline assessment score and dividing by the number of total years between these assessments. Pearson correlations were used to test for associations between annualized CERAD Change Scores and the annualized change scores of the MMSE, CDR Sum of Boxes, and BDRS. A moderate correlation was defined as ± 0.4 to ± 0.6.

Results

Demographic characteristics

Descriptive information for the NC and AD groups from the CERAD registry samples can be viewed in Table 1. NC subjects were significantly younger and more educated than AD subjects. Actual mean differences for the groups were small, but moderate effect sizes were observed (age $d = .41$, education $d = .40$). There was a larger proportion of females in the NC group compared to the AD sample, and in both groups there was a larger proportion of Caucasian subjects. Approximately 60% of the AD group obtained a CDR Stage 1 rating at entry.

CERAD Change Scores

Longitudinal CERAD Change Scores from baseline to each follow-up visit (i.e., the difference from baseline to Visit 1, baseline to Visit 2, etc.) for AD and NC subjects are presented in Figure 1A, and illustrate the relatively stable performance of the NC sample compared to the steadily declining trajectory of the AD group. The change in mean Total Score from baseline to each visit for the 64 AD and 118 NC subjects with follow-up data at each time interval is shown in Figure 1B. This subset of subjects showed a similar pattern of results to that found in the larger sample, with stable NC scores and declining performance in AD.

Significantly greater annual change in CERAD Total Score for the AD group versus controls was observed from the baseline visit to fourth visit, after accounting for the significant covariates (age at baseline, education, and race; time by group interaction: $F(4,1031) = 246.08$, $p < 0.001$), with AD subjects declining an average of 22.2 points compared to an average 2.8 point *gain* for NC subjects (see Table 2).

Reliable Change Index

Data for the RCI calculations are provided in Table 3. The majority of AD subjects did not exhibit a degree of cognitive change that exceeded the confidence intervals established by the RCI until Visit 3 when 65.2% of the AD group exhibited a decline in CERAD Change Scores that exceeded the RCI. When examining the annualized change in CERAD Total Score from baseline to each individual's last visit, the majority of AD subjects (55.1%) exceeded the cut-offs established by the RCI.

Impact of dementia severity

There was no significant difference in mean annual change by level of severity [CDR $\leq 1 = -7.03$, CDR $\geq 2 = -7.50$; $F(1, 653) = 1.34$, $MSE = 1.34$, $p = .247$, partial $\eta^2 = .002$], indicating that the level of dementia severity with which a subject entered the study did not impact change in CERAD performance over time.

Correlation with other measures

Total Scores of the AD sample on the MMSE, CDR Sum of Boxes, and BDRS at baseline, last visit, and the annualized Change Scores are presented in Table 4. Annualized CERAD Change Scores correlated significantly with all other measures ($p < .001$), and the highest correlation was seen with the MMSE, $r(654) = .66$, followed by the CDR Sum of Boxes, $r(181) = -.42$, and the BDRS, $r(654) = -.38$. Annualized CERAD Change Scores were at least moderately correlated with the annualized scores of the established measures, with the exception of the BDRS. Annualized MMSE Change Scores also correlated significantly ($p < .001$) with the CDR Sum of Boxes ($r = -.42$) and the BDRS ($r = -.41$), and annualized CDR Sum of Boxes Change Scores with the BDRS ($r = .60$).

Discussion

A composite score for the CERAD neuropsychological battery was recently developed[6], but until now the utility of the Total Score had not been examined in terms of AD progression. As expected, AD subjects showed significantly greater annual change in Total Scores than NC subjects. In order to establish a meaningful degree of cognitive change, the annual Change Scores of the AD and NC groups were compared. According to the 90% CI established by the RCI, CERAD Change Scores should exceed at least ± 10.51 points in order to represent clinically meaningful change. This threshold provided by the RCI increases the likelihood that an individual's change in performance reflects an actual change in cognitive abilities rather than extraneous factors such as subject variability. The mean annualized rate of AD progression as measured by the CERAD battery was determined to be -7.2 points, and it was not until Visit 3 that the majority of AD subjects exhibited a clinically significant decline in their performance as determined by the calculated RCIs.

Several factors were examined to better understand the lack of clinically significant change in years 1 and 2 (based on the RCI). To assess the potential impact of dementia severity, change scores were reviewed based on CDR at baseline and no differences were seen when comparing those with baseline CDRs of 1 versus 2. This was surprising given the substantial amount of research identifying dementia severity as a key factor in AD progression.[12] Although the mild severity group (CDR Stage ≤ 1; N=411) had significantly better CERAD

and MMSE total scores than the moderate to severe group (CDR Stage ≥ 2; N=244) at baseline and last visit, there was no significant difference in the rate of decline. However, it should be noted that the range of CDR scores was restricted at baseline (.5 – 3.0) and the majority of subjects (59%) were classified at CDR Stage 1, limiting the number of subjects in other stages that were available for comparison. Since the AD sample's CERAD performance was fairly impaired at baseline there may not have been much room for further decline, resulting in floor effects.

The potential impact of selective attrition on CERAD Change Scores was also examined to identify if lower-functioning individuals with AD may have dropped out earlier in the study than less-affected individuals, thus biasing the remaining sample with a disproportionate contribution from comparatively higher-functioning individuals. In fact, comparison of the average baseline CERAD performance of subjects with no follow-up data to those with at least one follow-up assessment revealed that those individuals who dropped out were significantly more impaired (CERAD M (*SD*) = 30.78 (*14.07*); MMSE M (*SD*) = 13.71 (*6.72*) than those who remained in the study (CERAD M (*SD*) = 39.46 (*12.97*); MMSE M (*SD*) = 18.50 (*5.10*). For the 64 AD subjects who had follow-up data at each time interval, the mean annualized decline was -5.4 (*3.69*), which is 1.8 points less than the mean of the larger sample (*N* = 655). These results provide some support for the idea that subjects who were experiencing less pronounced dementia remained in the study longer, which suggests that the progression rates found in the current study may not apply as readily to individuals with more severe dementia.

The annual CERAD Change Score was compared to the annual rate of change in the MMSE, CDR Sum of Boxes, and BDRS, which are established measures of cognitive decline.[13, 14, 15] The mean annualized rate of change in the AD sample was -7.20 *(6.93)* on the CERAD, -3.43 *(3.33)* on the MMSE, 2.28 *(2.13)* on the CDR, and 1.89 *(1.54)* on the BDRS. The mean rates of change for the MMSE and BDRS are in keeping with prior research.[16, 17, 18] The CERAD Change Score was significantly and moderately correlated with each measure with the exception of the BDRS, which fell slightly short of the 0.4 – 0.6 range (*r* = -.38), likely due to the fact that the BDRS is based upon caregiver report of functioning rather than an objective assessment of cognitive functioning per se. These results support the concurrent validity of the CERAD neuropsychological battery in measuring AD progression and provide comparative data for annualized change in the CERAD Total Score and other common measures.

This study was subject to the inherent limitations of longitudinal designs, namely, selective attrition that resulted in successively smaller *n*s and inconsistency in the subjects returning at each visit. However, when comparing subjects with consistent annual follow-ups to the total sample, the results were similar, as shown in Figure 1. This study was limited to Caucasian and African-American subjects, so the CERAD Total Score's utility as a measure of progression among other ethnic populations is unknown. As reflected in the composition of the CERAD registry, this sample has a relatively high level of education (*M* = 13.3 years). It is possible that the progression of individuals with limited education may manifest differently than the results reported here.

The RCI method of determining clinically significant change in scores utilized in this study is well researched; however, it is relatively conservative.[19] As such, it may have provided overly wide confidence intervals that required AD subjects to exhibit a larger degree of change in order to be categorized as "meaningful." Therefore, Change Scores that do not exceed the specified confidence interval may not necessarily be insignificant but should be interpreted within the context of other sources of clinical information.[19] Although the confidence interval provided by RCIs is useful, once the degree of change extends beyond

the critical range it does not provide information regarding the relative magnitude of the change.[11] Given that in a sample of AD affected individuals each annual assessment would be expected to reveal significant change in cognitive performance, it may be that the RCI method is in fact too conservative for this population in assessing cognitive progression. Furthermore, while the RCI method takes into account standard error of measurement, it only measures one point in time and may miss variability that occurred between assessments.

The current study establishes the CERAD Total Score as a measure of AD progression comparable to other common measures such as the MMSE, CDR, and BDRS. Additional research geared toward exploring the utility of the CERAD Total Score in the staging of AD would be valuable. The establishment of cut-off scores for mild, moderate, and severe impairment would allow the CERAD battery to be used more precisely as a dementia staging tool. In addition, the CERAD battery could be utilized as a measure of change in intervention studies with dementia populations by examining the change in CERAD Total Score to assess the effectiveness of pharmacological or other interventions.

Acknowledgments

References

1. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 1989;39:1159–1165. [PubMed: 2771064]

2. Fisher NJ, Rourke BP, Bieliauskas LA. Neuropsychological subgroups of patients with Alzheimer's disease: An examination of the first 10 years of CERAD data. Journal of Clinical and Experimental Neuropsychology 1999;21:488–518. [PubMed: 10550808]

3. Morris JC, Edland S, Clark CC, 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;4:2457–2465. [PubMed: 8255439]

4. Welsh KA, Butters N, Hughes JP, et al. Detection and staging of dementia in Alzheimer's disease. Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's Disease. Archives of Neurology 1992;49:448–452. [PubMed: 1580805]

5. Welsh-Bohmer KA, Mohs RC. Neuropsychological assessment of Alzheimer's disease. Neurology 1997;49:S11–S13. [PubMed: 9310507]

6. Chandler MJ, Lacritz LH, Hynan LS, et al. A total score for the CERAD neuropsychological battery. Neurology 2005;65:102–106. [PubMed: 16009893]

7. McKhann G, Drachmann D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939–944. [PubMed: 6610841]

8. Welsh KA, Butters N, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. Neurology 1994;44:609–614. [PubMed: 8164812]

9. Fillenbaum GG, Woodbury MA. Typology of Alzheimer's disease: Findings from CERAD data. Aging and Mental Health 1998;2:105–127.

10. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology 1991;59:12–19. [PubMed: 2002127]

11. Hermann BP, Seidenburg M, Schoenfeld J, et al. Empirical techniques for determining the reliability, magnitude, and pattern of neuropsychological change after epilepsy surgery. Epilepsia 1996;37:942–950. [PubMed: 8822692]

12. Storandt M, Grant EA, Miller JP, et al. Rates of progression in mild cognitive impairment and early Alzheimer's disease. Neurology 59:1034–1041. [PubMed: 12370458]

13. Berg L. Clinical Dementia Rating (CDR). Psychopharmacology Bulletin 1988;24:637–639. 2002. [PubMed: 3249765]

14. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral gray matter of elderly subjects. British Journal of Psychiatry 1968;114:797–811. [PubMed: 5662937]

15. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State:" A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research 1975;12:189–198. [PubMed: 1202204]

16. Van Belle G, Uhlmann RF, Hughes JP. Reliability of estimates of changes in mental status test performance in senile dementia of the Alzheimer type. Journal of Clinical Epidemiology 1990;43:589–595. [PubMed: 2348211]

17. Clark CM, Sheppard L, Fillenbaum GG, et al. Variability in annual Mini-Mental State Examination Score in patients with probable Alzheimer's disease. Archives of Neurology 1999;56:857–862. [PubMed: 10404988]

18. Holmes C, Lovestone S. Long-term cognitive and functional decline in late onset Alzheimer's disease: therapeutic implications. Age and Ageing 2003;32:200–204. [PubMed: 12615565]

19. Chelune GJ, Naugle RI, Luders H, et al. Individual change after epilepsy surgery: Practice effects and base-rate information. Neuropsychology 1993;7:41–52.
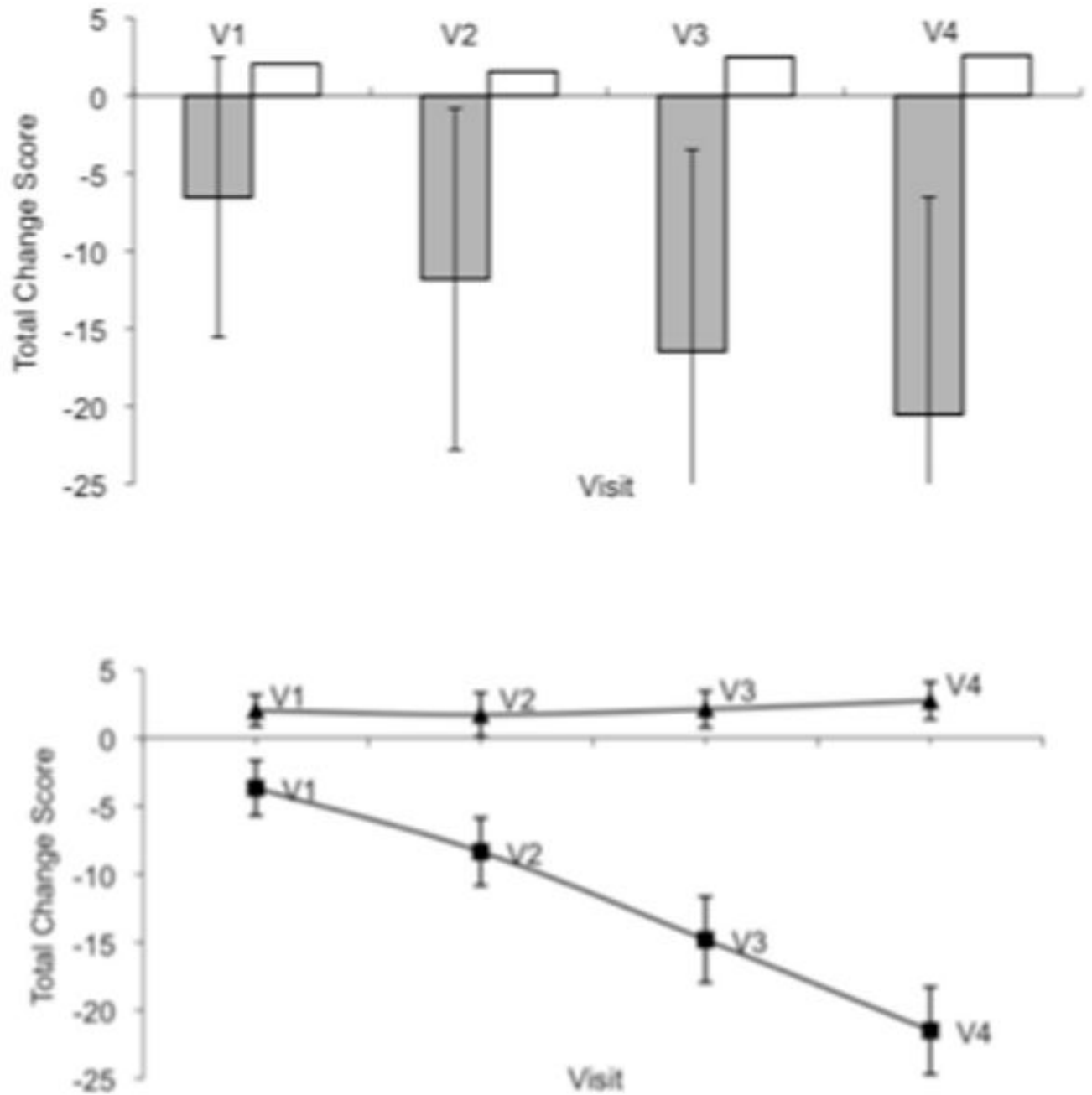
Figure 1.
Figure 1A. ■ AD □ NC
Mean CERAD total change score and standard deviations from baseline to each follow-up visit, N = 655 AD, 383 NC
Figure 1B. ▪ AD ▲ NC
Mean CERAD total change score and standard deviations from baseline to each follow-up visit for subjects with data at each time point, N = 64 AD, 118 NC.
Note. V1 through V4 denotes the first annual visit through the fourth annual visit.

Table 1
Characteristics of CERAD registry samples

Characteristic	AD (N = 655) M (SD)	NC (N = 383) M (SD)
Age at Baseline[*]	72.1 (7.9)	68.9 (7.9)
Education[*]	12.6 (3.7)	13.9 (3.1)
Baseline CERAD[*]	39.4 (13.0)	79.8 (9.4)
Baseline MMSE[*]	18.48 (5.1)	28.81 (1.6)
Baseline BDRS	4.3 (2.4)	-
Gender [N (%)][*]		
Male	281 (43)	129 (34)
Female	374 (57)	254 (66)
Race [N (%)][*]		
Caucasian	560 (86)	360 (94)
Black	95 (14)	23 (6)
CDR Stage [N (%)][^]		
.5	25 (3.8)	-
1	386 (58.9)	-
2	220 (33.6)	-
3	24 (3.7)	-

[*] Significant difference between AD and NC groups at p < .001 using independent samples t tests for continuous variables and χ^2 for categorical variables.

[^] All NCs were CDR Stage 0. Stage 0 = no dementia, 0.5 = questionable, 1 = mild, 2 = moderate, 3 = severe.

CERAD = Consortium to Establish a Registry for Alzheimer's Disease; NC = normal control; AD = Alzheimer's disease; MMSE = Mini-Mental State Examination; BDRS = Blessed Dementia Rating Scale, CDR = Clinical Dementia Rating Scale

Table 2

CERAD Total Score Means and Standard Errors with Covariates of Age at Baseline, Education, and Race

	AD	NC
Time Interval	**Mean (SE)**	**Mean (SE)**
Baseline	40.60 (0.49)	78.93 (0.64)
Visit 1	34.42 (0.53)	81.08 (0.66)
Visit 2	28.32 (0.59)	80.68 (0.71)
Visit 3	22.80 (0.70)	81.38 (0.78)
Visit 4	18.42 (0.87)	81.73 (0.84)

Note. p values for education and race were < 0.001 and age at baseline was < 0.002.

CERAD = Consortium to Establish a Registry for Alzheimer's Disease; NC = normal control; SE = standard error of the mean.

Table 3
Test-Retest Reliability Coefficients and Reliable Change Index Scores Based on Standard Errors of Measurement for CERAD NC Sample

Time Interval	Test-Retest Reliability	SE$_m$	SE$_{diff}$	RC (90%) CI
Baseline to Visit 1	.80	4.53	6.41	± 10.51
Baseline to Visit 2	.65	6.67	9.43	± 15.47
Baseline to Visit 3	.76	5.12	7.24	± 11.87
Baseline to Visit 4	.69	6.12	8.66	± 14.20
Baseline to Last Visit	.74	5.52	7.80	± 12.80

Note. Test-retest reliability coefficients based on the correlation between the mean CERAD Total Score at baseline, (M = 79.8, SD = 9.4), and each subsequent visit.

CERAD = Consortium to Establish a Registry for Alzheimer's Disease; NC = normal control; SE$_m$ = standard error of measurement; SE$_{diff}$ = standard error of difference; RC = reliable change; CI = confidence interval.

Table 4
Summary Measure Total Scores for AD Sample

	Total Score at Baseline M (SD)	Total Score at Last Visit M (SD)	Annualized Total Change Score M (SD)
CERAD	39.35 (13.01)	25.41 (16.37)	-7.20 (6.93)
MMSE	18.48 (5.07)	11.18 (7.46)	-3.43 (3.33)
CDR	7.43 (3.26)	12.71 (4.98)	2.28 (2.13)
BDRS	4.30 (2.36)	9.82 (4.32)	1.89 (1.54)

Note. $N = 655$ for all measures except the CDR ($N = 183$).

AD = Alzheimer's disease; CERAD = Consortium to Establish a Registry for Alzheimer's Disease; MMSE = Mini Mental State Examination; CDR = Clinical Dementia Rating Scale Sum of Boxes; BDRS = Blessed Dementia Rating Scale