NIH Public Access
Author Manuscript
Alzheimers Dement. Author manuscript; available in PMC 2015 November 01.

Published in final edited form as:
Alzheimers Dement. 2014 November ; 10(6): 666–674. doi:10.1016/j.jalz.2014.02.002.

An empirically derived composite cognitive test score with improved power to track and evaluate treatments for preclinical Alzheimer's disease

Jessica B.S. Langbaum[1,11,*], **Suzanne B. Hendrix**[2], **Napatkamon Ayutyanont**[1,11], **Kewei Chen**[1,4,11], **Adam S. Fleisher**[1,11], **Raj C. Shah**[6,7], **Lisa L. Barnes**[6,8,9], **David A. Bennett**[6,8], **Pierre N. Tariot**[1,3,11], and **Eric M. Reiman**[1,3,10,11]

[1]Banner Alzheimer's Institute, Phoenix, AZ, USA

[2]Pentara Corporation, Salt Lake City, UT, USA

[3]Department of Psychiatry, University of Arizona, Tucson, AZ, USA

[4]Department of Mathematics and Statistics, Arizona State University, Tempe, AZ, USA

[6]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA

[7]Department of Family Medicine, Rush University Medical Center, Chicago, IL, USA

[8]Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA

[9]Department of Behavioral Sciences, Rush University Medical Center, Chicago, IL, USA

[10]Neurogenomics Division, Translational Genomics Research Institute, Phoenix, AZ, USA

[11]Arizona Alzheimer's Consortium, Phoenix, AZ, USA

Abstract

Background—There is growing interest in the evaluation of preclinical Alzheimer's disease (AD) treatments. As a result, there is a need to identify a cognitive composite that is sensitive to tracking preclinical AD decline to be used as a primary endpoint in treatment trials.

Methods—Longitudinal data from initially cognitively normal, 70–85 year old participants in three cohort studies of aging and dementia from the Rush Alzheimer's Disease Center were examined to empirically define a composite cognitive endpoint that is sensitive to detecting and tracking cognitive decline prior to the onset of cognitive impairment. The mean to standard deviation ratios (MSDR) of change over time were calculated in a search for the optimal combination of cognitive tests/sub-tests drawn from the neuropsychological battery in cognitively

© 2014 Elsevier Inc. All rights reserved.

*To whom correspondence should be addressed: Jessica Langbaum, PhD, Banner Alzheimer's Institute, 901 E. Willetta Street, Phoenix, AZ 85006, telephone: 602.839.2548; Jessica.Langbaum@bannerhealth.com.

Portions of this study were presented at the 2011 Alzheimer's Association International Conference (AAIC), Paris, France and the 2011 Clinical Trials on Alzheimer's Disease (CTAD), San Diego, CA.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

normal participants who subsequently progressed to clinical stages of AD during a two and five year period, using data from those who remained unimpaired during the same time period to correct for aging and practice effects. Combinations that performed well were then evaluated for representation of relevant cognitive domains, robustness across individual years prior to diagnosis, and occurrence of selected items within top performing combinations.

Results—The optimal composite cognitive test score is comprised of 7 cognitive tests/sub-tests with an MSDR=0.964. By comparison, the most sensitive individual test score, *Logical Memory – Delayed Recall*, MSDR= 0.64.

Conclusions—We have identified a composite cognitive test score representing multiple cognitive domains that has improved power compared to the most sensitive single test item to track preclinical AD decline and evaluate preclinical AD treatments. We are confirming the power of the composite in independent cohorts, and with other analytical approaches, which may result in refinements, and have designated it as the primary endpoint in the Alzheimer's Prevention Initiative's preclinical treatment trials for individuals at high imminent risk for developing symptoms due to late-onset AD.

Introduction

Without an effective treatment that postpones the onset or completely prevents the clinical consequences of Alzheimer's disease (AD), the number of individuals afflicted by the disease will continue to rapidly increase (1;2). There is growing interest in the hypothesis that interventions may have their most profound effect if initiated in the preclinical AD phase (3), that is, in the absence of mild cognitive impairment (MCI) or AD dementia (4). Several such trials are underway or are in various planning stages, including those with the strategy of testing therapies in people who are at highest imminent risk of developing MCI or AD dementia due to factors such as age and genetic disposition or presence of biomarker evidence of AD (4–8). Traditional clinical outcomes, such as progression to clinical diagnosis, or cognitive outcomes developed for studies in MCI or AD dementia may not be well-suited for some preclinical treatment trials due to large sample size and long trial duration requirements or the psychometric properties of the tests themselves (9–12). Moreover, individually examining each cognitive assessment and treating as individual outcomes inflates Type 1 error if appropriate corrections are not made to guard against multiple comparisons. Use of an appropriate composite reduces the number of variables employed and thus risk of Type 1 error, it can be empirically derived and its sensitivity to detecting and tracking preclinical AD can be validated in multiple datasets. As a result, it affords a measure of multiple domains that can serve as a primary endpoint in preclinical treatment trials (13).

Small, but measurable cognitive decline occurs during preclinical AD. For instance, retrospective and prospective studies of cognitively healthy individuals who eventually progressed to AD dementia have shown episodic memory decline to be a defining feature of preclinical AD (14–18). In addition, decline in other cognitive domains, such as executive (19), visual spatial (16) and global cognitive functioning (16;20) occurs during the transition from normal aging to preclinical AD and into the clinical stages of AD. Studies of cognitively healthy individuals with significant fibrillar amyloid burden report decline

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

primarily in episodic memory, executive function and language (21–25). Long-term recall memory performance has been found to begin to decline in relationship to apolipoprotein E (*APOE*) ε4 gene dose, reflecting three levels of genetic risk for late-onset AD, despite maintenance of normal clinical status (26).

There are multiple approaches for selecting an appropriate cognitive endpoint for use in preclinical AD studies and therapeutic trials. For instance, a theoretically driven approach reasons that a composite should be constructed a priori from cognitive assessments known to decline during preclinical AD. A related approach is to construct composites specific to an individual cognitive domain, such as memory (27) or executive functioning (28). Yet another is an empirically driven approach, in which the endpoint or composite is selected based on analyses demonstrating sensitivity (e.g., has the greatest power) to detect and track the outcome of interest, such as preclinical AD decline. These approaches are not necessarily mutually exclusive; for instance, theoretical knowledge of preclinical AD can be taken into account when empirically deriving a composite cognitive test score. Several different analyses methods are available to developing composites, including but not limited to latent variable analyses or partial least squares regression (29–31), principal components (32), item-response theory (33), Rasch Measurement Theory (34) or item-level analysis (35). While there have been some efforts focused on refining existing cognitive assessments, this may be best suited for MCI and early AD trials (36).

Here we propose a strategy to empirically determine the combination of cognitive assessments most sensitive to the tracking of preclinical AD in individuals who subsequently progress to MCI or probable AD dementia, while controlling for practice and normal aging effects using data from individuals who did not progress to the clinical stages of AD over the same duration. The goal of the present study was to develop a composite with optimal sensitivity to decline, not limited to a single cognitive domain, corresponding to a change from baseline analysis. This approach differs from optimizing an endpoint for discriminating those who progress from those who remain stable, which would result in a composite that could be used as a progression endpoint in preclinical treatment trials. We hypothesize that the composite will be more sensitive (i.e., have greater power) to detecting and tracking preclinical AD decline compared to the most sensitive individual cognitive test/sub-test score given that the approach allows for the addition of assessments that improve sensitivity overall, despite perhaps being less sensitive individually to preclinical AD decline. Longitudinal data from three cohort studies of aging and dementia at the Rush Alzheimer's Disease Center in those who did and did not clinically progress over a two and five year period were used to develop a composite cognitive endpoint, employing the mean to standard deviation ratio (MSDR) of the change score as the measure of sensitivity to preclinical AD decline over time (31). The results from the present study are informing the design of trials for the Alzheimer's Prevention Initiative (API) focused on individuals at high imminent risk for symptoms of late-onset AD based on their age and genetics.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Materials and Methods

Participants

Data from participants enrolled in the Rush Alzheimer's Disease Center's Religious Orders Study [ROS], Memory and Aging Project [MAP], or the Minority Aging Research Study [MARS] was downloaded on June 7, 2010. Enrollment criteria for the three studies are quite similar and have been previously reported (37–39). Briefly, participants from each study are older adults without dementia at the time of enrollment and who agree to annual clinical and neuropsychological evaluations, and for those in ROS and MAP, agree to brain donation at the time of death. For the present study, longitudinal data from participants who subsequently progressed to either MCI or AD dementia (possible or probable) and who had at least 1 follow-up visit (n = 1073), 2 years of data (n = 528) or 5 years of data (n = 213) and from participants who remained cognitively normal during the same time period and had at least 2 (n = 831) or 5 years (n = 413) of follow-up data were included in the analyses (Table 1). The studies were approved by the Institutional Review Board of Rush University Medical Center and each participant signed an informed consent.

Cognitive and Clinical Evaluations

The three studies included uniform and structured annual clinical evaluations with medical history questions, neurological examination, and detailed cognitive testing using a battery of 19–21 neuropsychological tests (40;41). Diagnostic classification followed a multi-step procedure as previously described (42;43). Briefly, neuropsychological tests encompassing a wide range of cognitive functions were administered by trained technicians and scores were adjusted for education using an automated scoring algorithm computed in SAS (43;44). Participants were examined or records were reviewed by a clinician (primarily neurologists or geriatricians, supplemented by advanced practice geriatric nurses and neuropsychologists) and diagnostically classified using the recommendations of the joint working group of the National Institute of Neurologic and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) (45). The diagnosis of dementia and probable AD dementia followed a three step procedure and was validated pathologically (43;44;46). History of cognitive decline was determined by structured interview, evidence of impairment in memory and other cognitive abilities was based on neuropsychological performance tests summarized by an experienced neuropsychologist. Dementia due to AD and other causes was based on review of records and an interview and examination by a clinician with expertise in dementia evaluations. MCI was based on the presence of cognitive impairment as determined by the neuropsychologist and the absence of dementia as determined by the clinician (43;47;48).

Data Analysis

Data from tests were included in the following analyses only if all cohorts (ROS, MAP and MARS) received the particular assessment. For example, only MAP and MARS participants had Stroop Color Naming test data; as a result, this assessment was not included in the analyses. The scores for each test, T, were standardized to a range of 0.0 to 1.0 as shown in the following equation:

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$$x_{j,i}^{T}=\frac{S_{j,i}^{T}-\min(\mathrm{T}_{scr})}{\max(\mathrm{T}_{scr})-\min(\mathrm{T}_{scr})}$$

where $S_{j,i}^{T}$ is the score at time $t_i^j\ i\ (i=1,2)$ for subject *j*, and T_{scr} symbolizes all the possible scores this test T can take.

The annualized difference score y_j^T for subject *j*, test *T* is defined as

$$y_j^T=\frac{x_{j,2}^{T}-x_{j,1}^{T}}{t_2^j-t_1^j}$$

For assessments without a defined maximum score (such as category fluency), a maximum threshold was established that was 2 standard deviations above the mean. All analyses were conducted using SAS v. 9.2 (SAS Institute Inc, Cary, NC, USA).

Cognitive domains that were expected to change early in the disease process were identified prior to analyses. Sensitivity of each of the cognitive tests/sub-tests and correlations between them were examined, with the intention of identifying items to represent the relevant cognitive domains in a composite. Due to the complexity of constructing a multivariate composite based on univariate or bivariate summary statistics, an integrated approach was used to evaluate all possible combinations of items to optimize the sensitivity resulting in an analysis that is mathematically closely related to principal component analysis (43) as follows:

$X=\sum_{T=1}^{K} w_T y^T$, where $y^T=(\mathrm{y}_1^T, \mathrm{y}_2^T \ldots \mathrm{y}_N^T)'$, *N* is the number of subjects, *K* is the number of tests, w_T is the weight for test *T*, and we have $w_T \geq 0$ (*T=1*, …, *K*) and $\Sigma w_T = 1$.

The optimization criterion is the maximization of MSDR of X with regards to weights of 0 or 1 for each item for the exclusion or inclusion of that particular item:

$MSDR=\frac{\overline{X}}{\mathrm{std}(X)}$ where $\bar{X}$ is the mean of *X* and std(*X*) is its standard deviation.

Maximizing the MSDR across all combinations of items is not a statistical inference procedure; rather, it is a method that produces a metric that can guide our search for combinations that are sensitive. Those participants who eventually progressed to cognitive impairment (MCI or probable dementia AD), were assessed with this methodology, looking backwards zero to two (n = 528) or five years (n = 213) prior to diagnosis. To most accurately reflect an API preclinical treatment trial, a person was considered diagnosed when their clinical diagnosis first progressed from “no cognitive impairment” to either MCI or AD. In order to account for longitudinal aging and practice effects, which impact the sensitivity of outcome measures in this preclinical population (49;50), annualized MSDRs

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

were also calculated for those who remained cognitively normal during the two (n = 831) or five-year (n = 413) follow-up period. Adjusted MSDRs were calculated by subtracting the mean for the controls from that obtained from individuals who progressed to MCI or dementia prior to dividing by the standard deviation.

Results from these analyses were used as one way to assess the combinations and determine a “best” composite. Items that were consistently represented in the combinations with the highest sensitivity and that also demonstrated consistency within separate years of the 1–5 year time period were identified as robust items for measuring change. Construct validity was assessed by giving preference to combinations that represented multiple cognitive domains known to be important that also had consistent sensitivity across the two and five years of decline. Corresponding sample size estimates were calculated using the adjusted MSDRs, though it should be noted that these estimates are only to aid in gauging the comparable sensitivity of the tests and should not be directly used for powering a trial given that the estimates were calculated based on data from individuals who all subsequently progressed to MCI or AD dementia, and it would be impossible to enroll a similar population into a trial.

Following selection of the composite, we sought to determine whether the corresponding MSDR could be substantially increased (i.e., improved, resulting in smaller sample sizes) by weighting the individual assessments included in the composite (as opposed to the equal weights used in the initial analyses). A search of potential weighting combination, such that the sum of the weighting of the individual assessments equaled one, was conducted in those who progressed to cognitive impairment (MCI or AD) at two and five years prior to diagnosis.

Results

Participant Characteristics

There were some demographic differences between participants who progressed to clinical stages of AD and those who remained cognitively healthy during the same time period (Table 1). For instance, in the MSDR analyses of the two years and five years prior to diagnosis, those who progressed to clinical stages of AD were older ($p < 0.0001$) and had different percent distribution between the racial categories (two-year analysis $p < 0.0001$; five-year analysis $p = 0.01$).

Individual Cognitive Assessment Properties

The sensitivity of the individual cognitive assessments over the five years prior to a diagnosis of MCI or probable AD dementia, unadjusted for normal aging effects, is shown in Table 2. Among individuals who progressed to clinical stages of AD, *Category Fluency (fruits and vegetables)* had the highest sensitivity of all tests (MSDR 0.825), followed by *Symbol Digit Modalities* (0.71), and *MMSE Total* (0.665). Adjusting for longitudinal aging and practice effects using data from those participants who remained cognitively normal during the same time period resulted in an increased MSDR for some cognitive assessments (e.g., *Logical Memory Delayed Recall* adjusted MSDR = 0.64) due to observed increases in

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

normals, while others decreased (e.g., *Symbol Digit Modalities* adjusted MSDR = 0.385) due to observed worsening in normals (Table 2).

Deriving the Alzheimer's Prevention Initiative (API) Composite Cognitive Test Score

Results from the MSDR calculation for every possible combination of neuropsychological assessments indicate that the composite most sensitive to detecting preclinical cognitive decline related to AD, adjusting for longitudinal aging and practice effects, that has construct validity and also includes items that are robust across individual years consisted of: *Category Fluency (fruits and vegetables), Boston Naming Test (15 item), Logical Memory – Delayed Recall, East Boston Naming Test – Immediate Recall, Ravens Progressive Matrices Subset (9 items), Symbol Digit Modalities, and MMSE Orientation to Time* (Table 2). Based on the data five years prior to diagnosis, the total five-year MSDR of the composite is 0.9639. The best 50 combinations had annual MSDRs that ranged from 0.1928–0.1862 and were comprised of 6–7 test/sub-tests of episodic memory, working memory, language, global functioning, and visual spatial ability.

A composite with an MSDR of 0.9639 requires an estimated 264 completers per treatment arm to detect a 25% treatment effect in a five-year trial, noting that caveats mentioned previously of applying this sample size estimate when designing a prevention trial still apply. In comparison, the most sensitive individual cognitive assessment is Logical Memory Delayed Recall, with a five-year MSDR of 0.640 and an estimated sample size of 611 completers per treatment arm to detect a 25% treatment effect, making the API composite cognitive test score considerably more sensitive to tracking preclinical AD decline (Figure 1). Based on the data two years prior to diagnosis, a shorter study with the same composite cognitive test score would result in a lower total two-year MSDR of 0.3398. Results from analyses based on the search of potential weighting combinations revealed that weighting provided minimal improvement over the unweighted MSDRs (increase in MSDR < 5%).

Discussion

We empirically identified an API composite cognitive test score sensitive to preclinical AD decline, and suggest that it can be used in preclinical trials to evaluate treatment effects with smaller sample sizes and improved statistical power compared to the most sensitive individual cognitive assessments and larger test batteries, and in a manner that is reasonably likely to predict a treatment's effect on clinical progression to MCI or AD. This API composite cognitive test score and the approach taken to develop it appears to fit into the framework provided by the Food and Drug Administration's recent draft guidance concerning a cognitive assessment serving as a primary efficacy measure in preclinical AD trials (51). Moreover, the composite identified in the present study was comprised of the same cognitive domains/assessments with the exception of one (present study included a test of visual spatial ability – *Symbol Digits Modalities*) as a composite cognitive endpoint identified in cognitively healthy ADAD mutation carriers, despite substantial differences in the cohorts' neuropsychological test batteries (52). This consistency is noteworthy as confirmation of the composite's performance in an independent population, as well as suggesting that there is extensive overlap in the pattern of cognitive decline between the two

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

forms of the disease even though they strike at different ages, and may have different preclinical and clinical time courses, underlying etiologies, and biological processes.

We employed an empirical strategy refined by theoretical understanding of preclinical AD to develop a sensitive composite cognitive endpoint, controlling for aging and practice effects, by examining longitudinal data in the two and five years prior to clinical progression. With this approach, we focused primarily on the aspects of the disease that decline consistently across individuals in order to assess effectiveness of a treatment in slowing decline in a preclinical trial, rather than discrimination between those who progress and those that do not, or the neuropathological underpinnings of AD that result in a change in cognitive functioning. This approach has the added advantage in that it incorporates data from participants at various points along the preclinical AD continuum and does not presuppose the cognitive assessments sensitive to detecting and tracking this decline. Just as in a clinical trial, some participants may progress to cognitive impairment within months while others are several years away.

The optimal composite cognitive test score identified in the current study incorporates cognitive assessments from several different domains, complementing recent studies suggesting that multiple cognitive domains decline in preclinical AD (16;53), not just decline in episodic memory, though that remains a defining feature of preclinical AD (14–18). With the exception of the Ravens Progressive Matrices, all of the assessments that comprise the composite do have a language component. Although on an individual basis, the neuropsychological assessments have varying levels of sensitivity to tracking preclinical decline, as measured by their MSDR, this empirical strategy, focused on the years prior to diagnosis (while controlling for longitudinal aging and practice effects), identified these items as providing a sensitive combination of assessments across multiple cognitive domains. More sensitive cognitive tests/sub-tests may not be included in the composite endpoint, since these items may correlate with another assessment that captures the same information and has a higher MSDR (32). The cognitive tests/sub-tests that are included and have a smaller MSDR may measure variability not captured by other assessments. As a result, the optimal composite was not comprised solely of the cognitive tests with the largest MSDRs.

There are some limitations to the present study. Development of the optimal composite cognitive endpoint was constrained by the neuropsychological test battery used in the Rush cohort studies. That said, we achieved remarkably similar results with independent efforts with an empirically driven approach to derive a composite cognitive endpoint for ADAD mutation carriers despite differences in the cohorts' neuropsychological assessment battery (52). Likewise, scientists preparing for the Alzheimer's Disease Cooperative Study (ADCS) "A4" trial for cognitively healthy individuals with presence of fibrillar amyloid burden have undertaken a similar effort using datasets from other cohorts and have produced results comparable to those reported here (6). The sample size and power estimates reported herein serve as a guide in comparing sensitivity and power of the various measures and are not directly applicable to future trials given that they were calculated in a sample of individuals all of whom progressed to MCI or dementia. We acknowledge that there are multiple analytical methods for deriving composites and there are limitations to the approach used,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

including but not limited to the assumption of a simple model structure and that all variables are measured without error(54). Additional efforts under the auspices of the API are underway using other analysis methods, such as partial least squares, and will be reported separately. It should be noted that this study capitalized on longitudinal data from a cohort of cognitively unimpaired research participants who subsequent progressed to the clinical stages of AD in an effort to provide a more sensitive measure of the cognitive decline associated with AD and a cognitive endpoint could possibly qualify for use in preclinical AD trials. Since antemortem brain imaging or CSF biomarker measurements were not available in the Rush cohorts, additional studies may be helpful in clarifying the composite cognitive test score's power to track declines and evaluate preclinical AD treatments in cognitively unimpaired subjects who meet the recently proposed research criteria for preclinical AD (55). We are also confirming the generalizability and power of the composite cognitive endpoint in other cohorts followed to progression of cognitive impairment, and to estimate the statistical power in different at-risk groups (e.g., *APOE* ε4 homozygotes or heterozygotes, different age groups, older adults with or without biomarker evidence of AD). Indeed, it may be that certain cognitive assessments or combinations of assessments are more sensitive to cognitive decline depending on the time frame prior to (or following) diagnosis. Similarly, there may be issues of generalizability to other populations. For instance, the Rush cohorts are not population-based samples, and all participants in these cohort studies are expected to undergo annual assessments and some are required consent to brain donation upon death. That said, given expectations of participants in clinical trials (e.g., frequent in-person visits with cognitive testing, biomarker assessments), the composite developed using data from the Rush cohorts is likely to be sensitive to tracking preclinical decline in individuals who would enroll in a preclinical AD trial. However, it remains unknown whether the composite is sensitive to detecting a treatment effect.

In summary, we examined longitudinal cognitive data from three well-characterized cohorts and conducted a search of every combination of cognitive assessments to identify the optimal combination that is sensitive to tracking preclinical AD decline in the two and five years prior to diagnosis of MCI or dementia, while controlling for aging and practice effects in individuals who remained cognitively normal during the same time interval. This combination was also selected to include robust items and to represent domains important in early disease. The approach allowed us to empirically identify an API composite cognitive endpoint that consists of multiple cognitive domains and is sensitive to preclinical cognitive decline associated with AD. A similar composite cognitive endpoint with even greater statistical power was independently derived in cognitively normal, ADAD mutation carriers and is being used as the primary endpoint in the first API preclinical treatment trial set to begin in 2013 (52). As a result of these efforts, other preclinical trial investigators have extended the API composite cognitive test score development strategy for use in their planned trials and studies.

Acknowledgments

This work was supported by the National Institute on Aging (R01AG031581 and P30AG19610 to EMR; P30AG10161, R01AG15819, and R01AG17917 to DAB; and R01AG22018 to LLB), the state of Arizona, and contributions from the Banner Alzheimer's Foundation and the Nomis Foundation. We thank Dr. Richard Caselli, Laura Jakimovich, Carolyn Langlois, and Stephanie Stanworth for their assistance.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Reference List

1. Alzheimer's Association. 2013 Alzheimer's Disease Facts and Figures. Vol. 9. Alzheimer's & Dementia; 2013.
2. Brookmeyer R, Gray S, Kawas C. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. Am J Public Health. 1998 Sep; 88(9):1337–42. [PubMed: 9736873]
3. Reiman EM, Langbaum JBS, Tariot PN. Alzheimer's Prevention Initiative: a proposal to evaluate presymptomatic treatments as quickly as possible. Biomarkers in Medicine. 2010 Feb 1; 4(1):3–14. [PubMed: 20383319]
4. Reiman EM, Langbaum JBS, Fleisher AS, Caselli RJ, Chen K, Ayutyanont N, Quiroz YT, Kosik KS, Lopera F, Tariot PN. Alzheimer's Prevention Initiative: a plan to accelerate the evaluation of presymptomatic treatments. J Alzheimers Dis. 2011; 25(Suppl 3):293–301.
5. Mills SM, Mallmann J, Santacruz AM, Fuqua A, Carril M, Aisen PS, Althage MC, Belyew S, Benzinger TL, Brooks WS, Buckles VD, Cairns NJ, Clifford D, Danek A, Fagan AM, Farlow M, Fox N, Ghetti B, Goate AM, Heinrichs D, Hornbeck R, Jack C, Jucker M, Klunk WE, Marcus DS, et al. Preclinical trials in autosomal dominant AD: implementation of the DIAN-TU trial. Rev Neurol (Paris). 2013 Oct; 169(10):737–43. [PubMed: 24016464]
6. Sperling R, Donohue M, Aisen P. The A4 trial: Anti-amyloid treatment of asymptomatic Alzheimer's disease. Alzheimers Dement. 2012 Jul 1; 8(4):425–6.
7. Roses AD, Lutz MW, Amrine-Madsen H, Saunders AM, Crenshaw DG, Sundseth SS, Huentelman MJ, Welsh-Bohmer KA, Reiman EM. A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. Pharmacogenomics J. 2010 Oct; 10(5):375–84. [PubMed: 20029386]
8. Langbaum JB, Fleisher AS, Chen K, Ayutyanont N, Lopera F, Quiroz YT, Caselli RJ, Tariot PN, Reiman EM. Ushering in the study and treatment of preclinical Alzheimer disease. Nat Rev Neurol. 2013 Jul; 9(7):371–81. [PubMed: 23752908]
9. ADAPT Research Group. Cognitive Function Over Time in the Alzheimer's Disease Anti-inflammatory Prevention Trial (ADAPT): Results of a Randomized, Controlled Trial of Naproxen and Celecoxib. Archives of Neurology. 2008 Jul 1; 65(7):896–905. [PubMed: 18474729]
10. DeKosky ST, Williamson JD, Fitzpatrick AL, Kronmal RA, Ives DG, Saxton JA, Lopez OL, Burke G, Carlson MC, Fried LP, Kuller LH, Robbins JA, Tracy RP, Woolard NF, Dunn L, Snitz BE, Nahin RL, Furberg CD. Ginkgo biloba for prevention of dementia: a randomized controlled trial. JAMA. 2008 Nov 19; 300(19):2253–62. [PubMed: 19017911]
11. Silverberg NB, Ryan LM, Carrillo MC, Sperling R, Petersen RC, Posner HB, Snyder PJ, Hilsabeck R, Gallagher M, Raber J, Rizzo A, Possin K, King J, Kaye J, Ott BR, Albert MS, Wagster MV, Schinka JA, Cullum CM, Farias ST, Balota D, Rao S, Loewenstein D, Budson AE, Brandt J, et al. Assessment of cognition in early dementia. Alzheimer's and Dementia. 2011 May; 7(3):e60–e76.
12. Cano SJ, Posner HB, Moline ML, Hurt SW, Swartz J, Hsu T, Hobart JC. The ADAS-cog in Alzheimer's disease clinical trials: psychometric evaluation of the sum and its parts. J Neurol Neurosurg Psychiatry. 2010 Dec; 81(12):1363–8. [PubMed: 20881017]
13. Aisen PS, Andrieu S, Sampaio C, Carrillo 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, Bednar MM, Touchon J, Vellas B. Report of the task force on designing clinical trials in early (predementia) AD. Neurology. 2011 Jan 18; 76(3): 280–6. [PubMed: 21178097]
14. Elias MF, Beiser A, Wolf PA, Au R, White RF, D'Agostino RB. The preclinical phase of alzheimer disease: A 22-year prospective study of the Framingham Cohort. Arch Neurol. 2000 Jun; 57(6):808–13. [PubMed: 10867777]
15. Saxton J, Lopez OL, Ratcliff G, Dulberg C, Fried LP, Carlson MC, Newman AB, Kuller L. Preclinical Alzheimer disease: neuropsychological test performance 1.5 to 8 years prior to onset. Neurology. 2004 Dec 28; 63(12):2341–7. [PubMed: 15623697]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

16. Johnson DK, Storandt M, Morris JC, Galvin JE. Longitudinal Study of the Transition From Healthy Aging to Alzheimer Disease. Archives of Neurology. 2009 Oct 1; 66(10):1254–9. [PubMed: 19822781]
17. Wilson RS, Leurgans SE, Boyle PA, Bennett DA. Cognitive Decline in Prodromal Alzheimer Disease and Mild Cognitive Impairment. Archives of Neurology. 2011 Mar 1; 68(3):351–6. [PubMed: 21403020]
18. Amieva H, Le GM, Millet X, Orgogozo JM, Peres K, Barberger-Gateau P, Jacqmin-Gadda H, Dartigues JF. Prodromal Alzheimer's disease: successive emergence of the clinical symptoms. Ann Neurol. 2008 Nov; 64(5):492–8. [PubMed: 19067364]
19. Grober E, Hall CB, Lipton RB, Zonderman AB, Resnick SM, Kawas C. Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. J Int Neuropsychol Soc. 2008 Mar; 14(2):266–78. [PubMed: 18282324]
20. Small BJ, Backman L. Longitudinal trajectories of cognitive change in preclinical Alzheimer's disease: a growth mixture modeling analysis. Cortex. 2007 Oct; 43(7):826–34. [PubMed: 17941341]
21. Stonnington CM, Chen K, Lee W, Locke DE, Dueck AC, Liu X, Roontiva A, Fleisher AS, Caselli RJ, Reiman EM. Fibrillar amyloid correlates of preclinical cognitive decline. Alzheimers Dement. 2013 Apr 11.
22. Pike KE, Ellis KA, Villemagne VL, Good N, Chetelat G, Ames D, Szoeke C, Laws SM, Verdile G, Martins RN, Masters CL, Rowe CC. Cognition and beta-amyloid in preclinical Alzheimer's disease: data from the AIBL study. Neuropsychologia. 2011 Jul; 49(9):2384–90. [PubMed: 21529702]
23. Ellis KA, Lim YY, Harrington K, Ames D, Bush AI, Darby D, Martins RN, Masters CL, Rowe CC, Savage G, Szoeke C, Villemagne VL, Maruff P. Decline in Cognitive Function over 18 Months in Healthy Older Adults with High Amyloid-beta. J Alzheimers Dis. 2013 Jan 8.
24. Lim YY, Maruff P, Pietrzak RH, Ames D, Ellis KA, Harrington K, Lautenschlager NT, Szoeke C, Martins RN, Masters CL, Villemagne VL, Rowe CC. Effect of amyloid on memory and non-memory decline from preclinical to clinical Alzheimer's disease. Brain. 2013 Oct 30.
25. Sperling RA, Johnson KA, Doraiswamy PM, Reiman EM, Fleisher AS, Sabbagh MN, Sadowsky CH, Carpenter A, Davis MD, Lu M, Flitter M, Joshi AD, Clark CM, Grundman M, Mintun MA, Skovronsky DM, Pontecorvo MJ. Amyloid deposition detected with florbetapir F 18 ((18)F-AV-45) is related to lower episodic memory performance in clinically normal older individuals. Neurobiol Aging. 2013 Mar; 34(3):822–31. [PubMed: 22878163]
26. Caselli RJ, Dueck AC, Osborne D, Sabbagh MN, Connor DJ, Ahern GL, Baxter LC, Rapcsak SZ, Shi J, Woodruff BK, Locke DEC, Hoffman-Snyder C, Alexander GE, Rademakers R, Reiman EM. Longitudinal modeling of age-related memory decline and the APOE ε4 effect. N Engl J Med. 2009 Jul 16; 361(3):255–63. [PubMed: 19605830]
27. Crane PK, Carle A, Gibbons LE, Insel P, Mackin RS, Gross A, Jones RN, Mukherjee S, Curtis SM, Harvey D, Weiner M, Mungas D. Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). Brain Imaging Behav. 2012 Jul 11.
28. Gibbons LE, Carle AC, Mackin RS, Harvey D, Mukherjee S, Insel P, Curtis SM, Mungas D, Crane PK. A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. Brain Imaging Behav. 2012 May 27.
29. Dowling NM, Hermann B, La RA, Sager MA. Latent structure and factorial invariance of a neuropsychological test battery for the study of preclinical Alzheimer's disease. Neuropsychology. 2010 Nov; 24(6):742–56. [PubMed: 21038965]
30. Hayden KM, Kuchibhatla M, Romero HR, Plassman BL, Burke JR, Browndyke JN, Welsh-Bohmer KA. Pre-clinical Cognitive Phenotypes for Alzheimer Disease: A Latent Profile Approach. Am J Geriatr Psychiatry. 2013 Sep 27.
31. Hendrix SB. Measuring clinical progression in MCI and pre-MCI populations: enrichment and optimizing clinical outcomes over time. Alzheimers Res Ther. 2012 Jul 13.4(4):24. [PubMed: 22805433]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

32. Xiong C, van Belle G, Chen K, Tian L, Luo J, Gao F, Yan Y, Chen L, Morris JC, Crane P. Combining Multiple Markers to Improve the Longitudinal Rate of Progression: Application to Clinical Trials on the Early Stage of Alzheimer's Disease. Statistics in Biopharmaceutical Research. 2013 Feb 1; 5(1):54–66.
33. Balsis S, Unger AA, Benge JF, Geraci L, Doody RS. Gaining precision on the Alzheimer's Disease Assessment Scale-cognitive: a comparison of item response theory-based scores and total scores. Alzheimers Dement. 2012 Jul; 8(4):288–94. [PubMed: 22465173]
34. Hobart J, Cano S, Posner H, Selnes O, Stern Y, Thomas R, Zajicek J. Putting the Alzheimer's cognitive test to the test II: Rasch Measurement Theory. Alzheimers Dement. 2013 Feb; 9(1 Suppl):S10–S20. [PubMed: 23253779]
35. Shankle WR, Romney AK, Hara J, Fortier D, Dick MB, Chen JM, Chan T, Sun X. Methods to improve the detection of mild cognitive impairment. Proc Natl Acad Sci U S A. 2005 Mar 29; 102(13):4919–24. [PubMed: 15781874]
36. Raghavan N, Samtani MN, Farnum M, Yang E, Novak G, Grundman M, Narayan V, DiBernardo A. The ADAS-Cog revisited: Novel composite scales based on ADAS-Cog to improve efficiency in MCI and early AD trials. Alzheimers Dement. In press.
37. Barnes LL, Shah RC, Aggarwal NT, Bennett DA, Schneider JA. The Minority Aging Research Study: ongoing efforts to obtain brain donation in African Americans without dementia. Curr Alzheimer Res. 2012 Jul; 9(6):734–45. [PubMed: 22471868]
38. Bennett DA, Schneider JA, Buchman AS, Barnes LL, Boyle PA, Wilson RS. Overview and findings from the rush Memory and Aging Project. Curr Alzheimer Res. 2012 Jul; 9(6):646–63. [PubMed: 22471867]
39. Bennett DA, Schneider JA, Arvanitakis Z, Wilson RS. Overview and findings from the religious orders study. Curr Alzheimer Res. 2012 Jul; 9(6):628–45. [PubMed: 22471860]
40. Wilson RS, Beckett LA, Barnes LL, Schneider JA, Bach J, Evans DA, Bennett DA. Individual differences in rates of change in cognitive abilities of older persons. Psychol Aging. 2002 Jun; 17(2):179–93. [PubMed: 12061405]
41. Wilson RS, Barnes LL, Krueger KR, Hoganson G, Bienias JL, Bennett DA. Early and late life cognitive activity and cognitive systems in old age. J Int Neuropsychol Soc. 2005 Jul; 11(4):400–7. [PubMed: 16209420]
42. Bennett DA, Schneider JA, Buchman AS, Mendes de LC, Bienias JL, Wilson RS. The Rush Memory and Aging Project: study design and baseline characteristics of the study cohort. Neuroepidemiology. 2005; 25(4):163–75. [PubMed: 16103727]
43. Bennett DA, Wilson RS, Schneider JA, Evans DA, Beckett LA, Aggarwal NT, Barnes LL, Fox JH, Bach J. Natural history of mild cognitive impairment in older persons. Neurology. 2002 Jul 23; 59(2):198–205. [PubMed: 12136057]
44. Bennett DA, Schneider JA, Aggarwal NT, Arvanitakis Z, Shah RC, Kelly JF, Fox JH, Cochran EJ, Arends D, Treinkman AD, Wilson RS. Decision rules guiding the clinical diagnosis of Alzheimer's disease in two community-based cohort studies compared to standard practice in a clinic-based cohort study. Neuroepidemiology. 2006; 27(3):169–76. [PubMed: 17035694]
45. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984 Jul; 34(7):939–44. [PubMed: 6610841]
46. Barnes LL, Schneider JA, Boyle PA, Bienias JL, Bennett DA. Memory complaints are related to Alzheimer disease pathology in older persons. Neurology. 2006 Nov 14; 67(9):1581–5. [PubMed: 17101887]
47. Schneider JA, Arvanitakis Z, Leurgans SE, Bennett DA. The neuropathology of probable Alzheimer disease and mild cognitive impairment. Ann Neurol. 2009 Aug; 66(2):200–8. [PubMed: 19743450]
48. Bennett DA, Schneider JA, Bienias JL, Evans DA, Wilson RS. Mild cognitive impairment is related to Alzheimer disease pathology and cerebral infarctions. Neurology. 2005 Mar 8; 64(5): 834–41. [PubMed: 15753419]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

49. McEvoy LK, Edland SD, Holland D, Hagler DJ Jr, Roddey JC, Fennema-Notestine C, Salmon DP, Koyama AK, Aisen PS, Brewer JB, Dale AM. Neuroimaging enrichment strategy for secondary prevention trials in Alzheimer disease. Alzheimer Dis Assoc Disord. 2010 Jul; 24(3):269–77. [PubMed: 20683184]
50. Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, Harvey D, Jack CR, Jagust W, Liu E, Morris JC, Petersen RC, Saykin AJ, Schmidt ME, Shaw L, Siuciak JA, Soares H, Toga AW, Trojanowski JQ. The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. Alzheimers Dement. 2012 Feb; 8(1 Suppl):S1–68. [PubMed: 22047634]
51. Food and Drug Administration. Draft guidance. U.S. Department of Health and Human Services, Food and Drug Administratino, Center for Drug Evaluation and Research (CDER); 2013. Guidance for industry Alzheimer's disease: Developing drugs for the treatment of early stage disease.
52. Ayutyanont, N.; Langbaum, JBS.; Hendrix, SB.; Chen, K.; Fleisher, AS.; Friesenhahn, M.; Ward, M.; Aguirre, C.; Acosta-Baena, N.; Madrigal, L.; Munoz, C.; Tirado, V.; Moreno, S.; Tariot, PN.; Lopera, F.; Reiman, EM. The Alzheimer's Prevention Initiative composite cognitive test score: sample size estimates for the evaluation of preclinical Alzheimer's disease treatments in presenilin 1 E280A mutation carriers. submitted
53. Riley KP, Jicha GA, Davis D, Abner EL, Cooper GE, Stiles N, Smith CD, Kryscio RJ, Nelson PT, Van Eldik LJ, Schmitt FA. Prediction of preclinical Alzheimer's disease: longitudinal rates of change in cognition. J Alzheimers Dis. 2011; 25(4):707–17. [PubMed: 21498903]
54. Haenlein M, Kaplan AM. A beginner's guide to partial least squares analysis. Understanding Statistics. 2004; 3(4):283–97.
55. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging and the Alzheimer's Association workgroup. Alzheimers Dement. 2011 May; 7(3):280–92. [PubMed: 21514248]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Research In Context

Systematic Review

A PubMed search was conducted to identify relevant studies from 1980-present examining cognitive decline in preclinical AD, as well as methodological approaches for development of composite endpoints.

Interpretation

The analyses identified a composite cognitive test score with optimal statistical power to track preclinical AD decline compared to the most sensitive individual neuropsychological test score. The composite is similar to one independently identified for tracking preclinical decline in autosomal dominant AD and is well-suited for use as a primary endpoint in AD prevention treatment trials.

Future Directions

From a regulatory perspective it may be important for the field to reach consensus regarding the optimal approaches for identification and use of composite endpoints, particularly given the growing interest in preclinical AD treatment trials.

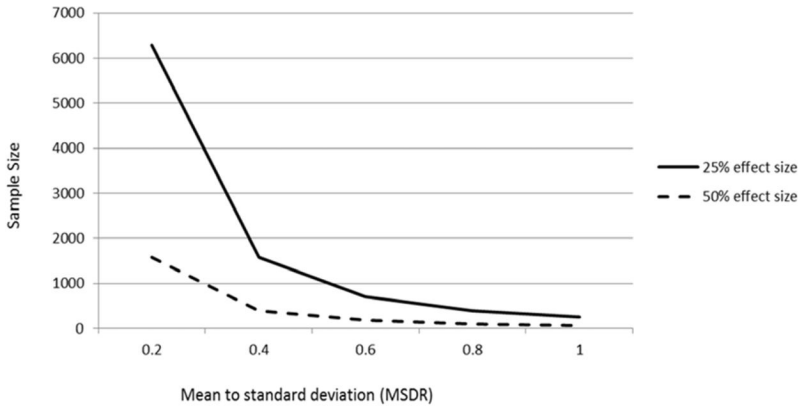

Figure 1.

Estimated sample size per group required to detect a 25% or 50% treatment effect with two-tailed p = 0.05 and 80% power in a randomized, placebo-controlled trial.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Alzheimers Dement. Author manuscript; available in PMC 2015 November 01.

Table 1

Demographic information for those who progressed to MCI or dementia (converters) compared to those who remained cognitively healthy (non-converters) in the two and five year MSDR analyses

	2-year MSDR analysis			**5-year MSDR analysis**		
	Non-Converter (n = 831)	**Converter (n = 528)**	**Test statistic, p-value**	**Non-Converter (n = 413)**	**Converter (n = 213)**	**Test statistic, p-value**
Age, years	74.1 (7.1)	77.6 (6.9)	t (1357) = 8.9, p <0.0001	73.3 (6.4)	76.0 (6.8)	t (624) = 4.8, p <0.0001
Sex, %						
Female	72.8	71.0	χ^2 (1)= 0.5, p = 0.48	72.6	66.7	χ^2 (1)= 2.4, p = 0.12
Male	27.2	29.0		27.4	33.3	
Educational level, years	16.2 (3.8)	16.5 (3.7)	t (1357) = 1.8, p = 0.07	16.8 (3.7)	17.2 (3.7)	t (624) = 1.5, p =0.14
APOE ε4 carrier status, %						
ε4 carriers	17.0	19.5	χ^2 (1)= 0.4, p =0.53	16.9	21.6	χ^2 (1)= 1.1, p =0.29
ε4 n on-carriers	69.7	73.1		72.9	74.2	
Missing	13.3	7.4		10.2	4.2	
Race, %						
Caucasian	78.3	89.4	χ^2 (3)= 28.8, p <0.0001	93.2	98.6	χ^2 (2)= 8.7, p = 0.01
Black, African American	21.2	10		6.3	1.4	
Native American, Indian	0.2	0.2		0	0	
Asian or Pacific Islander	0.2	0.4		0.5	0	
Missing	0	0		0	0	

Values are mean(standard deviation) or percentage.

p-values were calculated using t-test to compare participant groups for age and educational level, and with chi-square tests to compare the groups for sex, *APOE* ε4 carrier status and race.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Five-Year MSDRs for Individual Cognitive Assessment Test Items Considered for the API Composite Cognitive Test Score

Cognitive assessment	Domain	Unadjusted MSDR	Adjusted MSDR
Boston Naming Test (15 item) *	Language / Semantic memory	0.36	0.305
Category fluency – Animals	Language / Semantic memory	0.645	0.44
Category fluency – Fruits/Vegetables *	Language / Semantic memory	0.825	0.61
CERAD Word list recall (Immediate)	Episodic memory	0.545	0.51
CERAD Word list memory (Delayed recall)	Episodic memory	0.635	0.52
CERAD Word List Recognition	Episodic memory	0.415	0.4
Complex Ideational Material	Auditory comprehension	0.3	0.285
Digit Ordering	Working memory	0.39	0.24
Digit Span - forward	Working memory	0.35	0.175
Digit Span - backward	Working memory	0.42	0.175
East Boston Naming Test, Immediate recall (Memory I)*	Episodic memory	0.44	0.485
East Boston Naming Test, Delayed recall (Memory II)	Episodic memory	0.54	0.435
Judgment of Line Orientation	Visuospatial	0.375	0.295
Logical Memory Ia (Immediate)	Episodic memory	0.405	0.55
Logical Memory IIa (Delayed) *	Episodic memory	0.455	0.64
Mini-Mental Status Examination (MMSE) - Total	General / Global Cognition	0.665	0.545
MMSE – Orientation to Time *	Orientation	0.555	0.465
MMSE – Orientation to Place	Orientation	0.465	0.44
MMSE – Registration	Working memory	0.2	0.145
MMSE – Attention and Concentration	Attention and Concentration	0.225	0.185
MMSE – Recall	Episodic memory	0.145	0.11
MMSE – Language	Language	0.17	0.04
National Adult Reading Test (NART) – 10 items	General / Global Cognition	0.155	0.17
Number Comparison Test	Perceptual speed	0.56	0.355
Ravens Progressive Matrices – 16 items	Visuospatial / working memory	0.54	0.385
Ravens Progressive Matrices Subset – 9 items *	Visuospatial / working memory	0.5	0.43
Symbol Digit Modalities *	Perceptual speed	0.71	0.385
Wide Range Achievement Test (WRAT) – 15 items	General / Global Cognition	0.17	0.12

CERAD = Consortium to Establish a Registry for Alzheimer's Disease

* Item included in the composite cognitive test score