



Published in final edited form as:

Neuropsychology. 2013 July ; 27(4): . doi:10.1037/a0032707.

Age and APOE genotype influence rate of cognitive decline in non-demented elderly

David P. Salmon¹, Steven H. Ferris², Ronald G. Thomas^{1,3}, Mary Sano⁴, Jeffery L. Cummings⁵, Reisa A. Sperling⁶, Ronald C. Petersen⁷, and Paul S. Aisen¹ for the Alzheimer's Disease Cooperative Study

¹Department of Neurosciences, University of California, San Diego, CA 92093 USA

²Department of Psychiatry, New York University, New York City, NY 10016 USA

³Department of Family and Preventive Medicine, University of California, San Diego, CA 92093 USA

⁴Department of Psychiatry, Mount Sinai School of Medicine, New York City, NY 10029 USA

⁵Cleveland Clinic Lou Ruvo Center for Brain Health, Cleveland Clinic Neurological Institute, Las Vegas, NV 89106 USA

⁶Department of Neurology, Harvard Medical School, Boston, MA 02114 USA

⁷Department of Neurology, Mayo Clinic, Rochester, MN 55905 USA

Abstract

Objective—This study examined the impact of age and apolipoprotein E (APOE) genotype on the rate of cognitive decline in non-demented elderly participants in a simulated Alzheimer's disease (AD) primary prevention treatment trial carried out by the Alzheimer's Disease Cooperative Study.

Method—Cognitive tests were administered at baseline and at four subsequent annual evaluations to 417 non-demented participants (172 men, 245 women) between the ages of 74 and 93 (mean=79.13 ± 3.34). APOE genotyping was available for 286 of the participants.

Results—Four-year decline was evident on measures of orientation, memory, executive function and language. Faster decline was evident in APOE 4+ (a genetic risk factor for AD; n=73) than 4- participants (n=213), even after controlling for education, gender, ethnicity, and baseline functional and cognitive abilities. This discrepancy increased with increasing age indicating an age X genotype interaction.

Conclusions—These results are consistent with population-based studies, and extend the findings to a carefully-screened sample that meets inclusion and exclusion criteria for an AD primary prevention trial. The interaction between age and APOE genotype on rate of decline suggests that preclinical disease may be over represented in older 4+ individuals. Thus, APOE genotype and age should be considered in the design of AD primary prevention treatment trials.

Address correspondence to: David P. Salmon, Department of Neurosciences (0948), University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0948, Telephone: (858) 246-1253, dsalmon@ucsd.edu.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/neu

Keywords

Cognitive decline; Apolipoprotein E; Aging

Introduction

A number of longitudinal studies have shown that the cognitive test performance of non-demented elderly adults often declines over the course of only a few years and that the rate of decline may be influenced by factors that confer risk for Alzheimer's disease (AD; Jonkers et al., 1998; O'Hara et al., 1998; Small et al., 1998; Wilson et al., 2002; Yaffe et al., 1997). Mayeux and colleagues (2001), for example, found that a large group of healthy elderly individuals assessed with a battery of cognitive tests on three occasions over a seven year period declined significantly on a memory composite score, but not on composite scores for tests of language or visuospatial and general cognitive abilities. A faster rate of decline on the memory score was associated with the presence of an apolipoprotein E (APOE) 4 allele, a genetic risk factor for late onset AD, even after age, education and ethnicity were taken into account. Although the cohorts in this and other studies with similar findings (Jonkers et al., 1998; O'Hara et al., 1998; Small et al., 1998; Yaffe et al., 1997) averaged more than 70 years of age, subsequent studies that included healthy younger adults showed that decline in memory performance and the association between rate of decline and APOE 4 genotype occurred as early as age 50 to 60 years (Caselli et al., 2004; 2009; 2011). Studies that followed the outcome of participants beyond the preclinical stage suggest that the relationship between faster memory decline and APOE 4 genotype in non-demented elderly may reflect the over representation of preclinical AD in those with the risk factor (e.g., Bondi et al., 1999; Bunce et al., 2004).

Clear understanding of the relationship between cognitive decline and APOE genotype, age, and other demographic characteristics in non-demented elderly is important as clinical trials are developed to test treatments during the preclinical stage of AD that precedes the manifestation of the dementia syndrome (Aisen et al., 2011). Any treatment effect will have to be detected against the background of cognitive changes related to these genetic and demographic factors. The studies reviewed above are informative in this regard, but most were based on population-based or convenience samples that may differ from the typical cohort recruited for a clinical trial, and they usually had a lower frequency of evaluation (e.g., every 2 to 3 years) than most clinical trials. The present study examined the rate of change in the cognitive test performance of a large group of non-demented elderly individuals who were assessed annually over the course of four years in a simulated clinical trial. As part of the simulated trial, all participants satisfied inclusion and exclusion criteria for a typical AD primary prevention trial including being at least 75 years of age, non-demented, in good physical and mental health, and not having medical conditions or medications that would typically cause their exclusion from a clinical trial (see Ferris et al., 2006 for a full list of exclusion criteria). All participants were also required to have a qualified study partner who could knowledgeably rate the participant's behavior and ability to perform usual activities of daily living.

Participants were assessed at baseline and at yearly intervals with a battery of neuropsychological tests that included standard measures of global cognitive function, episodic memory, language, executive function, and visual attention that have been widely used to detect and track early AD in previous studies (for review, see Salmon & Bondi, 2009). This design allowed a number of questions to be addressed: 1) How rapidly does performance of non-demented elderly change over time on neuropsychological tests that are sensitive to early AD? 2) Does decline occur on tests across all cognitive domains? 3) Is the

rate of decline influenced by AD risk factors such as age and APOE 4 genotype? 4) Is the influence of AD risk factors on the rate of decline the same across all tests, or are some tests or domains more susceptible than others? To address these questions, the interaction between age and APOE genotype on rate of cognitive change was assessed. The impact of education, gender, and ethnicity on test performance was also evaluated given that some of these demographic factors influenced the performance of non-demented elderly in previous studies (Jonkers et al., 1998; Mayeux et al., 2001; O'Hara et al., 1998; Small et al., 1998; Wilson et al., 2002; Yaffe et al., 1997).

Method

Participants

A total of 644 non-demented elderly individuals (269 men, 375 women) between the ages of 74 and 93 were initially enrolled in this study. The participants were recruited from 29 Alzheimer's Disease Cooperative Study (ADCS) sites. Each site recruited at least 16 participants and 20% of those enrolled at each site had to be non-white (see Ferris et al., 2006 for a detailed description of the recruitment procedures). The total sample averaged 79.05 (± 3.61) years of age, 14.96 (± 3.08) years of education, and 95.28 (± 3.68) points on the Modified Mini-Mental State Examination (mMMSE) at the time of enrollment (i.e., baseline). There were 497 participants (77.2%) with a Clinical Dementia Rating (CDR) of 0 and 147 (22.8%) with a CDR of 0.5. The sample had 497 (77.2%) white and 147 (22.8%) non-white participants. The APOE allele distribution was 3 2 2, 40 2 3, 266 3 3, 12 2 4, 86 3 4, and 4 4 4. APOE genotype was not determined for 233 of the 644 participants. Information regarding family history of dementia was not obtained. The study was approved by the institutional review boards at UCSD and each of the participating ADCS sites. Informed consent was obtained from all study participants after the procedures of the study had been fully explained.

Over the four-year course of the study 227 participants dropped out at various points before the last evaluation leaving a final sample of 417 participants (172 men, 245 women) who completed the study. Only data from the 417 participants who completed the study are analyzed and included in this report. The final sample averaged 79.13 (± 3.34) years of age, 15.27 (± 2.98) years of education, and 95.83 (± 3.37) points on the mMMSE at baseline. There were 335 participants (80.3%) with a Clinical Dementia Rating (CDR; Hughes et al., 1982) of 0 and 82 (19.7%) with a CDR of 0.5. The final sample had 340 (81.5%) white and 77 (18.5%) non-white participants. The APOE allele distribution was 1 2 2, 31 2 3, 181 3 3, 8 2 4, 62 3 4, and 3 4 4. APOE genotype was not determined for 131 participants who completed the study. Participants who did not complete the study were significantly older (80.26 ± 3.98 ; $t(642)=3.83$; $p<.001$, Cohen's $d=.31$), less educated (14.38 ± 3.17 ; $p<.001$, $d=.29$), and had lower mMMSE scores (94.26 ± 4.01 ; $p<.0001$, $d=.43$) than those who completed. Non-whites had a higher drop-out rate (47.6%) than whites (31.6%). The percentage of APOE 4+ (i.e., at least one 4 allele) and 4- subjects who did not complete the study was similar in the Young-Old (as determined by a median split on age) (27% versus 28%, respectively) and in the Old-Old (33% versus 35%, respectively). For participants who completed the study, there was no significant difference in age, education, mMMSE score, gender distribution or ethnicity distribution between those genotyped or not genotyped (all p 's $> .05$).

Procedure

As part of the ADCS Prevention Instrument Project (Ferris et al., 2006) a battery of cognitive tests was administered to participants at baseline and at four subsequent annual evaluations (i.e., baseline and months 12, 24, 36, and 48). Participants were tested

individually in a quiet, well-lit room by a trained psychometrist. The cognitive battery consisted of a test of global cognitive function (mMMSE) and specific measures of episodic memory (Free and Cued Selective Reminding Test, NYU Paragraph Recall Test), language (modified Boston Naming Test, Semantic Verbal Fluency Test), executive function (Trail-Making Test: Parts A and B; Digit Symbol Substitution Test), and visual attention (Cancellation Test). The tests have been described in detail previously and are briefly described below:

Modified Mini-Mental State Examination (mMMSE; Teng & Chui, 1987)—This adaptation of the standard MMSE is a global test of mental status that briefly assesses orientation, language, verbal recall and recognition, and constructional ability. The mMMSE is scored on a 100 point scale with higher scores indicative of better cognitive function.

Free and Cued Recall Selective Reminding Test (FCSRT; Grober et al., 1988; Grober & Kawas, 1997)—On the first trial of this episodic memory test, subjects are asked to remember a 16-item word list that is presented visually and verbally with semantic cuing to facilitate encoding. Immediately after presentation, the words must be recalled freely or following a semantic category cue (to facilitate retrieval). Items recalled spontaneously (i.e., free) and after a cue are recorded. On subsequent trials, the subject is reminded only of items that were not recalled on the previous trial, and then must recall all 16 items freely or following a cue. Three trials are presented and the number of correctly recalled items is summed across trials making the maximum possible score 48 points.

NYU Paragraph Recall Test (Kluger et al., 1999)—This task requires subjects to remember a brief story that is read aloud to them. Recall of the story is assessed immediately after it is read and again after a 5 minute delay filled with unrelated testing. One of six alternative versions that are of equal difficulty is used at each test session. The test is scored for the number of story components (or ideas) that are correctly recalled immediately and after a delay for a total of 21 possible points for each condition.

Trail Making Test: Parts A and B (Army Individual Test Battery, 1944; Reitan, 1958)—Part A of this test consists of 25 circles numbered 1 through 25 distributed over a white sheet of 8 1/2 × 11 paper. The subject is instructed to connect the circles with a pencil line as quickly as possible in ascending numerical order. Part B also consists of 25 circles, but these circles are either numbered (1 through 13) or contain letters (A through L). Now the subject must connect the circles while alternating between numbers and letters in an ascending order (e.g., A to 1; 1 to B; B to 2; 2 to C). Parts A and B of the Trail-Making Test both require visuomotor and perceptual scanning skills, but Trails B also requires cognitive flexibility in shifting from number to letter sets under time pressure. The subject's performance is judged in terms of the time (in seconds) required to complete each trail (Trails A: 150 second maximum; Trails B: 300 second maximum) with higher scores indicative of poorer performance. To more effectively assess the executive function aspect of the Trail-Making Test the analysis was repeated using a derived difference score between time to complete Part B and time to complete Part A. This measure corrects for the attention or psychomotor speed aspects of the test that are common to both Parts A and B.

WAIS-R Digit Symbol Substitution Test (Wechsler, 1981)—In this timed test, subjects are presented with a key which associates nine unfamiliar symbols with the numbers 1 through 9. They are then asked to use the key to draw the appropriate symbols below a random series of their associated numbers as quickly as possible for 90 seconds. The number of correctly completed symbols is the score of interest and can range from 0 to 64 with higher scores indicative of better performance.

Boston Naming Test (BNT):10-item version (Kaplan et al., 1983)—This abbreviated version of the BNT requires the subject to name 10 objects depicted in outline drawings. The drawings are graded in difficulty with the easiest drawings presented first. If a subject encounters difficulty in naming an object, a stimulus or phonemic cue is provided. Correct responses produced spontaneously and after semantic cues are summed to provide the score of interest for a maximum score of 10.

Verbal Fluency Test (“Animals”) (Borkowski et al., 1967; Butters et al., 1987)—This test of language requires subjects to verbally generate as many different “animals” as possible in one minute. The total number of correct exemplars produced, excluding repetitions and variants (e.g., horse, horses), is the score of interest.

Cancellation Test (Mohs et al., 1997)—In this timed test of visual attention the subject is presented with an 8.5 × 11 inch piece of paper which contains 240 digits placed randomly in 8 rows of 30 and asked to cross-off (i.e., cancel), as quickly as possible, all of 2 specified target numbers (e.g., 2 and 8). Each quadrant of the paper contains 10 target stimuli for a total of 40 targets. The number of targets correctly crossed-off in 45 seconds is recorded.

Data Analysis

Data analyses were limited to completers because there were a number of significant demographic and cognitive differences at baseline between those who completed or did not complete the study (see above). First, one-way repeated-measures Analysis of Variance (ANOVA) was used to examine change over time (i.e., baseline, 12-month, 24-month, 36-month, 48-month) on the primary measure from each cognitive test. Student’s t-tests were used to determine if change scores (between baseline and the final annual evaluation) for each test were significantly different from zero. Second, a linear mixed effects model fit by restricted maximum likelihood (REML; Laird & Ware, 1982) was derived for each cognitive test to examine the effects of age and APOE genotype (4– vs. 4+) on change over time on that measure. Each model included terms for time, age, genotype, education, gender and ethnicity, as well as age X time, time X genotype, age X genotype and age X genotype X time interaction terms. Each model also included baseline scores on a measure derived from the components of the CDR known as “sum of boxes” (CDR-SB) to control for initial level of global cognitive status. CDR sum of boxes was used to control for initial level of global cognitive status, rather than mMMSE score, so that change over time on the mMMSE could be examined. The REML analyses were limited to those participants with known APOE genotype (n=286). The mean age, years of education, mMMSE, and CDR sum of boxes scores for participants with no 4–) or at least one 4+) APOE 4 allele who completed the study and included in the REML analyses are shown in Table 1. The percentage of participants in these groups who were CDR 0.5, male, or white is also shown. There were no significant differences between the groups on these factors.

Results

The mean scores achieved on each cognitive test at each annual evaluation are presented in Table 2. Performance declined significantly on all measures except the Boston Naming Test over the four year duration of the study. The mean changes in score from baseline to each subsequent evaluation are shown for each measure in Table 3. As the Table shows, scores on the Boston Naming Test remained at the same level over the course of four years, Cancellation Test scores initially improved significantly and then declined by the fourth year, and all other measures declined steadily over time and reached a significant level of decline by the fourth year (i.e., a change score significantly different from zero). Other than the Boston Naming Test, scores declined from 1.4% (Free and Cued Selective Reminding

Test) to 16% (NYU Paragraph Recall) between the baseline and 48-month (Year 4) evaluations.

Memory and Orientation

The REML analysis examining the effects of age, APOE genotype and other demographic factors on change in mMMSE scores showed that there was a significant decline over the course of the study [$F(1,1111)=26.81$; $p<.0001$], older performed worse than younger participants [$F(1,277)=8.31$; $p=.0042$], less educated performed worse than more educated participants [$F(1,277)=33.93$; $p<.0001$], and non-white performed worse than white participants [$F(1,277)=8.41$; $p=.004$]. Participants with higher CDR-SB scores (scores increase with worse function) performed worse than those with lower scores [$F(1,277)=14.58$; $p=.0002$]. Rate of decline on the mMMSE was influenced by age and APOE genotype [$F(1,1111)=15.24$; $p=.0001$] such that 4+ declined faster than 4- and this discrepancy grew with increasing age (even though 4- baseline-to-48 months decline was significantly greater than zero [$t(211)=2.39$; $p=.018$]). This three-way interaction was apparent even after controlling for education, gender, ethnicity and CDR-SB scores. To illustrate this interaction effect, the performance of APOE 4+ and 4- participants on the mMMSE across the four years of the study are compared separately for young-old and old-old participants (determined by a median split of age) in Figure 1A.

There was also a significant decline on the FCSRT over the course of the study [$F(1,1104)=23.83$; $p<.0001$], older performed worse than younger participants [$F(1,277)=7.58$; $p=.0063$], and 4+ performed worse than 4- participants [$F(1,277)=6.13$; $p=.014$]. Participants with higher CDR-SB scores performed worse than those with lower scores [$F(1,277)=7.03$; $p=.0085$]. After controlling for education, gender, ethnicity and CDR-SB scores, a significant three-way interaction effect showed that rate of decline on the FCSRT was influenced by age and APOE genotype [$F(1,1104)=30.83$; $p<.0001$] such that 4+ declined faster than 4- and this discrepancy grew with increasing age (see Figure 1B) (even though 4- baseline-to-48 months decline was significantly greater than zero [$t(209)=2.89$; $p=.004$]).

There was a significant interaction between APOE genotype and time on the NYU Paragraph Recall Test [$F(1,1113)=7.83$; $p=.0052$] indicative of more rapid decline on this measure in 4+ participants than in 4- participants regardless of age (see Figure 1C). There was also a general effect of age with older participants performing worse than younger participants [$F(1,278)=12.19$; $p=.0006$]. Less educated performed worse than more educated participants [$F(1,278)=36.72$; $p<.0001$], and non-white performed worse than white participants [$F(1,278)=7.71$; $p=.006$]. Participants with higher CDR-SB scores performed worse on the NYU Paragraph Recall Test than those with lower scores [$F(1,278)=8.64$; $p=.0036$]. The three-way age X genotype X time interaction effect was not significant [$F(1,1112)=2.20$; $p=.14$].

Executive Functions

Scores on the Digit Symbol Substitution Test declined significantly over the course of the study [$F(1,1105)=24.01$; $p<.0001$], older participants performed worse than younger participants [$F(1,276)=11.34$; $p=.0009$], less educated performed worse than more educated participants [$F(1,276)=31.48$; $p<.0001$], male participants performed worse than females [$F(1,276)=9.14$; $p=.0027$], and non-white performed worse than white participants [$F(1,276)=10.65$; $p=.0012$]. Participants with higher CDR-SB scores performed worse than those with lower scores [$F(1,276)=7.09$; $p=.008$]. After controlling for education, gender, ethnicity and CDR-SB scores, rate of decline on the Digit Symbol Substitution Test was influenced by age and APOE genotype [three-way interaction: $F(1,1105)=5.60$; $p=.018$]

such that 4+ declined faster than 4- and this discrepancy grew with increasing age (see Figure 2A) (even though 4- baseline-to-48 months decline was significantly greater than zero [$t(210)=2.98$; $p=.003$]).

There was a significant decline on Part A of the Trail-Making Test over the course of the study [$F(1,1108)=4.34$; $p=.037$], older performed worse than younger participants [$F(1,277)=19.33$; $p<.0001$], less educated participants performed worse than more educated participants [$F(1,277)=12.71$; $p=.0004$], and non-white performed worse than white participants [$F(1,277)=12.06$; $p=.0006$]. Participants with higher CDR-SB scores performed worse on this test than those with lower scores [$F(1,277)=5.21$; $p=.023$] (see Figure 2B). There were no interactions between time and age or APOE genotype on this measure [three-way age X genotype X time interaction effect: $F(1108)<1$; $p=.98$].

There was a significant interaction between APOE genotype and time on Part B of the Trail-Making Test [$F(1,1109)=7.15$; $p=.008$] indicative of more rapid decline on this measure in 4+ participants than in 4- participants regardless of age. There was also a general effect of age with older participants performing worse than younger participants [$F(1,278)=14.83$; $p=.0001$]. In addition, less educated performed worse than more educated participants [$F(1,278)=42.90$; $p<.0001$], non-white performed worse than white participants [$F(1,278)=7.16$; $p=.0079$], and participants with higher CDR-SB scores performed worse than those with lower scores [$F(1,277)=13.43$; $p=.0003$] (see Figure 2C). The three-way age X genotype X time interaction effect was not significant [$F(1108)=1.72$; $p=.19$].

Trail Making Test B-A difference scores declined significantly over the course of the study [$F(1,854)=5.96$; $p<.01$], older participants performed worse than younger participants [$F(1,213)=11.78$; $p=.001$], less educated performed worse than more educated participants [$F(1,213)=9.35$; $p<.001$], and non-white performed worse than white participants [$F(1,213)=14.16$; $p=.001$]. Older participants decline more rapidly than younger participants [$F(1,213)=11.78$; $p=.001$] on this measure. After controlling for education, gender, ethnicity and CDR-SB scores, rate of decline on the Trail Making Test difference score was influenced by age and APOE genotype [three-way interaction: $F(1,854)=3.82$; $p=.05$] such that 4+ declined faster than 4- and this discrepancy grew with increasing age (4- baseline-to-48 months decline was not significantly greater than zero [$t(208)=0.98$; $p=.33$]).

Language

Although there was no overall decline on the BNT, older participants performed worse than younger participants [$F(1,277)=6.90$; $p=.009$], less educated participants performed poorer than those with more education [$F(1,277)=53.75$; $p<.0001$], and non-white performed worse than white participants [$F(1,277)=23.90$; $p<.0001$]. Participants with higher CDR-SB scores performed worse than those with lower scores [$F(1,277)=10.26$; $p=.0015$]. Rate of change on the BNT was influenced by age and APOE genotype [three-way interaction: $F(1,1112)=6.72$; $p<.01$] (while controlling for education, gender, ethnicity and CDR-SB scores) such that 4+ changed faster than 4- and this discrepancy grew with increasing age (see Figure 3A) (4- baseline-to-48 months decline was not significantly greater than zero [$t(212)=0.51$; $p=.61$]).

There was a significant decline on the Category Fluency Test over the course of the study [$F(1,1110)=12.47$; $p=.0004$], older performed worse than younger participants [$F(1,277)=9.78$; $p=.0019$], less educated participants performed worse than more educated participants [$F(1,277)=40.49$; $p<.0001$], and participants with higher CDR-SB scores performed worse than those with lower scores [$F(1,277)=11.60$; $p=.0008$] (see Figure 3B). There were no interactions between time and age or APOE genotype on this measure [three-way age X genotype X time interaction effect: $F(1110)=2.15$; $p=.14$].

Visual Attention

There was significant decline over the course of the study on the Cancellation Test [$F(1,276)=51.64$; $p<.0001$], but no interactions between time and age or APOE genotype [three-way age X genotype X time interaction effect: $F(1106)=2.39$; $p=.12$]. Older participants performed worse than younger participants [$F(1,276)=9.72$; $p=.002$] and less educated performed worse than more educated participants [$F(1,276)=9.27$; $p=.0028$] on this test (see Figure 3C).

Discussion

The present results show that under procedures appropriate for an AD primary prevention trial non-demented older adults' performance declined over four years on commonly used tests of episodic memory, executive function, language, and attention. Four-year decline was evident on all cognitive test measures examined, with the exception of a test of confrontation naming, and ranged from -1.4% to -16.0% of baseline score. The largest overall declines as a percentage of baseline score occurred on the NYU paragraph delayed recall test (-16.0%), Trail-Making test B-A difference score (-15.9%), Part B of the Trail-Making test (-10.9%), and the Cancellation test (-9.9%). It should be noted, however, that performance on the NYU paragraph delayed recall and the Cancellation tests did not decline in a steady fashion, but fluctuated over the four-year observation period. Scores on the Cancellation test actually improved over baseline levels at the 12-month, 24-month, and 36-month evaluations, perhaps due to a practice effect, and declined only on the 48-month evaluation. NYU paragraph delayed recall showed its greatest decline relative to baseline at the 12-month evaluation (i.e., -2.15 point decline), with less decline evident at subsequent evaluations (e.g., a -1.23 point decline after 48 months), perhaps due to slight variations in the multiple versions that were used. In contrast, a steady, though not necessarily large, decline was observed for several measures including the mMMSE (approximately -0.35 points per year), the FCSRT (approximately -0.17 points per year), the Digit Symbol Substitution Test (approximately -0.73 points per year), and Part B of the Trail-Making Test (approximately 3.43 seconds increase per year). The magnitude of the observed decline on many of these measures was similar to that reported in previous studies (e.g., Yaffe et al., 1997).

Consistent with previous studies that examined population-based or convenience samples of older adults (Jonkers et al., 1998; Mayeux et al., 2001; O'Hara et al., 1998; Small et al., 1998; Wilson et al., 2002; Yaffe et al., 1997), the rate of decline on most of the neuropsychological tests was influenced by APOE genotype. Faster decline was evident in $4+$ participants than $4-$ participants on measures of memory (mMMSE, FCSRT, NYU Paragraph Recall Test), executive functions (Digit Symbol Substitution Test, Trail-Making Test, Part B and B-A difference score) and language (Boston Naming Test). This effect remained apparent after controlling for years of education, gender, ethnicity, and initial level of functional and cognitive ability assessed by the CDR-SB measure. The magnitude of the impact of APOE $4+$ genotype on rate of cognitive decline was not particularly large and may not have clinically important consequences, but could influence the ability to detect a relatively subtle treatment effect in a clinical trial. Although the present sample was limited to those at least 74 years of age, previous studies have shown that the effect of APOE on rate of cognitive decline may begin at a much younger age (Caselli et al., 2004; 2009).

The impact of APOE genotype on rate of decline was modulated by age for five of the cognitive measures: mMMSE, FCSRT, Digit Symbol Substitution Test, the derived difference score between time to complete Part B and time to complete Part A of the Trail Making Test, and the Boston Naming Test (although there was no overall decline on this latter test). In each case, $4+$ participants declined faster than $4-$ participants, and this discrepancy increased with increasing age. The interaction between age and APOE genotype

on cognitive decline was not specifically examined in many previous studies, but there is some evidence consistent with the present findings. A population-based study by Nilsson et al. (2006) showed that healthy adults between 70 and 85 years of age experienced more rapid decline over five years on tests of verbal recall if they carried at least one APOE ϵ 4 allele than if they did not. Caselli et al. (2009) found that the effect of aging on decline in memory performance on the Rey Auditory Verbal Learning Test was greater in ϵ 4+ than in ϵ 4- older adults. A similar interaction was not seen on tests of verbal fluency, visuo-perceptual abilities, or simple mental status.

Faster cognitive decline in elderly ϵ 4+ participants than in ϵ 4- participants, particularly in those who are oldest, may reflect an over-representation of AD in the former group. This possibility is consistent with previous findings that age and APOE ϵ 4 genotype are two of the strongest risk factors associated with sporadic (i.e. non-familial) AD (Mayeux & Stern, 2012). Also consistent with this possibility, differences between the ϵ 4+ and ϵ 4- participants in the present study were evident on measures that are most sensitive to cognitive changes associated with very early (e.g., Salmon et al., 2002) or preclinical (Mickes et al., 2009) AD. In particular, rate of decline on measures of orientation and memory, the mMMSE and the FCSRT, were most sensitive to the interaction between age and APOE genotype, consistent with evidence that the earliest cognitive change to occur in AD is a decline in episodic memory (e.g., Petersen et al., 1999). It is somewhat surprising that a similar interaction was not seen on another episodic memory test, the NYU Paragraph Recall Test, but this may reflect less sensitivity to early AD for this test compared to word-list learning tests, such as the FCSRT, because memory for prose passages provides semantic structure that can aid in encoding and retrieval processes. More accurate differentiation between early AD and normal aging with list-learning tests than with tests of memory for prose passages has been previously reported (Salmon et al., 2002).

To the extent that faster cognitive decline in elderly ϵ 4+ participants than ϵ 4- participants reflects an over-representation of AD in the elderly ϵ 4+ group, the present results suggest that APOE genotype and age should be considered in the design of AD primary prevention treatment trials. The results of such clinical trials could be distorted if participants with an APOE ϵ 4 allele are over-represented in an unbalanced way in either the placebo or treatment arm of a trial. Thus, APOE genotype should be stratified across placebo and treatment groups to insure that any group differences in onset of dementia are related to treatment effects. The results also suggest that AD primary prevention treatment trials could be enriched with elderly APOE ϵ 4 carriers (with stratification in the treatment and placebo groups) in order to enhance the possibility that sufficient numbers of non-demented participants would develop AD dementia within a reasonable trial timeframe to provide statistical power to detect a group difference. It should be noted, however, that vascular contributions to faster cognitive decline related to older age and APOE ϵ 4 genotype in the present study cannot be ruled out since vascular risk factors were not examined (although individuals with significant vascular disease were excluded). Therefore, a trial strategy that initially enriches with elderly APOE ϵ 4 carriers and then screens for the presence of preclinical biomarkers of AD pathology (e.g., amyloid imaging, cerebrospinal fluid levels of tau and β -amyloid proteins) might be most efficient.

Because initial level of cognitive and functional performance was accounted for in the analytic models by including the baseline CDR-SB score, it is unlikely that the interaction between age and APOE genotype on cognitive decline was driven by those participants with CDR scores of 0.5. To insure that this was the case, we repeated the REML analyses for all cognitive measures while limiting the sample to participants with CDR scores of 0 ($n=222$). The same pattern of results was obtained with significant age X APOE genotype X time interaction effects for the mMMSE [$F(1,856)=4.34$; $p=.038$] and the FCSRT

[$F(1,851)=6.10$; $p=.014$]. In each case, 4+ participants declined faster than 4- participants and this discrepancy grew with increasing age (see Supplementary Table 1.). The age X APOE genotype X time interaction effect was no longer observed for the Digit Symbol Substitution Test, but was apparent for Part A of the Trail-Making Test [$F(1,854)=4.07$; $p=.043$] with a faster decline in 4+ than 4- participants that grew larger with increasing age. As in the previous analyses, no other significant age X APOE genotype X time interaction effects were observed. These findings support the conclusions drawn from the larger sample.

Demographic factors influenced performance on a number of the cognitive tests. Participants with low education performed worse than those with higher education on the mMMSE, NYU paragraph delayed recall, Digit Symbol Substitution Test, BNT, Category Fluency Test, and the Cancellation test. White participants performed better than non-white participants on the mMMSE, NYU paragraph delayed recall, Digit Symbol Substitution Test, BNT, and Parts A and B of the Trail-Making Test. Participants with higher scores on the CDR-SB measure performed worse than those with lower scores on all of the cognitive measures except the Cancellation Test. The impact of these demographic factors on cognitive test performance in the general population has been reported previously (for reviews, see Heaton et al., 1991; Walker et al., 2009). The present results indicate that they can influence performance even in a carefully-screened sample that meets inclusion and exclusion criteria for an AD primary prevention trial and should be controlled in such trials.

Several limitations of the present study should be noted. First, there was a relatively high drop-out rate over the course of the study which reduced the sample size and the power to detect potential effects of APOE genotype and age on rate of decline on some measures. Data from participants who dropped out were not included in analyses because they were older, less educated and initially less functionally intact than those who completed the study. The loss of these poorer functioning subjects may lead to an underestimation of rate of decline on the cognitive tests. It is unlikely, however, that their loss alters the interpretation of the interaction between age and APOE genotype on rate of decline since there were similar percentages of 4+ and 4- non-completers in both Young-Old and Old-Old age groups. While the reduced sample size precludes strong claims about non-significant results, it does not diminish claims based on statistically significant results. However, the high drop-out rate could impact the generalizability of the findings to true clinical treatment trials in which drop-out rates may be lower because of the motivating effects of a potential treatment (it should be noted, however, that the demographic characteristics of the sample of those who completed the current study are similar to those in previously reported primary prevention trials; e.g., Snitz et al., 2009). Furthermore, an actual randomized treatment trial would require an intention-to-treat analysis with imputation of the outcome measure using last-observation-carried-forward or multiple imputation methods. That is less important for the explanatory analysis carried out in the present study. Second, the sample was selected to simulate an AD primary prevention treatment trial and may not be representative of the elderly population at large. Inclusion and exclusion criteria were strict with regard to health status and the availability of a knowledgeable study partner who could provide information about the participant's activities of daily living. Therefore, extrapolation of the present findings to the general elderly population should be done with caution. Third, the study was limited to individuals at least 74 years old and this precluded an examination of possible differences in rate of cognitive decline in 4+ and 4- participants that might emerge at a younger age. As previously mentioned, Caselli et al (2009) found that differences in rate of memory decline in 4+ and 4- participants began before the age of 60. It should be noted, however, that little difference in rate of decline was observed in the younger (i.e. 74-80 years of age) 4+ and 4- participants in the present study (see Figures 1-3), perhaps because of the stringent inclusion and exclusion criteria, or the inclusion of individuals with very mild cognitive or functional problems (i.e., $CDR = 0.5$) in both genotype groups.

Finally, the present study addressed the specific hypothesis that age and APOE 4+ genotype, two factors strongly associated with AD, would influence rate of cognitive decline. The possibility that family history of dementia beyond that accounted for by APOE genotype would be associated with rate of decline, or that demographic factors other than age might interact with APOE genotype to influence rate of decline, was not examined.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was supported by NIH award AG10483 for the Alzheimer's Disease Cooperative Study. The authors report no actual or potential conflicts of interest associated with this study. The study was approved by the institutional review boards at UCSD and each of the participating ADCS sites. Informed consent was obtained from all study participants after the procedures of the study had been fully explained.

References

- Aisen PS, Andrieu S, Sampario C, Carillo M, Khachaturian ZS, Dubois B, Feldman HH, Petersen RC, Siemers E, Doody RS, Hendrix SB, Grundman M, Schneider LS, Schindler RJ, Salmon E, Potter WZ, Thomas RG, Salmon D, Donohue M, Bedner MM, Touchon J, Vellas B. Report of the task force on designing clinical trials in early (pre-dementia) Alzheimer's disease. *Neurology*. 2011; 76:280–286. [PubMed: 21178097]
- Army Individual Test Battery: Manual of Directions and Scoring. War Department's Adjutant General's Office; Washington, DC: 1944.
- Bondi MW, Salmon DP, Galasko D, Thomas RG, Thal LJ. Neuropsychological function and apolipoprotein E genotype in the preclinical detection of Alzheimer's disease. *Psychology and Aging*. 1999; 14:295–303. [PubMed: 10403716]
- Borkowski JG, Benton A, Spreen O. Word fluency and brain damage. *Neuropsychologia*. 1967; 5:135–140.
- Bunce D, Fratiglioni L, Small BJ, Winblad B, Backman L. APOE and cognitive decline in preclinical Alzheimer's disease and non-demented aging. *Neurology*. 2004; 63:816–821. [PubMed: 15365129]
- Butters N, Granholm E, Salmon DP, Grant I, Wolfe J. Episodic and semantic memory: A comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology*. 1987; 9:479–497. [PubMed: 2959682]
- Caselli RJ, Dueck AC, Locke DEC, Hoffman-Snyder CR, Woodruff BK, Rapcsak SZ, Reiman EM. Longitudinal modeling of frontal cognition in APOE e4 homozygotes, heterozygotes, and noncarriers. *Neurology*. 2011; 76:1383–1388. [PubMed: 21502596]
- Caselli RJ, Dueck AC, Osborne D, Sabbagh MN, Conner DJ, Ahern GL, Baxter LC, Rapcsak SZ, Shi J, Woodruff BK, Locke DE, Snyder CH, Alexander GE, Rademakers R, Reiman EM. Longitudinal modeling of age-related memory decline and the APOE 4 effect. *New England Journal of Medicine*. 2009; 361:255–263. [PubMed: 19605830]
- Caselli RJ, Reiman EM, Osborne D, Hentz JG, Baxter LC, Hernandez JL, Alexander GG. Longitudinal changes in cognition and behavior in asymptomatic carriers of the APOE 4 allele. *Neurology*. 2004; 62:1990–1995. [PubMed: 15184602]
- Ferris SH, Aisen PS, Cummings JL, Galasko D, Salmon DP, Schneider L, Sano M, Whitehouse PJ, Edland S, Thal LJ. for the Alzheimer's Disease Cooperative Study Group. ADCS Prevention Instrument Project: Overview and initial results. *Alzheimer's Disease and Associated Disorders*. 2006; 20(Suppl 3):S109–S123.
- Grober E, Buschke H, Crystal H, Bang S, Dresner R. Screening for dementia by memory testing. *Neurology*. 1988; 38:900–903. [PubMed: 3368071]
- Grober E, Kawas C. Learning and retention in preclinical and early Alzheimer's disease. *Psychology and Aging*. 1997; 12:183–188. [PubMed: 9100279]

- Heaton, RK.; Miller, SW.; Taylor, MJ.; Grant, I. Revised comprehensive norms for an expanded Halstead-Reitan battery. Psychological Assessment Resources; Florida: 1991.
- Hughes CP, Berg L, Danzinger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *British Journal of Psychiatry*. 1982; 140:566–572. [PubMed: 7104545]
- Jonker C, Schmand B, Lindeboom J, Havekes LM, Launer LJ. Association between apolipoprotein E epsilon4 and the rate of cognitive decline in community-dwelling elderly individuals with and without dementia. *Archives of Neurology*. 1998; 55:1065–1069. [PubMed: 9708956]
- Kaplan, E.; Goodglass, H.; Weintraub, S. Boston Naming Test. Lea & Febiger; Philadelphia, PA: 1983.
- Kluger A, Ferris SH, Golomb J, Mittelman MS, Reisberg B. Neuropsychological prediction of decline to dementia in nondemented elderly. *Journal of Geriatric Psychiatry and Neurology*, 1999. 1999; 12:168–179.
- Laird NM, Ware JH. Random-Effects Models for Longitudinal Data. *Biometrics*. 1982; 38:963–974. [PubMed: 7168798]
- Mayeux R, Small SA, Tang MX, Tycko B, Stern Y. Memory performance in healthy elderly without Alzheimer's disease: effects of time and apolipoprotein-E. *Neurobiology of Aging*. 2001; 22:683–689. [PubMed: 11445269]
- Mayeux, R.; Stern, Y. Epidemiology of Alzheimer's disease. In: Selkoe, DJ.; Mandelkow, E.; Holtzman, DM., editors. *The Biology of Alzheimer Disease*. Cold Spring Harbor, NY: Cold Spring Harbor Press; 2012. p. 115-132.
- Mickes L, Wixted JT, Fennema-Notestine C, Galasko D, Bondi MW, Thal LJ, Salmon DP. Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. *Neuropsychology*. 2007; 21:696–705. [PubMed: 17983283]
- Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, Sano M, Bieliauskas L, Geldmacher D, Clark C, Thal LJ. Development of cognitive instruments for use in clinical trials of antidementia drugs: Additions to the Alzheimer's Disease Assessment Scale (ADAS) that broaden its scope. *Alzheimer's Disease and Associated Disorders*. 1997; 11 (Suppl 2):S13–S21.
- Nilsson LG, Adolfsson R, Backman L, Cruts M, Nyberg L, Small BJ, van Broeckoven C. The influence of APOE status on episodic and semantic memory: Data from a population-based study. *Neuropsychology*. 2006; 20:645–657. [PubMed: 17100509]
- O'Hara R, Yesavage JA, Kraemer HC, Mauricio M, Friedman LF, Murphy GM. The APOE epsilon4 allele is associated with decline on delayed recall performance in community-dwelling older adults. *Journal of the American Geriatric Society*. 1998; 46:1493–1498.
- Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild 'cognitive impairment: clinical characterization and outcome. *Archives of Neurology*. 1999; 56:303–308. [PubMed: 10190820]
- Reitan R. Validity of the trail making test as an indicator of organic brain disease. *Perceptual and Psychomotor Skills*. 1958; 8:271–276.
- Salmon DP, Bondi MW. Neuropsychological assessment of dementia. *Annual Review of Psychology*. 2009; 60:257–282.
- Salmon DP, Thomas RG, Pay MM, Booth A, Hofstetter CR, Thal LJ, Katzman R. Alzheimer's disease can be accurately diagnosed in very mildly impaired individuals. *Neurology*. 2002; 59:1022–1028. [PubMed: 12370456]
- Small BJ, Basun H, Backman L. Three-year changes in cognitive performance as a function of apolipoprotein E genotype: evidence from very old adults without dementia. *Psychology and Aging*. 1998; 13:80–87. [PubMed: 9533192]
- Snitz BE, O'Meara ES, Carlson MC, Arnold AM, Ives DG, Rapp SR, Saxton J, Lopez OL, Dunn LO, Sink KM, DeKosky ST. for the Ginkgo Evaluation of Memory (GEM) study investigators. Ginkgo biloba for preventing cognitive decline in older adults. *Journal of the American Medical Association*. 2009; 302:2663–2670. [PubMed: 20040554]
- Teng EL, Chui HC. The Modified Mini Mental State (3MS) Examination. *Journal of Clinical Psychiatry*. 1987; 48:314–318. [PubMed: 3611032]

- Walker AJ, Batchelor J, Shores A. Effects of education, and cultural background on performance on WAIS-III, WAIS-R and WMS-R measures: systematic review. *Australian Psychologist*. 2009; 44:216–223.
- Wechsler, D. Wechsler Adult Intelligence Scale-Revised. Psychological Corp; New York, NY: 1981.
- Wilson RS, Schneider JA, Barnes LL, Beckett LA, Aggarwal NT, Cochran EJ, Berry-Kravis E, Bach J, Fox JH, Evans DA, Bennett DA. The apolipoprotein E e4 allele and decline in different cognitive systems during a 6-year period. *Archives of Neurology*. 2002; 59:1154–1160. [PubMed: 12117364]
- Yaffe K, Cauley J, Sands L, Browner W. Apolipoprotein E phenotype and cognitive decline in a prospective study of elderly community women. *Archives of Neurology*. 1997; 54:1110–1114. [PubMed: 9311354]

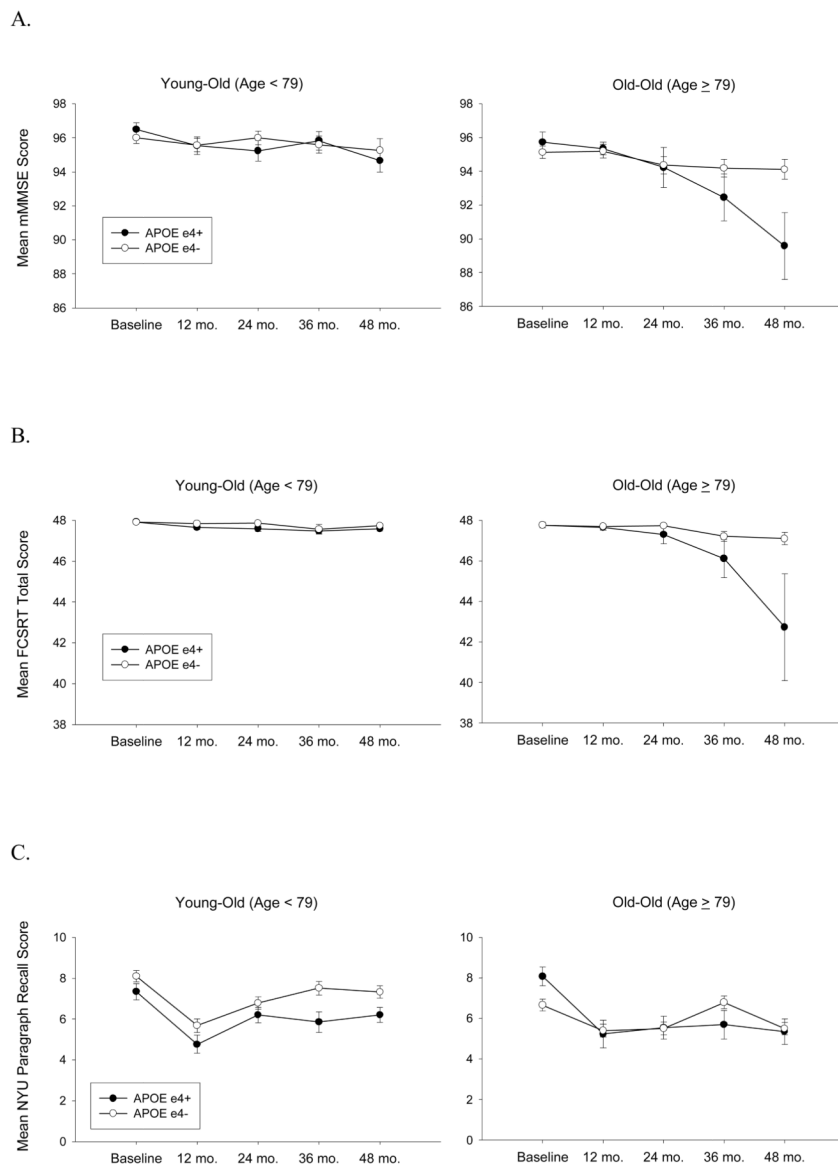


Figure 1. The mean scores achieved on tests of orientation and memory at each time point as a function of age (based on median split) and apolipoprotein E (APOE) genotype. Panel A: modified Mini-Mental State Exam (mMMSE); Panel B: Free and Cued Selective Reminding Test (FCSRT); Panel C: New York University (NYU) Paragraph Recall Test. N's = 115 Young-Old 4-, 98 Old-Old 4-, 46 Young-Old 4+, and 27 Old-Old 4+.

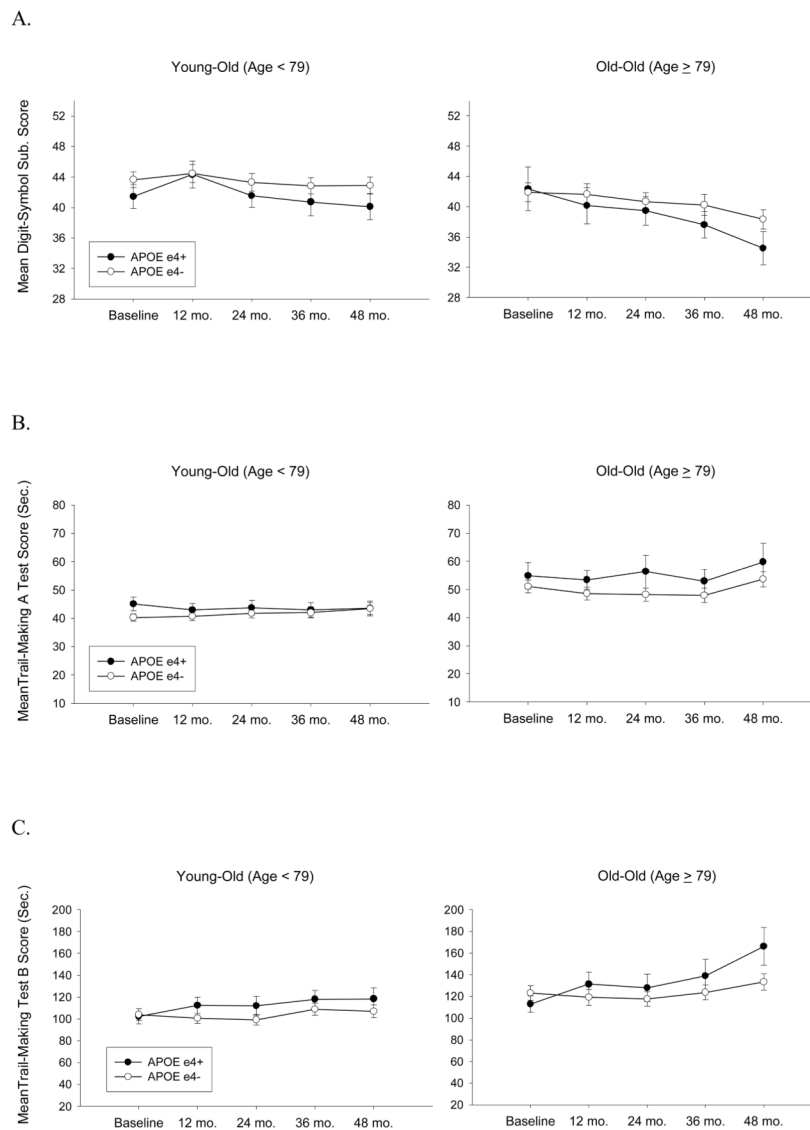


Figure 2. The mean scores achieved on measures of executive functions at each time point as a function of age (based on median split) and apolipoprotein E (APOE) genotype. Panel A: Digit-Symbol Substitution Test; Panel B: Trail-Making Test, Part A; Panel C: Trail-Making Test, Part B. N's = 115 Young-Old 4-, 98 Old-Old 4-, 46 Young-Old 4+, and 27 Old-Old 4+.

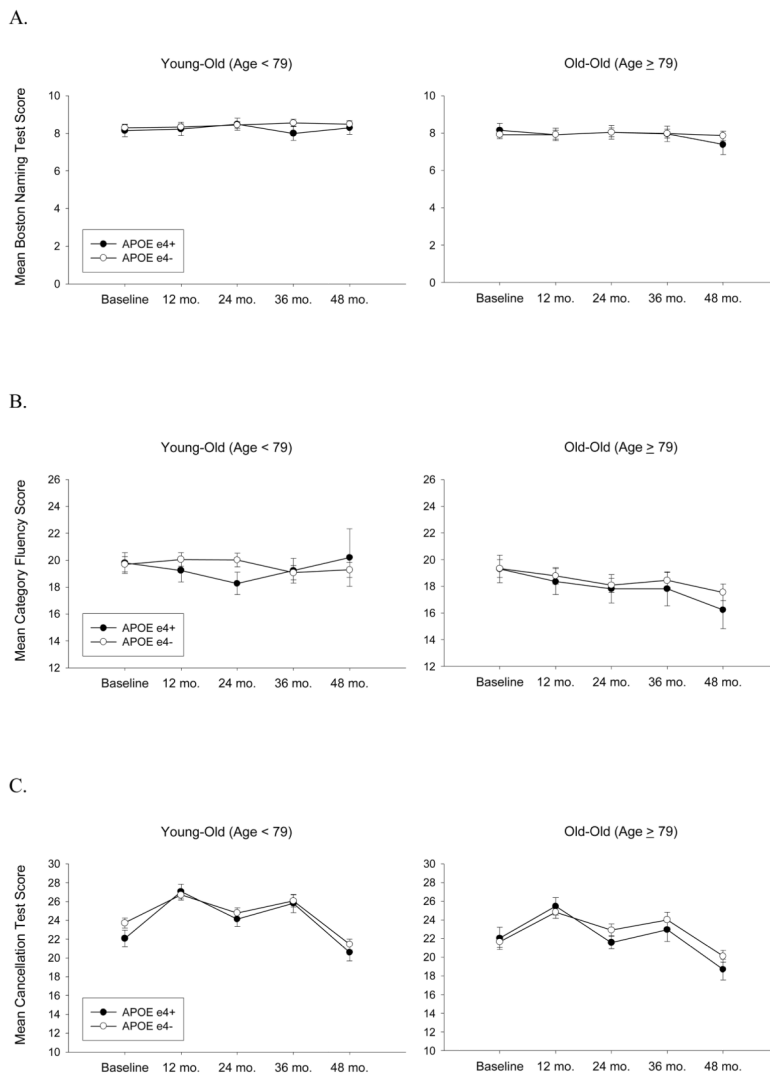


Figure 3. The mean scores achieved on measures of language and attention at each time point as a function of age (based on median split) and apolipoprotein E (APOE) genotype. Panel A: Boston Naming Test (10-item version); Panel B: Category Fluency Test (“Animals”); Panel C: Cancellation Test. N’s = 115 Young-Old 4–, 98 Old-Old 4–, 46 Young-Old 4+, and 27 Old-Old 4+.

Table 1

The mean age, years of education, modified Mini-Mental State Examination (mMMSE) scores and Clinical Dementia Rating (CDR) Sum of Boxes (SOB) scores for participants with zero (4-) or at least one (4+) apolipoprotein E (APOE) 4 allele who completed the study (standard deviations are shown in parentheses). The percentage of participants in these groups who were CDR 0.5, male, or white is also shown

	APOE 4+ (n=73)	APOE 4- (n=213)	
Age	78.19 (3.24)	78.86 (3.29)	t (284)=1.51; p=.13, <i>d</i> =.20
Education	15.81 (2.68)	15.15 (3.21)	t (284)=1.56; p=.12, <i>d</i> =.21
mMMSE	96.36 (2.68)	95.57 (3.67)	t (284)=1.68; p=.09, <i>d</i> =.23
% CDR 0.5	30.1%	19.7%	X ² =3.40; p=.07
CDR SOB	0.34 (0.50)	0.27 (0.44)	t (284)=1.02; p=.31, <i>d</i> =.15
% Male	39.7%	43.7%	X ² =0.34; p=.56
% White	76.7%	81.2%	X ² =1.74; p=.78

Table 2

The mean scores achieved by the 417 participants who completed the study on each cognitive measure at Baseline and subsequent evaluations. Standard deviations are shown in parentheses. Note that higher scores are indicative of worse performance on the Trail-Making Test A and B. The results of one-way, repeated-measures Analysis of Variance comparing scores on each measure across time are shown in the right-most column. Small differences in degrees of freedom across tests reflect occasional missing data due to participant refusal, lack of time to complete testing, or administrator error

	Baseline	12-month	24-month	36-month	48-month	
mMSE	95.84 (3.31)	95.48 (3.97)	95.36 (4.47)	94.20 (4.85)	94.44 (6.67)	F (1,393) = 18.60 p < .001; $\eta^2 = .045$
Free & Cued Selective Reminding	47.84 (0.48)	47.75 (0.73)	47.73 (0.86)	47.40 (2.16)	47.13 (3.99)	F (1,388) = 15.10 p < .001; $\eta^2 = .037$
NYU Paragraph Recall	7.56 (2.81)	5.38 (3.18)	6.19 (2.94)	6.87 (3.26)	6.35 (3.02)	F (1,395) = 6.92 p = .009; $\eta^2 = .017$
Trail Making A (seconds)	44.56 (17.76)	43.45 (16.98)	44.03 (19.11)	44.37 (20.94)	46.82 (23.08)	F (1,391) = 5.23 p = .023; $\eta^2 = .013$
Trail Making B(seconds)	107.17 (51.67)	107.71 (54.14)	106.68 (56.31)	114.68 (61.32)	119.65 (68.06)	F (1,388) = 25.19 p < .001; $\eta^2 = .061$
Trail Making: B-A Difference Score (seconds)	62.87 (44.58)	64.33 (45.60)	62.46 (46.29)	70.50 (51.36)	72.89 (55.53)	F (1,385) = 18.66 p < .001; $\eta^2 = .046$
Digit Symbol Substitution	43.29 (11.14)	43.48 (12.09)	42.93 (11.50)	42.02 (11.78)	40.92 (12.17)	F (1,390) = 29.79 p < .001; $\eta^2 = .071$
Cancellation Test	22.82 (5.92)	26.35 (5.87)	24.02 (5.86)	25.23 (7.06)	20.70 (6.24)	F (1,391) = 77.22 p < .001; $\eta^2 = .165$
Boston Naming Test	8.14 (1.99)	8.24 (1.94)	8.33 (1.99)	8.20 (2.07)	8.20 (2.16)	F (1,395) = 0.24 p = .624; $\eta^2 = .001$
Verbal Fluency: Animals	19.40 (5.61)	19.36 (5.29)	19.02 (5.22)	18.98 (5.75)	18.66 (7.06)	F (1,392) = 6.78 p = .010; $\eta^2 = .017$

Table 3

The mean change in score from Baseline to each annual evaluation is shown for the various cognitive measures for the 417 participants who completed the study

	12-month	24-month	36-month	48-month	% Change from Baseline at 48 mo.
mMMSE	-0.39 (3.38)	-0.53 (3.83)	-1.28 (8.20)	-1.42*** (5.93)	-1.5%
Free & Cued Selective Reminding	-0.10 (0.802)	-0.22 (1.80)	-0.45 (2.14)	-0.68*** (3.83)	-1.4%
NYU Paragraph Recall	-2.15 (3.20)	-1.32 (2.81)	-0.67 (3.38)	-1.23*** (2.97)	-16.0%
Trail Making A (seconds)	-0.70 (15.90)	-0.27 (16.41)	-0.01 (18.63)	2.40** (19.81)	-5.3%
Trail Making B (seconds)	1.13 (43.32)	-0.01 (42.91)	7.94 (45.20)	12.09*** (55.08)	-10.9%
Trail Making B-A Difference Score (seconds)	-2.08 (43.30)	0.17 (43.15)	-7.64 (46.02)	-10.28*** (52.72)	-15.9%
Digit Symbol Substitution	0.15 (9.08)	-0.27 (8.56)	-1.19 (8.86)	-2.30*** (9.56)	-5.3%
Cancellation Test	3.54 (5.15)	1.23 (5.25)	2.42 (5.52)	-2.24*** (5.32)	-9.9%
Boston Naming Test	0.10 (1.20)	0.20 (1.36)	0.04 (1.36)	0.04 (1.45)	+0.7%
Verbal Fluency: Animals	-0.05 (4.48)	-0.35 (4.63)	-0.44 (5.14)	-0.70* (6.43)	-3.7%

Standard deviations are shown in parentheses. Note that positive change scores are indicative of worse performance on the Trail-Making Test A and B. The percent (%) change in score from Baseline to 48 months is shown for each measure in the right-most column. Baseline-to-48 months change scores that are significantly different from zero (by student's t-test) are indicated by *at the p < .05 level, **at the p < .01 level, and ***at the p < .001 level.