Published in final edited form as:
Alzheimer Dis Assoc Disord. 2015 ; 29(1): 32–44. doi:10.1097/WAD.0000000000000042.

Normative Data for Eight Neuropsychological Tests in Older Blacks and Whites from the Atherosclerosis Risk in Communities (ARIC) Study

Andrea L.C. Schneider, PhD[1], **A. Richey Sharrett, MD DrPH**[1], **Rebecca F. Gottesman, MD PhD**[1,2], **Josef Coresh, MD PhD**[1], **Laura Coker, PhD**[3], **Lisa Wruck, PhD**[4], **Ola A. Selnes, PhD**[2], **Jennifer Deal, PhD**[1], **David Knopman, MD**[5], and **Thomas H. Mosley 6, PhD**[6]

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, 21205, USA

[2]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA

[3]Department of Social Sciences and Health Policy, Wake Forest University School of Medicine, Winston-Salem, North Carolina, 27106, USA

[4]Department of Biostatistics, University of North Carolina Gillings School of Global Public Health, Chapel Hill, North Carolina, 27514, USA

[5]Department of Neurology, Mayo Clinic, Rochester, Minnesota, 55905, USA

[6]Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi, 39216, USA

Abstract

Accurate assessment of cognitive impairment requires comparison of cognitive performance in individuals to performance in a comparable healthy normative population. Few prior studies have included a large number of black participants and few have excluded participants from the normative sample with subclinical/latent neurologic disease or dementia. This study provides age, race, and education specific normative data for 8 cognitive tests derived from 320 black and 392 white participants aged 61–82 years (mean 71 years) in the Atherosclerosis Risk in Communities (ARIC) Study without clinical or subclinical/latent neurological disease. Normative data are provided for the Delayed Word Recall Test, Logical Memory Parts I and II, the Word Fluency Test, Animal Naming, the Trail Making Test Parts A and B and the Digit Symbol Substitution Test. Age, race, and education specific mean and −1.5 standard deviation scores are given in tabular form and graphically, as well as regression-based equations to derive adjusted score cut-points. These robust normative data should enhance comparison across studies of cognitive aging, where these measures are widely used, and improve interpretation of performance on these tests for the diagnosis of cognitive impairment not only within the ARIC cohort, but also among older blacks and whites with similar demographics.

Corresponding Author: Thomas H. Mosley PhD, Department of Medicine (Geriatrics), University of Mississippi Medical Center, 2500 North State St., Jackson, MS 39216-4505, Phone: 601-984-2763, Fax: 601-815-3422, tmosley@umc.edu.

Keywords

Neuropsychological testing; Normative data; Cognitive performance; ARIC Study; Race

INTRODUCTION

The burden of dementia and mild cognitive impairment (MCI) among the elderly in the United States is high, affecting approximately 25% of adults aged 70 years or older [1, 2] and the number of persons with these conditions is expected to triple over the next 20 years along with the growing elderly population [3]. Neuropsychological testing is often part of the evaluation performed in the identification, diagnosis, and subtyping of dementia and MCI. Unbiased identification of cognitive impairment requires a person's cognitive performance to be compared to an appropriate normative sample derived from a comparable healthy population.

Level of performance on cognitive tests in persons without evidence of neurologic disease is related to several demographic factors, including age [4, 5], race [4–6], and education level [4–7]. Several studies have shown that blacks (and other ethnic minority groups) tend to score lower than whites of similar age on neuropsychological tests even at apparently equivalent education and socioeconomic levels, resulting in reduced specificity and misclassification of cognitive impairment [4–6]. For these reasons, normative data for older adults, and more recently for specific racial/ethnic populations (in particular blacks) have been published [8–14]. While the use of race-specific norms has been shown to substantially reduce misclassification of cognitive impairment to a level comparable to that observed in whites, a limitation of previous normative studies is the inclusion of a relatively small number of black participants with diverse ages and education levels [8–13]. Additionally, previous normative studies generally have not been "robust" and have not excluded participants with subclinical/latent neurologic disease/risk factors (e.g. lacunes, APOE genotype), or those who subsequently develop dementia after a short follow up[8–13]. Inclusion of individuals with these factors in the "conventional" normative population may lower cognitive test score means and decrease sensitivity to detect cognitive impairment. Therefore, robust normative data may increase sensitivity to detect cognitive impairment[15, 16], but it is important to note that this increased sensitivity may also result in an increase in false positive cases.

The Atherosclerosis Risk in Communities Neurocognitive Study (ARIC-NCS) was funded to determine the prevalence of MCI and dementia and to evaluate the association of risk factors, particularly midlife vascular risk factors and markers, with late life cognitive impairment in the biracial ARIC cohort. Given the sparsity of cognitive norms in blacks and key limitations of existing conventional norms noted above, there was a need to develop age, race, and education specific norms for a battery of cognitive measures to use in identifying cognitive impairment in ARIC participants. Taking advantage of the extensively characterized population-based ARIC cohort, a robust normative sample of whites and blacks was defined, comprised of individuals free of both clinical and subclinical/latent neurological disease.

In the current study, we present robust normative data for 8 neuropsychological tests measuring a range of cognitive functions. The measures were selected not only to provide consistency with longitudinal assessments within ARIC, but also to provide comparability with related studies of cognitive aging [17–19], where these tests have been widely used, including many from the core battery recommended in the Uniform Data Set by the National Institute on Aging (NIA) Alzheimer Disease Centers [20, 21]. Normative results are reported for: Delayed Word Recall Test (DWRT), Logical Memory Part I (LM I), Logical Memory Part II (LM II), Word Fluency Test (WFT),Animal Naming, Trail Making Test Part A (TMT-A), Trail Making Test Part B (TMT-B), and Digit Symbol Substitution Test (DSST).

METHODS

Study Population

The Atherosclerosis Risk in Communities (ARIC) study is an ongoing, community-based prospective cohort of 15,792 middle-aged adults from four U.S. communities: Washington County, Maryland; suburbs of Minneapolis, Minnesota; Forsyth County, North Carolina; and Jackson, Mississippi. ARIC participants were seen at four in-person visits which occurred approximately 3 years apart, from 1987–89 for visit 1, through 1996–98 for visit 4. A fifth visit was conducted in 2011–2013. A subset of ARIC participants (n=1,134) from the Forsyth County, North Carolina and Jackson, Mississippi field centers attended the ARIC Brain MRI visit (2004–2006) and underwent a more comprehensive cognitive assessment than was performed on the entire cohort at ARIC visits 2 and 4. Therefore, the ARIC Brain MRI visit participants serve as the source population for the present analysis.

Of the 1,134 participants who attended the ARIC Brain MRI visit, we excluded (1) 101 participants with clinical neurologic disease. Clinical neurologic disease was defined as: (a) stroke or transient ischemic attack prior to the Brain MRI exam, (b) prior diagnosis of multiple sclerosis, Parkinson's disease, brain tumor, (c) history of surgery or radiation to brain or skull, (d) prior diagnosis of dementia, (e) use of cholinomimetic medication. (2) 279 participants with subclinical neurologic disease or possible latent dementia. Subclinical neurologic disease or possible latent dementia was defined as: (a) white matter grade ≥6 [22], (b) two or more lacunar infarcts measuring >3 mm, (c) Mini-Mental Status Exam (MMSE) score <22 at the ARIC Brain MRI visit or at the ARIC Neurocognitive Study visit (2011–2013), (d) two APOE ε4 alleles, (e) self-report of often misplacing or losing items around the house, (f) self-report of often having trouble remembering conversations that occurred a few days earlier, (g) ICD-9 discharge code for dementia at any hospitalization (obtained from annual telephone contact with study participants and through active surveillance of all hospitalizations in the study communities) occurring after the Brain MRI visit through December 31, 2011 (ICD-9 codes: 290.0, 290.1, 290.2, 290.3, 290.9, 294.1, 294.2, 294.8, 294.9, 331.1, 331.2, 331.8, 331.9). And (3) 1 participant with missing education data, and 41 participants with missing cognitive test score data for DWRT, LM I, LM II, WFT, Animal Naming, TMT-A, and DSST. After exclusions, 712 participants remained for the analysis (Figure 1). All included participants were required to have complete data for model covariates (age, race, and education). Our normative data for TMT-B is derived from a subset of participants (n=651) who were not missing data on this test.

Measures of Cognitive Function

Trained examiners administered the cognitive tests in a standardized order during one session in a quiet room. Examiner performance was monitored by audio tape recording. Recordings were reviewed locally and shared across centers to ensure consistency with testing procedures and standardization across study sites.

DWRT [23] is a test of verbal learning and recent memory. In this test, participants were given 10 common nouns that they were asked to learn by using each word in a sentence. Two exposures to the words were given. After a five-minute delay, participants were given 60 seconds to recall the words. The score for the DWRT is the number of words correctly recalled.

Logical memory test, from the Wechsler Memory Scale-Revised (WMS-R) [24] is a test of immediate (LM I) and delayed (LM II) memory. In this test, two short stories (story A and story B) are presented, each containing a total of 25 pieces of information. Immediately after each story is presented, free recall of the story is elicited. The score for LM I is derived as the total number of story elements recalled. After a 30-minute delay, free recall of each story is elicited. The score for LM II is derived as the total number of story elements recalled. We present normative data for LM I and LM II as well as data for LM I Story A and LM II Story A to facilitate comparison with the Uniform Data Set from the NIA Alzheimer Disease Centers [20].

WFT, also known as the Controlled Oral Word Association Test (COWA) [25, 26], is a test of executive function and language. Participants were given 60 seconds to generate as many words as possible for the letters F, A and S (60 seconds for each letter), avoiding proper nouns. The word fluency score is the total number of acceptable words generated for the three letters.

Animal Naming [27] is a test of semantic category fluency in which the participant is asked to spontaneously generate words from a specific category (in this test, animals). Participants could name multiple words in the same subcategory (e.g., dog, poodle, dalmation). The score is the total number of animals generated within 60 seconds.

TMT-A [28] is primarily a test of processing speed in which participants are asked to draw a line connecting circles numbered 1 to 25 that are randomly distributed on the page as fast as possible. The score is the time (seconds) for completion of this task, with a maximum allotted time of 240 seconds.

TMT-B [28] is a test of executive function and processing speed in which participants are asked to draw a line connecting the numbers 1 to 13 and the letters A to L that are randomly distributed on the page, alternating between numbers and letters. The score is the time (seconds) for completion of this task, with a maximum allotted time of 240 seconds.

DSST, of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) [29], is a test of processing speed and new learning, where participants were asked to translate numbers to symbols using a key. The score is the total number of numbers correctly translated to symbols within 90-seconds and the range of possible scores is 0 to 93.

Statistical Analysis

Means (standard deviations) and percentages for characteristics of participants included and excluded from our analytic population stratified by race were calculated to facilitate comparison with other populations. We also report how many participants in each group are missing data on each covariate. T-test and chi-squared statistics were used to compare covariates across groups. Variables included are demographic factors (age, gender, education [< high school; high school or vocational school; college, graduate or professional school], income [<$35,000/year; ≥$35,000/year; not reported]), lifestyle factors (cigarette smoking, alcohol consumption), cardiovascular and genetic risk factors (diabetes, hypertension, APOE genotype), cognitive test scores and the 11-Item Center for Epidemiological Studies-Depression Scale score.

To create cognitive test normative data, we modeled the association between age and cognitive test score and included terms for education, race and the interaction of education X race. A linear model was determined to have the best fit using the criteria of standardized residuals, root-mean squared error, and graphical display. Using this model we report the mean and −1.5 SD (defined as −1.5 times the root mean squared error from each model) for each age, race, and education group. We chose to report the −1.5 SD normative score because this is a common cut-point for defining cognitive impairment in MCI and dementia [30, 31]. Age categories were defined as follows: 65- < 70, 70- < 75, and 75- < 80. The midpoint age was used in the model to define the mean and −1.5 SD score for each 5-year age category. We performed a sensitivity analysis to compare the −1.5 SD scores derived from our model to scores calculated using stratified data (stratified by age, race, and education). In this sensitivity analysis, we calculated the stratified −1.5 SD scores for each age category using all individuals within 5 years of the midpoint age, but the scores calculated are applied only to individuals within 2.5 years of the midpoint age. For example, stratified −1.5 SD scores for midpoint age 67.5 are compared to scores from the model for age category 65- < 70 and are derived from all individuals between the ages of 62.5 and 72.5 who are within the same race and education category.

In order to assess the associations of age (per 10 years) and education (comparing < high school to > high school [college, graduate, or professional school]) with cognitive test scores we calculated standardized z-scores for each test by subtracting the test mean and dividing by the standard deviation. Means and standard deviations for each test were calculated separately by race. As a sensitivity analysis, we also defined standardized z-scores for each test using the overall (combined black and white) means and standard deviations.

All analyses were performed using Stata Version 13 (StataCorp, College Station, Texas).

RESULTS

Table 1 shows the characteristics of the 712 participants included and the 422 participants excluded from the normative population, stratified by race. Comparing the 320 included blacks to the 274 excluded blacks, those who were included were younger (71.0 years versus 73.1 years, $p < 0.001$), more likely to have college, graduate, or professional school education (49.7% versus 32.1%, p=0.002), and, as expected, performed better on all 8

cognitive tests (p<0.001 for all tests). Whites who were included (n=392) and excluded (n=148) did not differ by education category (p=0.639), but those who were included were younger (72.4 years vs. 73.5 years, p=0.009), and performed better on all tests (p<0.010) except the WFT (35.2 versus 34.1 words, p=0.322). Comparing included blacks to included whites, blacks were younger (71.0 years versus 72.4 years, p<0.001), were more likely to have < high school education (26.6% versus 8.4%, p<0.001), and had lower scores on all 8 cognitive tests (p<0.001 for all tests).

Table 2 shows the regression coefficients and root mean squared errors from the linear regression models for each cognitive test. This table can be used to generate age, race, and education specific normative scores. For example, to calculate the −1.5 SD normative score on the DWRT for a black individual who is 72 years old with < high school education: −1.5 SD DWRT score = [6.785+(−0.068*(72-71))+(−0.636)+(−0.428)+(−0.024)]–(1.5*1.458) = 3.442 words. Tables 3 and 4 show the mean and −1.5 SD scores by age and education category for blacks (Table 3)and whites (Table 4) derived from the linear regression model. The −1.5 SD score for each cognitive test by age, race, and education category are depicted graphically in Figure 2. In sensitivity analysis, we compared −1.5 SD scores derived from the model to −1.5 SD scores calculated in stratified analyses (Appendix Figure 1). Model derived and stratified −1.5 SD scores were largely similar in both whites and blacks for DWRT, TMT-A, TMT-B, and DSST. However, for LM I, LM II, WFT, and Animal Naming, −1.5 SD scores from stratified analyses for blacks tended to be higher than those derived from the model, especially for lower levels of education. Alternatively, for whites, the −1.5 SD scores from stratified analyses tended to be lower than those derived from the model, especially for higher levels of education.

The association of age (per 10 years) with worse performance on cognitive tests tended to be stronger for tests of executive function (TMT-B: 0.51 standard deviations worse [95% CI: 0.35, 0.68], DSST: 0.48 standard deviations worse [95% CI: 0.33, 0.64]) compared to tests of verbal fluency (WFT: 0.19 standard deviations worse [95% CI: 0.03, 0.35], Animal Naming: 0.36 standard deviations worse [95% CI: 0.20, 0.53]). The association of age (per 10 years) with scores on tests of memory tended to fall in between the association of age (per 10 years) with scores on tests of executive function and verbal fluency. The association of education (comparing < high school to > high school [college, graduate, or professional school]) tended to be weaker for tests of memory in both blacks and whites compared to tests of executive function or verbal fluency (Appendix Table 1). In sensitivity analysis, there were no substantial differences in associations of age or education with cognitive test scores using the z-score derived from the overall population compared to the race-specific z-scores.

DISCUSSION

In this study, we provide age, race, and education specific robust normative data for 8 cognitive tests derived from an extensively well-characterized population of older black and white individuals free of both clinical and subclinical/latent neurological disease and dementia. Additionally, we show that although age, race, and education are each associated with cognitive performance, associations with age tended to be greatest for tests of executive

function, while associations with education tended to be greatest for measures of executive function and word fluency.

Formal comparison of our norms with prior studies is difficult because of differences in study design. Many prior normative studies [8–13] reported "conventional" norms based on a population which excluded persons with evident clinical neurologic disease and dementia, while our study reports "robust" norms based on a population which excluded participants with both clinical and subclinical/latent neurological disease and dementia. Robust normative data have usually been shown to be more appropriate for studies designed to distinguish normal aging from early transitions to dementia because persons who develop significant cognitive impairments shortly after baseline are excluded to assure that the normative sample does not include participants already in a prodromal stage[15, 16]. The importance of this is increasingly recognized because of the search for early interventions. However, one study, by Ritchie et al[32], suggested that conventional norms performed similarly to robust norms in identifying early cognitive impairment.

Formal comparison is also difficult due to differences in reporting of the normative data. Nevertheless, some observations are noted. In whites, our −1.5 SD robust normative scores are approximately similar to the 3rd-5th percentile conventional normative scores reported by Mayo's Older Americans Normative Studies (MOANS) for LM I, LM II, WFT, Animal Naming, TMT-A, and TMT-B [33, 34]. Similarly, our −1.5 SD robust normative scores for whites are comparable to the 5th percentile conventional normative scores reported from the Monongahela-Youghiogheny Healthy Aging Team (MYHAT) cohort for LM I, LM II, Animal Naming, TMT-A, and TMT-B [35]. These similarities are likely attributable to similarities in design, including adjustment for age and education [33, 34,35].

Differences with prior studies are noted for normative scores in blacks. In our sample, −1.5 SD robust normative scores for blacks were higher than those reported for the 3rd-5th percentile conventional normative scores in Mayo's Older African American Normative Studies (MOAANS) for LM I and Animal Naming [8]; this pattern was most pronounced for younger age groups. These differences may also be due to MOAANS including a wider age range (56–99 years) or MOAANS not reporting education specific normative data (MOANNS reported normative data adjustment for IQ score)[8–10]. Scores for blacks in our study also differ in comparison with the (publically available) data from the National Alzheimer Disease Centers, notably for Animal Naming and TMT-A. Compared to the National Alzheimer Disease Centers data, −1.5 SD robust normative scores for blacks in our data tended to be higher on Animal Naming and lower on TMT-A, indicating better performance on these tests among our participants. These differences likely reflect differences in sample selection criteria (robust versus conventional normative populations, demographic factors (age, education), and point to the importance of considering comparability on these factors when selecting a normative reference sample.

Floor effects were noted in blacks for LM II and TMT-B, suggesting that the diagnostic utility of these measures, particularly for those with lower levels of education, may be limited. Similar findings were reported for TMT-B in MOAANS [8]. We allowed a maximum time of four minutes to complete TMT-B (compared to five minutes in MOAANS). While a

longer time limit could be considered, this may not be feasible in many clinical or research settings, where efficiency of the total testing battery and limiting participant/patient frustration may be of principal concern.

The use of race-specific cognitive norms has been an issue of debate [12, 36, 37]. It is important to note that white and black race represent social constructs rather than scientifically-based categories [38] and using these race categories ignore the heterogeneity of performance observed within these groups. Perhaps the largest drawback to depending on race-specific norms is that it leaves the reasons for these differences largely unexamined. Cognitive performance differences across race/ethnic categories, especially those observed in cross-sectional studies, may provoke unsupported views that the differences are biologic. Although we have not attempted to explain the racial differences observed in our study, when used appropriately, the current norms should help reduce misclassification of cognitive impairment in blacks with comparable demographic characteristics and facilitate identification of preclinical cognitive decline for those with MCI [12, 36].

Certain limitations of this study are noted. The applicability of any norms is dependent upon the degree of similarity between the individual test taker and the educational, cultural, and related demographic characteristics of those in the normative sample. The normative estimates in our study were based upon ARIC study participants from two geographic regions (Forsyth County, North Carolina and Jackson, Mississippi) in the US and their educational and cultural experiences may not be representative of blacks and whites from other regions. Although ARIC is a population-based study, those who agreed to be in the study and who further agreed to be participants in the more comprehensive ARIC Brain MRI study may be different from the general community in important ways (e.g., they may be in better health or have more interest in their health). It is up the individual clinician/investigator to determine whether and how best to apply these norms. To facilitate comparisons to other populations of interest, we have provided detailed characteristics of our sample.

Another limitation of our norms is the age range to which they are applicable. Although our participants ranged from 61 to 82 years in age, we were only able to create norms for ages 65 to 80 years due to the small number of participants outside of this age range. Additionally, although sex is a predictor of performance on some cognitive tests, we did not stratify on sex. Stratifying on sex, in addition to age, race, and education group (all of which were more strongly associated with cognitive performance than sex), would have resulted in unacceptably small cell sizes and less reliable normative estimates. We also did not have a measure of literacy in this study, which some have argued may be a better indicator of educational achievement and quality than years of education, which may vary by geographic region and race [37]. We hope to revisit this limitation in future work. It should also be noted that our sensitivity analysis comparing stratified −1.5 SD scores to −1.5 SD scores derived from the regression model should be interpreted with caution as the number of individuals in each cell for the stratified analysis varied and some cells contained an inadequate number of individuals for reliable estimation.

The measures in our study where selected to provide an efficient but relatively comprehensive assessment of cognitive functioning in the ARIC cohort, and have been widely used in studies of cognitive aging and dementia. The reported data will facilitate comparisons across cohort studies examining MCI and dementia. While several of the measures are currently used clinically, newer versions exist for the WAIS and WMS tests. Some clinicians, however, may prefer the versions presented here, particularly those who evaluate black participants or who desire consistency with the National Alzheimer Disease Centers recommended core cognitive battery [20]. Because all the tests presented were normed in a single sample, relative performance can be directly compared across tests, further facilitating clinical interpretation.

Strengths of the current study include a biracial sample with a wide age span and range of education levels. Most prior studies have excluded participants based on only overt clinical conditions that may affect performance on cognitive testing [5, 11, 12, 33]. Our study used rigorous criteria to exclude individuals with subclinical/latent disease or known risk factors for cognitive impairment (e.g. lacunar infarcts on brain MRI, APOE ε4 genotype, subsequent diagnosis of dementia) that may affect performance on cognitive testing. Our ability to identify and exclude these individuals from the normative sample provides increased sensitivity to detect cognitive impairment and addresses a key limitation of existing cross-sectional norms.

The rapidly expanding elderly population and the growing proportion of older ethnic minorities in the United States highlights the need for appropriate normative data to accurately identify cognitive impairment and dementia in diverse populations. This study provides age, race, and education specific robust normative data for 8 cognitive tests derived from a large population-based sample of blacks and whites, free of clinical and subclinical/latent neurological disease and dementia. These norms are intended to provide an additional option for clinicians and researchers, particularly in the evaluation of black patients/participants. Applied appropriately, the current norms should help improve the interpretation of performance on these measures for the diagnosis of cognitive impairment not only within the ARIC cohort, but also among older blacks and whites with similar demographic characteristics.

ACKNOWLEDGMENTS

The authors thank the staff and participants of the ARIC study for their important contributions.

FUNDING

The Atherosclerosis Risk in Communities (ARIC) Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C). The ARIC Brain MRI study was supported by grant R01-HL70825. A.L.C.S. was supported by NIH/NHLBI training grant T32 HL007024.

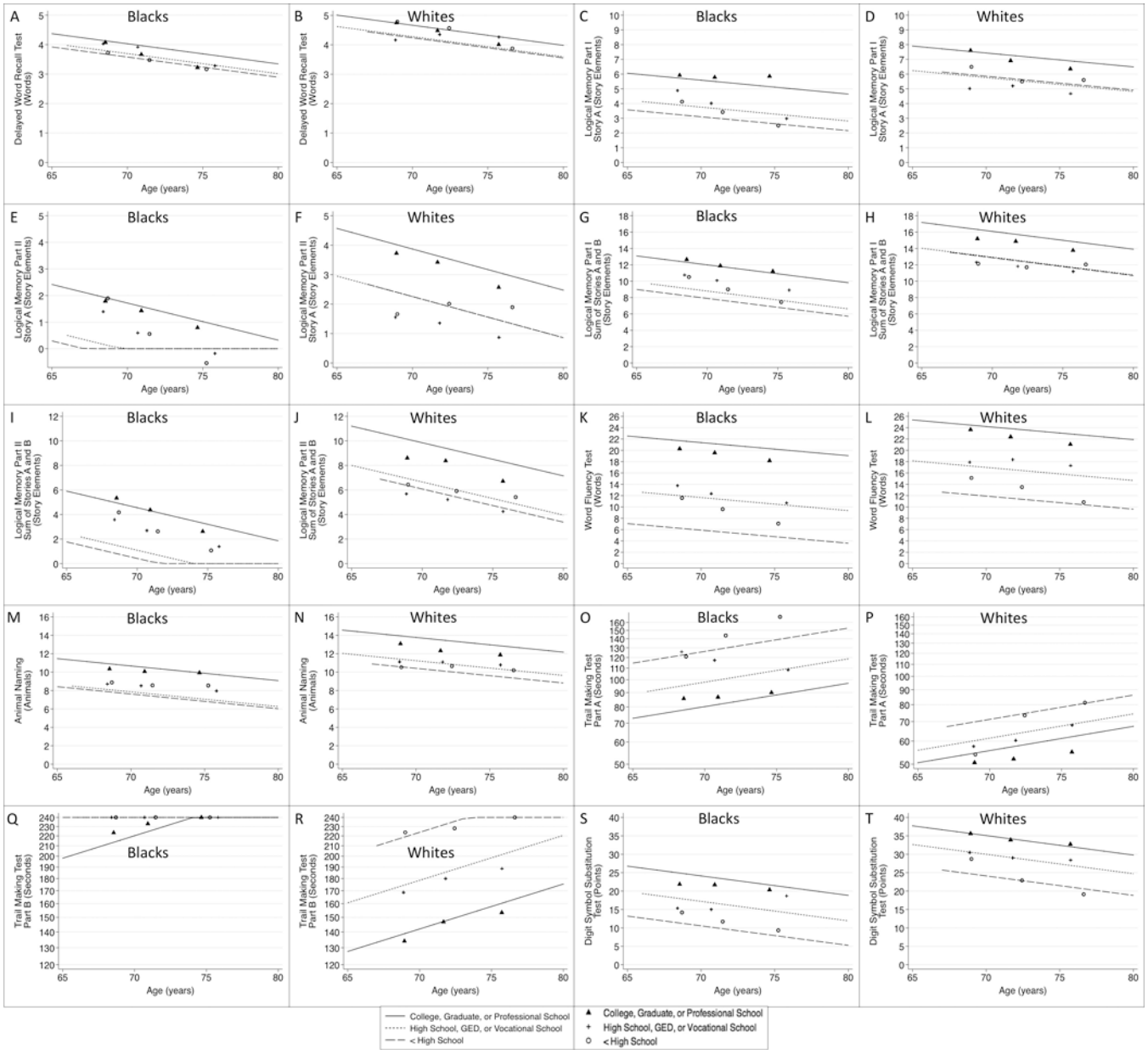
REFERENCES

1. Plassman BL, Langa KM, Fisher GG, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. Neuroepidemiology. 2007; 29:125–132. [PubMed: 17975326]

2. Anstey KJ, Cherbuin N, Eramudugolla R, et al. Characterizing mild cognitive disorders in the young-old over 8 years: Prevalence, estimated incidence, stability of diagnosis, and impact on IADLs. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2013
3. Hebert LE, Scherr PA, Bienias JL, Bennett DA, Evans DA. Alzheimer disease in the US population: prevalence estimates using the 2000 census. Arch Neurol. 2003; 60:1119–1122. [PubMed: 12925369]
4. Cerhan JR, Folsom AR, Mortimer JA, et al. Correlates of cognitive function in middle-aged adults. Atherosclerosis Risk in Communities (ARIC) Study Investigators. Gerontology. 1998; 44:95–105. [PubMed: 9523221]
5. Snitz BE, Unverzagt FW, Chang CC, et al. Effects of age, gender, education and race on two tests of language ability in community-based older adults. International psychogeriatrics / IPA. 2009; 21:1051–1062. [PubMed: 19586563]
6. Mayeux R, Reitz C, Brickman AM, et al. Operationalizing diagnostic criteria for Alzheimer's disease and other age-related cognitive impairment-Part 1. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2011; 7:15–34.
7. Schneider AL, Sharrett AR, Patel MD, et al. Education and cognitive change over 15 years: the atherosclerosis risk in communities study. Journal of the American Geriatrics Society. 2012; 60:1847–1853. [PubMed: 23013064]
8. Lucas JA, Ivnik RJ, Smith GE, et al. Mayo's Older African Americans Normative Studies: norms for Boston Naming Test, Controlled Oral Word Association, Category Fluency, Animal Naming, Token Test, WRAT-3 Reading, Trail Making Test, Stroop Test, and Judgment of Line Orientation. Clin Neuropsychol. 2005; 19:243–269. [PubMed: 16019707]
9. Lucas JA, Ivnik RJ, Smith GE, et al. Mayo's Older African Americans Normative Studies: WMS-R norms for African American elders. Clin Neuropsychol. 2005; 19:189–213. [PubMed: 16019704]
10. Lucas JA, Ivnik RJ, Smith GE, et al. A brief report on WAIS-R normative data collection in Mayo's Older African Americans Normative Studies. Clin Neuropsychol. 2005; 19:184–188. [PubMed: 16019703]
11. Lucas JA, Ivnik RJ, Willis FB, et al. Mayo's Older African Americans Normative Studies: normative data for commonly used clinical neuropsychological measures. Clin Neuropsychol. 2005; 19:162–183. [PubMed: 16019702]
12. Dotson VM, Kitner-Triolo M, Evans MK, Zonderman AB. Literacy-based normative data for low socioeconomic status African Americans. Clin Neuropsychol. 2008; 22:989–1017. [PubMed: 18609322]
13. Holtzer R, Goldin Y, Zimmerman M, Katz M, Buschke H, Lipton RB. Robust norms for selected neuropsychological tests in older adults. Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists. 2008; 23:531–541. [PubMed: 18572380]
14. Heaton, RK.; Heaton, RK. Revised comprehensive norms for an expanded Halstead-Reitan battery : demographically adjusted neuropsychological norms for African American and Caucasian adults, professional manual. Lutz, Fla.: Psychological Assessment Resources; 2004. Psychological Assessment Resources Inc. Updated ed.
15. De Santi S, Pirraglia E, Barr W, et al. Robust and conventional neuropsychological norms: diagnosis and prediction of age-related cognitive decline. Neuropsychology. 2008; 22:469–484. [PubMed: 18590359]
16. Holtzer R, Goldin Y, Zimmerman M, Katz M, Buschke H, Lipton RB. Robust norms for selected neuropsychological tests in older adults. Arch Clin Neuropsychol. 2008; 23:531–541. [PubMed: 18572380]
17. Roberts RO, Geda YE, Knopman DS, et al. The Mayo Clinic Study of Aging: design and sampling, participation, baseline measures and sample characteristics. Neuroepidemiolog. 2008; 30:58–69.
18. Singh-Manoux A, Britton AR, Marmot M. Vascular disease and cognitive function: evidence from the Whitehall II Study. J Am Geriatr Soc. 2003; 51:1445–1450. [PubMed: 14511166]
19. Murabito JM, Beiser AS, Decarli C, Seshadri S, Wolf PA, Au R. Parental longevity is associated with cognition and brain ageing in middle-aged offspring. Age Ageing. 2013

20. Morris JC, Weintraub S, Chui HC, et al. The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. Alzheimer Disease and Associated Disorders. 2006; 20:210–216. [PubMed: 17132964]
21. Shirk SD, Mitchell MB, Shaughnessy LW, et al. A web-based normative calculator for the uniform data set (UDS) neuropsychological test battery. Alzheimer's research & therapy. 2011; 3:32.
22. Manolio TA, Kronmal RA, Burke GL, et al. Magnetic resonance abnormalities and cardiovascular disease in older adults. The Cardiovascular Health Study. Stroke; a journal of cerebral circulation. 1994; 25:318–327.
23. Knopman DS, Ryberg S. A verbal memory test with high predictive accuracy for dementia of the Alzheimer type. Arch Neurol. 1989; 46:141–145. [PubMed: 2916953]
24. Wechsler, D. Wechsler Memory Scale Revised Manual. San Antonio: Psychological Corp; 1987.
25. Spreen, O.; Benton, A. Neurosensory Center Comprehensive Examination for Aphasia (NCCEA). Manual of instructions Victoria, BC: University of Victoria; 1969.
26. Benton, AL.; Hamsher, K. Multilingual Aphasia Examination. 2nd Edition. 1989.
27. Benton A. Differential Behavioral Effects in Frontal Lobe Disease. Neuropsychologia. 1968; 6:53–60.
28. Reitan R. Validity of the Trail Making Test as an Indicator of Organic Brain Damage. Percept Mot Skills. 1958; 8:271–276.
29. Wechsler D. Wechsler Adult Intelligence Scale Revised Manual. 1981
30. Petersen RC. Mild cognitive impairment as a diagnostic entity. Journal of Internal Medicine. 2004; 256:183–194. [PubMed: 15324362]
31. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. Journal of Internal Medicine. 2004; 256:240–246. [PubMed: 15324367]
32. Ritchie LJ, Frerichs RJ, Tuokko H. Effective normative samples for the detection of cognitive impairment in older adults. Clin Neuropsychol. 2007; 21:863–874. [PubMed: 17853155]
33. Steinberg BA, Bieliauskas LA, Smith GE, Ivnik RJ. Mayo's Older Americans Normative Studies: Age- and IQ-Adjusted Norms for the Wechsler Memory Scale--Revised. Clin Neuropsychol. 2005; 19:378–463. [PubMed: 16120536]
34. Steinberg BA, Bieliauskas LA, Smith GE, Ivnik RJ. Mayo's Older Americans Normative Studies: Age- and IQ-Adjusted Norms for the Trail-Making Test, the Stroop Test, and MAE Controlled Oral Word Association Test. Clin Neuropsychol. 2005; 19:329–377. [PubMed: 16120535]
35. Ganguli M, Snitz BE, Lee CW, Vanderbilt J, Saxton JA, Chang CC. Age and education effects and norms on a cognitive test battery from a population-based cohort: the Monongahela-Youghiogheny Healthy Aging Team. Aging & mental health. 2010; 14:100–107. [PubMed: 20155526]
36. Campbell AL Jr, Ocampo C, DeShawn Rorie K, et al. Caveats in the neuropsychological assessment of African Americans. Journal of the National Medical Association. 2002; 94:591–601. [PubMed: 12126285]
37. Manly JJ. Advantages and disadvantages of separate norms for African Americans. Clin Neuropsychol. 2005; 19:270–275. [PubMed: 16019708]
38. Herman AA. Toward a conceptualization of race in epidemiologic research. Ethnicity & disease. 1996; 6:7–20. [PubMed: 8882832]

Appendix Figure 1



−1.5 SD Scores for 8 Cognitive Tests from Linear Regression Models (lines) by Age at Time of Cognitive Testing and Education Category in Blacks and Whites Compared to −1.5 SD Scores Calculated in Stratified Analyses (symbols)

Panels A and B: Delayed Word Recall Test, Panels C and D: Logical Memory Part I Story A, Panels E and F: Logical Memory Part II Story A, Panels G and H: Logical Memory Part I Sum of Stories A and B, Panels I and J: Logical Memory Part II Sum of Stories A and B, Panels K and L: Word Fluency Test, Panels M and N: Animal Naming, Panels O and P:

Trail Making Test Part A, Panels Q and R: Trail Making Test Part B, Panels S and T: Digit Symbol Substitution Test.

Footnote.

Higher score represents worse performance on Trail Making Test Parts A and B. Data for Trail Making Test Part B is from a subsample of participants who completed this test (n=651). The length of the lines reflects age distribution of each age and education category.

Number of participants included in analysis for each stratum of the stratified analysis:

Blacks:

1.) Age 65- < 70; < High School (n=47 main analysis, n=31 TMT-B analysis), 2.) Age 65- < 70; High School (n=50 main analysis, n=43 TMT-B analysis), 3.) Age 65- < 70; > High School (n=112 main analysis, n=105 TMT-B analysis), 4.) Age 70- < 75; < High School (n=73 main analysis, n=44 TMT-B analysis), 5.) Age 70- < 75; High School (n=55 main analysis, n= 46 TMT-B analysis), 6.) Age 70- < 75; > High School (n=135 main analysis, n=122 TMT-B analysis); 7.) Age 75- < 80; < High School (n=41 main analysis, n=23 TMT-B analysis), 8.) Age 75- < 80; High School (n=31 main analysis, n=27 TMT-B analysis), 9.) Age 65- < 70; > High School (n=57 main analysis, n=49 TMT-B analysis).

Whites:

1.) Age 65- < 70; < High School (n=10 main analysis and TMT-B analysis), 2.) Age 65- < 70; High School (n=92 main analysis, n=91 TMT-B analysis), 3.) Age 65- < 70; > High School (n=106 main analysis and TMT-B analysis), 4.) Age 70- < 75; < High School (n=23 main analysis and TMT-B analysis), 5.) Age 70- < 75; High School (n=135 main analysis, n=133 TMT-B analysis), 6.) Age 70- < 75; > High School (n=134 main analysis and TMT-B analysis); 7.) Age 75- < 80; < High School (n=24 main analysis and TMT-B analysis), 8.) Age 75- < 80; High School (n=94 main analysis, n=93 TMT-B analysis), 9.) Age 65- < 70; > High School (n=89 main analysis, n=88 TMT-B analysis).

Appendix Table 1

Estimated Mean Race-Specific Z-Score Differences (95% Confidence Intervals) (from Linear Regression Model*) for 8 Cognitive Tests Comparing <High School Education to >High School Education by Race.

Domain	Test	Blacks	Whites
Memory	Delayed Word Recall Test	−0.27 (−0.52, −0.01)	−0.33 (−0.54, −0.01)
	Logical Memory Part I, Story A	−0.72 (−0.97, −0.47)	−0.40 (−0.76, −0.03)
	Logical Memory Part II, Story A	−0.58 (−0.83, −0.32)	−0.38 (−0.75, −0.01)
	Logical Memory Part I, Sum of Stories A and B	−0.68 (−0.93, −0.43)	−0.46 (−0.82, −0.10)
	Logical Memory Part II, Sum of Stories A and B	−0.64 (−0.89, −0.39)	−0.48 (−0.84, −0.12)

Domain	Test	Blacks	Whites
Verbal Fluency	Word Fluency Test	−1.23 (−1.45, −1.01)	−1.08 (−1.43, −0.73)
	Animal Naming	−0.70 (−0.94, −0.45)	−0.75 (−1.11, −0.39)
Executive Function	Trail Making Test Part A **	1.02 (0.78, 1.25)	0.79 (0.43, 1.14)
	Trail Making Test Part B **†	0.96 (0.68, 1.25)	1.17 (0.83, 1.50)
	Digit Symbol Substitution Test	−1.13 (−1.35, −0.90)	−1.10 (−1.44, −0.76)

* Model includes age and education.

** Represents ln(Trail Making Test Part A) and ln(Trail Making Test Part B) scores. Higher score represents worse performance.

† Subsample of participants who completed Trail Making Test Part B (n=262 blacks; n=389 whites)

Note: Education is defined as: 1.) <High school, 2.) High school or equivalent (includes high school, GED, or vocational school), and 3.) >High school (includes college, graduate, or professional school).

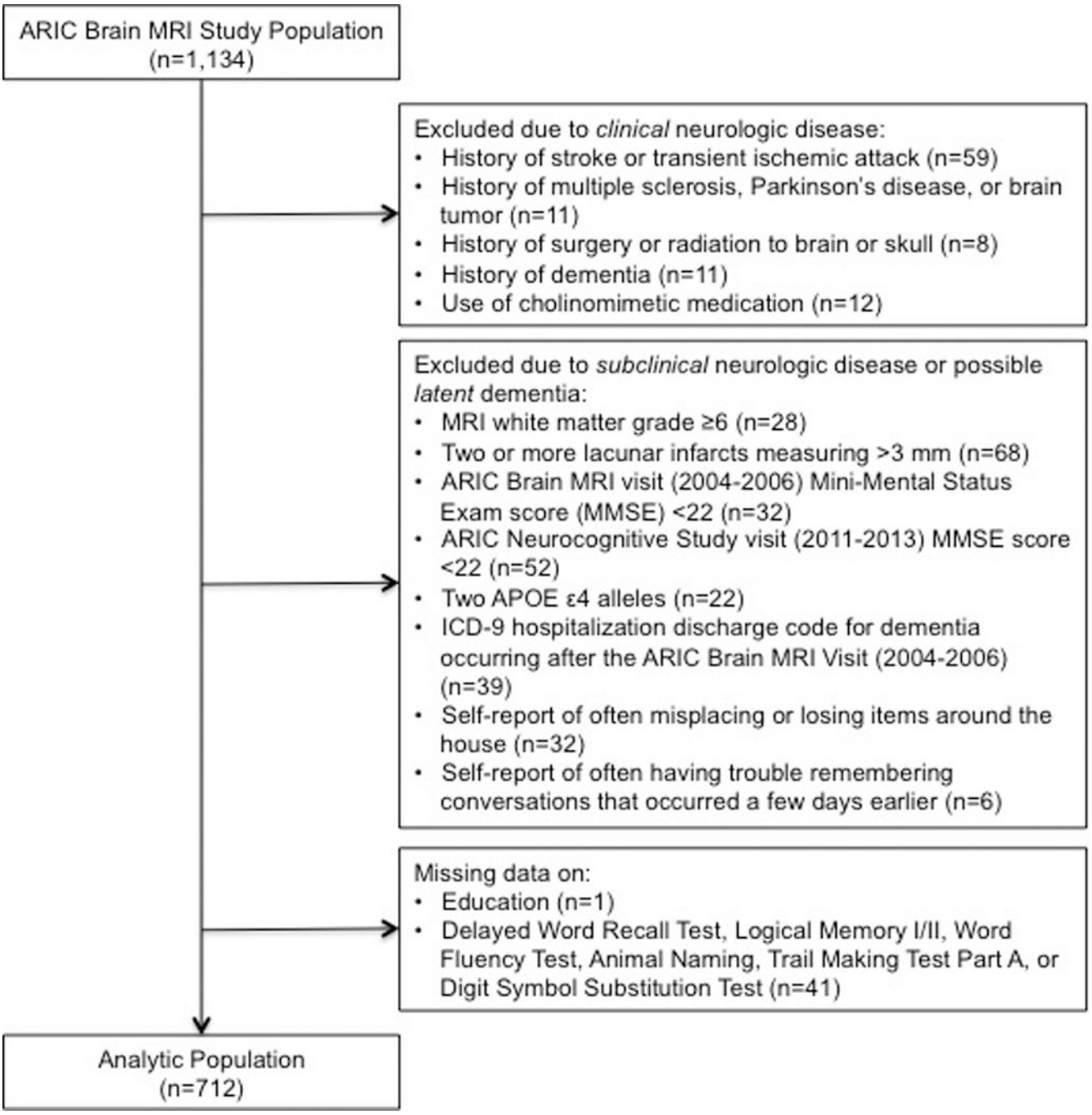


Figure 1. Participant Selection

Footnote.

Figure shows sequential exclusion criteria. The analytic population (n=651) for Trail Making Test Part B consists of a subset of the analytic population who were not missing data on this test (n=61 missing data).

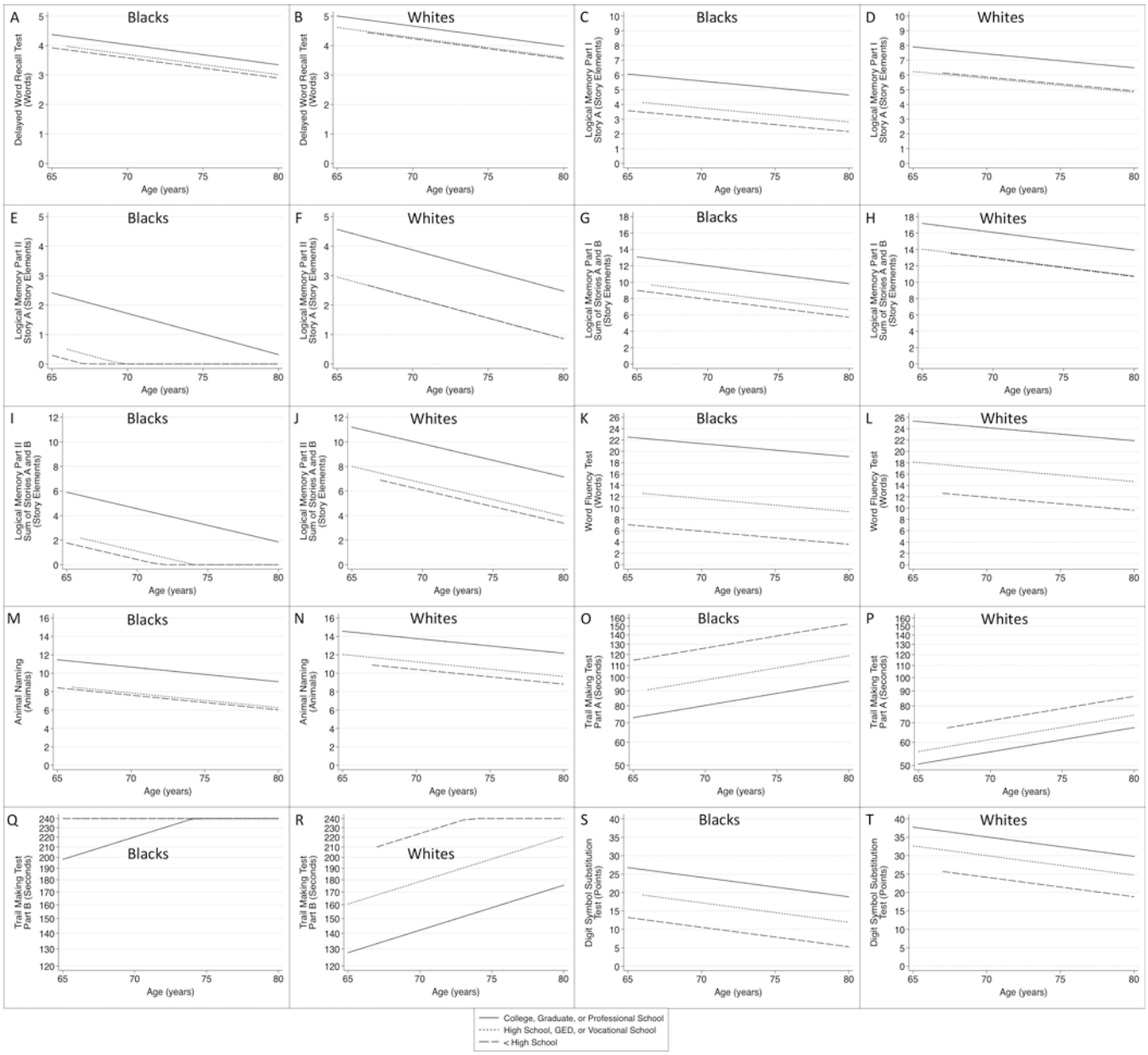


Figure 2. −1.5 SD Scores for 8 Cognitive Tests from Linear Regression Models by Age at Time of Cognitive Testing and Education Category in Blacks and Whites

Panels A and B: Delayed Word Recall Test, Panels C and D: Logical Memory Part I Story A, Panels E and F: Logical Memory Part II Story A, Panels G and H: Logical Memory Part I Sum of Stories A and B, Panels I and J: Logical Memory Part II Sum of Stories A and B, Panels K and L: Word Fluency Test, Panels M and N: Animal Naming, Panels O and P: Trail Making Test Part A, Panels Q and R: Trail Making Test Part B, Panels S and T: Digit Symbol Substitution Test.

Footnote.

Higher score represents worse performance on Trail Making Test Parts A and B. Data for Trail Making Test Part B is from a subsample of participants who completed this test (n=651). The length of the lines reflects age distribution of each age and education category.

Table 1

ARIC Brain MRI Study Participant Characteristics (ARIC Brain MRI Visit, 2004–2006).

	Included (n=712)			Excluded (n=422)			
	Blacks (n=320)	Whites (n=392)	P-Value*	Blacks (n=274)	Whites (n=148)	P-Value**	P-Value***
Age (years), mean (SD)	71.0 (4.1)	72.4 (4.3)	<0.001	73.1 (4.6)	73.5(4.1)	<0.001	0.009
Female, %	69.4	57.1	0.001	61.3	57.4	0.039	0.952
Education, %			<0.001			<0.001	0.293
<High school	26.6	8.4		44.5	12.8		
High school, GED, or vocational school	23.8	45.2		22.6	43.9		
College, graduate, or professional school	49.7	46.4		32.1	43.2		
Income, %			<0.001			0.002	0.639
<$35,000 per year	63.8	39.8		76.6	43.2		
≥$35,000 per year	24.7	56.4		14.2	52.0		
Not reported	11.6	3.8		9.1	4.7		
Cigarette smoking status, %			0.034			0.016	0.384
Current smoker	9.4	7.9		4.0	5.4		
Former smoker	30.0	39.0		36.9	44.6		
Never smoker	59.1	50.8		57.7	48.0		
Alcohol consumption status, %			<0.001			0.002	0.132
Current alcohol use	27.5	49.0		20.8	43.9		
Former alcohol use	52.8	20.9		46.7	29.1		
Never alcohol use	19.7	29.6		32.1	26.4		
Hypertension, %	75.6	58.4	<0.001	83.9	67.6	0.025	0.045
Diabetes, %	28.1	15.8	<0.001	40.9	21.0	0.001	0.159
APOE ε4 genotype, %			0.014			<0.001	<0.001
0 alleles	60.6	70.9		44.9	62.8		
1 allele	27.2	20.4		32.1	23.7		
2 alleles	0.0	0.0		9.1	6.1		
Mini-Mental Status Exam (points), mean (SD)	27.2 (1.9)	28.5 (1.5)	<0.001	25.0 (3.4)	27.9 (2.1)	<0.001	<0.001
11-Item Center for Epidemiological Studies-Depression Scale, mean (SD)	3.7 (3.2)	2.8 (2.7)	<0.001	5.3 (4.2)	4.0 (3.8)	<0.001	<0.001

	Included (n=712)			Excluded (n=422)			
	Blacks (n=320)	Whites (n=392)	P-Value*	Blacks (n=274)	Whites (n=148)	P-Value**	P-Value***
Delayed Word Recall Test (words), mean (SD)	5.95 (1.53)	6.48 (1.47)	<0.001	4.84 (1.92)	5.75 (1.86)	<0.001	<0.001
Logical Memory Part I, Story A (story elements), mean (SD)	9.8 (3.5)	11.7 (3.9)	<0.001	7.9 (3.8)	10.5 (4.1)	<0.001	0.002
Logical Memory Part II, Story A (story elements), mean (SD)	6.4 (3.7)	8.5 (4.2)	<0.001	4.4 (3.6)	7.1 (4.4)	<0.001	<0.001
Logical Memory Part I, Sum of Stories A and B (story elements), mean (SD)	19.4 (6.0)	23.3 (7.0)	<0.001	15.9 (6.6)	21.5 (7.2)	<0.001	0.009
Logical Memory Part II, Sum of Stories A and B (story elements), mean (SD)	12.8 (6.5)	17.9 (7.9)	<0.001	9.3 (6.2)	15.4 (8.0)	<0.001	0.001
Word Fluency Test (words), mean (SD)	30.6 (12.4)	35.2 (11.5)	<0.001	23.4 (11.7)	34.1 (10.9)	<0.001	0.322
Animal Naming (animals), mean (SD)	15.3 (4.3)	18.2 (4.6)	<0.001	13.3 (4.5)	16.7 (4.6)	<0.001	<0.001
Trail Making Test Part A (seconds), mean (SD)	64.5 (34.5)	39.3 (14.0)	<0.001	96.3 (55.5)	48.7 (24.5)	<0.001	0.002
Trail Making Test Part B (seconds), mean (SD)	163.0 (63.6)	107.3 (44.8)	<0.001	199.9 (53.0)	133.0 (55.6)	<0.001	<0.001
Digit Symbol Substitution Test (points), mean (SD)	32.8 (11.9)	45.0 (10.1)	<0.001	24.2 (11.0)	39.9 (10.4)	<0.001	<0.001

* P-value comparing included blacks to included whites.

** P-value comparing included blacks to excluded blacks.

*** P-value comparing included whites to excluded whites.

Note:
Of the 320 included blacks, the following data were missing: cigarette smoking status (n=5), hypertension status (n=15), APOE ε4 genotype (n=39), Trail Making Test Part B (n=58).
Of the 392 included whites, the following data were missing: cigarette smoking status (n=9), alcohol consumption (n=2), hypertension status (n=7), APOE ε4 genotype (n=34), 11-Item Center for Epidemiological Studies-Depression Scale (n=4), Trail Making Test Part B (n=3).
Of the 274 excluded blacks, the following data were missing: education (n=2), cigarette smoking status (n=4), alcohol consumption (n=1), hypertension status (n=8), APOE ε4 genotype (n=38), Mini-Mental Status Exam (n=1), 11-Item Center for Epidemiological Studies-Depression Scale (n=3), Delayed Word Recall Test (n=7), Logical Memory Part I Story A (n=2), Logical Memory Part II Story A (n=4), Logical Memory Part I Sum of Stories A and B (n=9), Logical Memory Part II Sum of Stories A and B (n=20), Word Fluency Test (n=10), Animal Naming (n=13), Trail Making Test Part A (n=24), Trail Making Test Part B (n=110), Digit Symbol Substitution Test (n=19).
Of the 148 excluded whites, the following data were missing: cigarette smoking status (n=3), alcohol consumption (n=1), hypertension status (n=3), APOE ε4 genotype (n=11), 11-Item Center for Epidemiological Studies-Depression Scale (n=1), Delayed Word Recall Test (n=1), Logical Memory Part I Sum of Stories A and B (n=4), Logical Memory Part II Sum of Stories A and B (n=3), Animal Naming (n=1), Trail Making Test Part A (n=4), Trail Making Test Part B (n=4), Digit Symbol Substitution Test (n=4).

Table 2

Regression β-Coefficients and Root Mean Squared Errors from Linear Regression Models from 8 Cognitive Tests.

	Intercept	Age*	Black	<High School	High School or Equivalent	Black× <High School Interaction	Black× High School or Equivalent Interaction	Root Mean Squared Error
Delayed Word Recall Test	**6.785**	**−0.068**	**−0.636**	-0.428	**-0.387**	−0.024	0.053	1.458
Logical Memory Part I, Story A	**12.719**	**−0.094**	**−1.846**	**−1.575**	**−1.669**	−0.910	−0.159	3.587
Logical Memory Part II, Story A	**9.519**	**−0.140**	**−2.152**	**−1.62**	**−1.617**	−0.510	−0.158	3.859
Logical Memory Part I, Sum of Stories A and B	**25.309**	**−0.218**	**−4.094**	**−3.222**	**−3.153**	−0.878	−0.0381	6.291
Logical Memory Part II, Sum of Stories A and B	**20.006**	**−0.270**	**−5.294**	**−3.770**	**−3.197**	−0.365	−0.268	6.952
Word Fluency Test	**39.811**	**−0.231**	**−2.841**	**−12.286**	**−7.220**	−3.171	−2.484	10.560
Animal Naming	**19.874**	**−0.159**	**−3.096**	**−3.354**	**−2.527**	0.306	−0.287	4.180
Trail Making Test Part A **	**3.526**	**0.019**	**0.365**	**0.247**	**0.098**	**0.206**	0.102	0.341
Trail Making Test Part B **†	**4.426**	**0.021**	**0.439**	**0.455**	**0.229**	−0.029	0.015	0.367
Digit Symbol Substitution Test	**48.967**	**−0.529**	**−10.953**	**−10.957**	**−5.077**	−2.606	−1.840	9.612

*Age is continuous and centered at median (71 years).

**Coefficients in table represent ln(Trail Making Test Part A) and ln(Trail Making Test Part B) scores. Higher score represents worse performance.

†Subsample of participants who completed Trail Making Test Part B (n=651).

Note: Education is defined as: 1.) <High school, 2.) High school or equivalent (includes high school, GED, or vocational school), and 3.) >High school (includes college, graduate, or professional school). Reference groups are white race and education >high school.
Bold coefficients represent p<0.05.

Table 3

Normative Data for 8 Cognitive Tests (Mean and −1.5 SD Scores) Derived from Linear Regression Model* by Age and Education Category in Blacks.

Test	Age**	Education	Mean	−1.5 SD
Delayed Word Recall Test (words)	65-<70	<High School	5.937	3.750
		High School or Equivalent	6.055	3.867
		>High School	6.389	4.201
	70-<75	<High School	5.595	3.408
		High School or Equivalent	5.713	3.525
		>High School	6.047	3.859
	75-<80	<High School	5.253	3.066
		High School or Equivalent	5.371	3.183
		>High School	5.705	3.517
Logical Memory Part I, Story A (story elements)	65-<70	<High School	8.718	3.337
		High School or Equivalent	9.374	3.993
		>High School	11.203	5.822
	70-<75	<High School	8.246	2.866
		High School or Equivalent	8.902	3.522
		>High School	10.731	5.351
	75-<80	<High School	7.774	2.394
		High School or Equivalent	8.430	3.050
		>High School	10.259	4.879
Logical Memory Part II, Story A (story elements)	65-<70	<High School	5.733	0.000†
		High School or Equivalent	6.080	0.293
		>High School	7.855	2.067
	70-<75	<High School	5.035	0.000†
		High School or Equivalent	5.382	0.000†
		>High School	7.157	1.369
	75-<80	<High School	4.337	0.000†
		High School or Equivalent	4.684	0.000†
		>High School	6.459	0.671
Logical Memory Part I, Sum of Stories A and B (story elements)	65-<70	<High School	17.878	8.442
		High School or Equivalent	18.777	9.341
		>High School	21.979	12.542
	70-<75	<High School	16.787	7.351
		High School or Equivalent	17.687	8.250

Test	Age**	Education	Mean	−1.5 SD
		>High School	20.888	11.452
	75-<80	<High School	15.697	6.261
		High School or Equivalent	16.596	7.160
		>High School	19.798	10.361
Logical Memory Part II, Sum of Stories A and B (story elements)	65-<70	<High School	11.521	1.094
		High School or Equivalent	12.192	1.765
		>High School	15.657	5.230
	70-<75	<High School	10.172	0.000†
		High School or Equivalent	10.842	0.415
		>High School	14.307	3.880
	75-<80	<High School	8.822	0.000†
		High School or Equivalent	9.492	0.000†
		>High School	12.957	2.530
Word Fluency Test (words)	65-<70	<High School	22.321	6.481
		High School or Equivalent	28.075	12.235
		>High School	37.778	21.938
	70-<75	<High School	21.166	5.326
		High School or Equivalent	26.920	11.080
		>High School	36.623	20.783
	75-<80	<High School	20.011	4.171
		High School or Equivalent	25.765	9.925
		>High School	35.468	19.628
Animal Naming (animals)	65-<70	<High School	14.288	8.018
		High School or Equivalent	14.521	8.251
		>High School	17.336	11.066
	70-<75	<High School	13.491	7.221
		High School or Equivalent	13.725	7.455
		>High School	16.539	10.269
	75-<80	<High School	12.694	6.424
		High School or Equivalent	12.928	6.658
		>High School	15.742	9.472
Trail Making Test – Part A (seconds) //	65-<70	<High School	72.065	120.245
		High School or Equivalent	55.992	93.426
		>High School	45.809	76.436
	70-<75	<High School	79.315	132.343

Test	Age**	Education	Mean	−1.5 SD
		High School or Equivalent	61.625	102.826
		>High School	50.418	84.126
	75-<80	<High School	87.295	145.657
		High School or Equivalent	67.825	113.170
		>High School	55.491	92.590
Trail Making Test – Part B (seconds) //‡	65-<70	<High School	184.364	240.000§
		High School or Equivalent	153.721	240.000§
		>High School	120.410	208.846
	70-<75	<High School	204.930	240.000§
		High School or Equivalent	170.868	240.000§
		>High School	133.841	232.141
	75-<80	<High School	227.789	240.000
		High School or Equivalent	189.927	240.000§
		>High School	148.770	240.000§
Digit Symbol Substitution Test (points)	65-<70	<High School	26.303	11.885
		High School or Equivalent	32.950	18.531
		>High School	39.866	25.448
	70-<75	<High School	23.657	9.239
		High School or Equivalent	30.304	15.885
		>High School	37.220	22.802
	75-<80	<High School	21.011	6.593
		High School or Equivalent	27.658	13.239
		>High School	34.574	20.156

*Model includes terms for age, race, education, and race X education interaction.

**Scores derived using the midpoint of each age category.

†Minimum possible score is 0 story elements recalled.

//Higher score represents worse performance.

‡Subsample of participants who completed Trail Making Test Part B (n=262 blacks).

§Maximum possible score is 240 seconds.

Note: Education is defined as: 1.) <High school, 2.) High school or equivalent - high school, GED, or vocational school, and 3.) >High school - college, graduate, or professional school.

Table 4

Normative Data for 8 Cognitive Tests (Mean and −1.5 SD Scores) Derived from Linear Regression Model* by Age and Education Category in Whites.

Test	Age**	Education	Mean	−1.5 SD
Delayed Word Recall Test (words)	65-<70	<High School	6.597	4.409
		High School or Equivalent	6.638	4.450
		>High School	7.025	4.837
	70-<75	<High School	6.255	4.067
		High School or Equivalent	6.296	4.108
		>High School	6.683	4.495
	75-<80	<High School	5.913	3.725
		High School or Equivalent	5.954	3.766
		>High School	6.341	4.153
Logical Memory Part I, Story A (story elements)	65-<70	<High School	11.474	6.094
		High School or Equivalent	11.380	5.999
		>High School	13.049	7.669
	70-<75	<High School	11.003	5.622
		High School or Equivalent	10.908	5.527
		>High School	12.577	7.197
	75-<80	<High School	10.531	5.150
		High School or Equivalent	10.436	5.056
		>High School	12.106	6.725
Logical Memory Part II, Story A (story elements)	65-<70	<High School	8.395	2.608
		High School or Equivalent	8.391	2.603
		>High School	10.007	4.220
	70-<75	<High School	7.697	1.909
		High School or Equivalent	7.693	1.905
		>High School	9.309	3.521
	75-<80	<High School	6.999	1.211
		High School or Equivalent	6.994	1.207
		>High School	8.611	2.823
Logical Memory Part I, Sum of Stories A and B (story elements)	65-<70	<High School	22.850	13.414
		High School or Equivalent	22.919	13.483
		>High School	26.072	16.636
	70-<75	<High School	21.760	12.323
		High School or Equivalent	21.829	12.392
		>High School	24.982	15.545

Test	Age**	Education	Mean	−1.5 SD
	75-<80	<High School	20.669	11.233
		High School or Equivalent	20.738	11.302
		>High School	23.891	14.455
Logical Memory Part II, Sum of Stories A and B (story elements)	65-<70	<High School	17.180	6.753
		High School or Equivalent	17.753	7.326
		>High School	20.950	10.523
	70-<75	<High School	15.831	5.403
		High School or Equivalent	16.403	5.976
		>High School	19.601	9.173
	75-<80	<High School	14.481	4.054
		High School or Equivalent	15.054	4.626
		>High School	18.251	7.824
Word Fluency Test (words)	65-<70	<High School	28.334	12.494
		High School or Equivalent	33.400	17.560
		>High School	40.620	24.780
	70-<75	<High School	27.179	11.339
		High School or Equivalent	32.245	16.405
		>High School	39.465	23.625
	75-<80	<High School	26.024	10.184
		High School or Equivalent	31.090	15.250
		>High School	38.310	22.470
Animal Naming (animals)	65-<70	<High School	17.078	10.808
		High School or Equivalent	17.905	11.635
		>High School	20.432	14.162
	70-<75	<High School	16.281	10.011
		High School or Equivalent	17.108	10.838
		>High School	19.635	13.365
	75-<80	<High School	15.484	9.215
		High School or Equivalent	16.311	10.041
		>High School	18.838	12.569
Trail Making Test – Part A (seconds)†	65-<70	<High School	40.688	67.890
		High School or Equivalent	35.079	58.531
		>High School	31.792	53.048
	70-<75	<High School	44.781	74.721
		High School or Equivalent	38.608	64.420
		>High School	34.991	58.385

Test	Age**	Education	Mean	−1.5 SD
	75-<80	<High School	49.286	82.238
		High School or Equivalent	42.493	70.902
		>High School	38.511	64.259
Trail Making Test – Part B (seconds) †//	65-<70	<High School	122.364	212.235
		High School or Equivalent	97.632	169.339
		>High School	77.650	134.680
	70-<75	<High School	136.013	235.909
		High School or Equivalent	108.523	188.228
		>High School	86.311	149.703
	75-<80	<High School	151.186	240.000 ‡
		High School or Equivalent	120.628	209.224
		>High School	95.939	166.402
Digit Symbol Substitution Test (points)	65-<70	<High School	39.862	25.444
		High School or Equivalent	45.742	31.324
		>High School	50.819	36.401
	70-<75	<High School	37.216	22.798
		High School or Equivalent	43.096	28.678
		>High School	48.173	33.755
	75-<80	<High School	34.570	20.152
		High School or Equivalent	40.450	26.032
		>High School	45.527	31.109

*Model includes terms for age, race, education, and race X education interaction.

**Scores derived using the midpoint of each age category.

†Higher score represents worse performance.

//Subsample of participants who completed Trail Making Test – Part B (n=389 whites).

‡Maximum possible score is 240 seconds.

Note: Education is defined as: 1.) <High school, 2.) High school or equivalent - high school, GED, or vocational school, and 3.) >High school - college, graduate, or professional school.