HHS Public Access

Author manuscript

Neuroepidemiology. Author manuscript; available in PMC 2015 November 05.

Published in final edited form as:

Neuroepidemiology. 2014 ; 43(2): 131–139. doi:10.1159/000365733.

Subsets of a Large Cognitive Battery Better Power Clinical Trials on Early Stage Alzheimer Disease

Chengjie Xiong, PhD[1,6], **Hua Weng, PhD**[1], **David A. Bennett, MD**[2], **Patricia A. Boyle, PhD**[2], **Raj C. Shah, MD**[2], **Scot Fague, MS**[1], **Charles B. Hall, PhD**[3], **Richard B. Lipton, MD**[4], and **John C. Morris, MD**[5,6]

[1]Division of Biostatistics, Washington University in St. Louis, St. Louis, MO

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

[3]Department of Epidemiology & Population Health, Albert Einstein College of Medicine, New York, NY

[4]The Saul R. Korey Department of Neurology, Albert Einstein College of Medicine, New York, NY

[5]Departments of Pathology and Immunology, Department of Neurology, Washington University in St. Louis, St. Louis, MO

[6]Knight Alzheimer's Disease Research Center, Washington University in St. Louis, St. Louis, MO

Abstract

Background/Aims—Cognitive batteries routinely used by the Alzheimer disease (AD) research community may contain items uninformative for tracking disease progression to power clinical trials on early stage AD. We aim to identify subsets of the most informative items from an existing cognitive battery for better powering clinical trials on early AD.

Methods—Longitudinal change in item scores from the battery was associated with the onset of Mild Cognitive Impairment (MCI) in 1513 elderly individuals. Items whose longitudinal changes were correlated with the onset of MCI were selected as informative for tracking the early cognitive progression.

Results—226 items in the battery were annually assessed over a follow-up of up to 13 years. Changes of item scores over time from 187 items were significantly correlated with the onset of MCI. For clinical trials on preclinical AD and on MCI, informative items permit smaller or similar sample sizes as compared to the entire battery, whereas uninformative items require much larger sample sizes.

Conclusions—Longitudinal changes in item scores from about 17% of items in the cognitive battery are uninformative for tracking early disease progression. Clinical trials on early AD can be better powered using informative items rather than the entire battery.

Corresponding Author: Chengjie Xiong, Ph.D., Division of Biostatistics, Campus Box 8067, Washington University in St. Louis, St. Louis, MO 63110, U.S.A, Phone: (314)3623635, Fax: (314)362-2693, chengjie@wubios.wustl.edu.

Introduction

Alzheimer's disease (AD) is an age-related neurodegenerative disorder that results in progressive cognitive impairment and death. Accumulating research suggests that the neurodegenerative processes associated with AD begin years prior to the symptomatic onset of AD when the disease is clinically at the early prodromal stage or a latent stage.[1–3] Many recent clinicopathologic studies also have demonstrated that asymptomatic individuals can manifest the neuropathological changes of AD, notably senile plaques and neurofibrillary tangles. [4–6] These observations, coupled with the fact that there are currently no pharmaceutical treatments that reverse the pathological processes of AD, have led to a major paradigm shift in the search of efficacious treatments of AD, that is, the focus of modern AD clinical trials now is on individuals at the earliest clinical stages, such as Mild Cognitive Impairment [7] (MCI) and/or very mild dementia (i.e., a Clinical Dementia Rating [8] (CDR) of 0.5), or even the preclinical stage [9] prior to the substantial development of clinical symptoms as these may be the groups of individuals in which targeted therapies may have the greatest chance of preserving brain function.

The paradigm shift in clinical trials on AD subsequently has led to three major inter-related biomedical decisions that must be made by investigators in designing modern clinical trials at the early stages of AD: the cognitive outcome measure, sample size, and disease duration. Because cognitive batteries routinely used by the AD research community have been traditionally designed to track the disease progression after symptomatic onset and to identify cases of fully developed AD dementia in comparison to normal controls, they often show significant ceiling and floor effects. Therefore, the current cognitive batteries may contain items that are neither sensitive nor specific for tracking early stage disease progression. As a result, they only exhibit subtle changes during the very early stage or the preclinical stage of AD. Several recent randomized clinical trials (RCTs) using existing instruments (e.g., the Alzheimer's Disease Assessment Scale-Cognitive subscale[10]) failed to detect significant decline in placebo groups with MCI. Especially for RCTs on early stage or preclinical stage of AD, the lack of progression on existing cognitive outcomes has become an important challenge to the feasibility of such trials because of the need for a large number of individuals to be followed over many years to allow meaningful statistical conclusions to be drawn.[11–13] Large, long-duration RCTs are time-consuming and prohibitively costly. Although emerging cerebrospinal fluid (CSF) biomarkers and neuroimaging markers[14–18] have been reported to show early changes in AD progression, recently revised FDA guidelines for RCTs on early stage AD mandate that treatments of AD be only approved if they demonstrate cognitive and functional benefits (FDA Guidance for Industry Alzheimer's Disease: Developing Drugs for the Treatment of Early Stage Disease:[19] http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM338287.pdf).

The objective of this manuscript is to provide better cognitive outcomes for future RCTs on early stage AD. We posit that subsets of items, identified from a large cognitive battery that are shown to be most informative to early disease progression, can better power clinical trials on early stage AD than the other uninformative items and even the entire battery by reducing the sample sizes and improving the efficiency. In order to identify the most

informative items, we focused on tracking the disease progression (not necessarily for the prediction of the disease), and analyzed the longitudinal changes of item scores of individual items from the cognitive battery administered to a large longitudinal cohort study, and correlated the time to the onset of MCI with the time of the item score changes (i.e., from endorsement to non-endorsement of the item over time). Finally, we estimated the sample sizes to adequately power future RCTs on MCI and on preclinical AD using the composite cognitive scores from the items identified as the most informative for early disease progression, and compared them to the composite scores derived from the entire battery as well as the items not identified as informative.

Materials and Methods

The longitudinal cognitive database of the Rush Memory and Aging Project (MAP), a longitudinal clinical-pathologic cohort study of aging and dementia, was analyzed first for identifying the most informative items to track cognitive progression of early stage AD and then for powering clinical trials on early stage AD.

Participants

Participants were individuals from the MAP, an ongoing longitudinal, community study of common chronic conditions of old age that began in 1997. Participants were recruited primarily from continuing-care retirement communities throughout the Chicago metropolitan area because the ability to maintain high rates of clinical follow-up and to obtain autopsy is key to the MAP mission. This was supplemented by recruitment at senior and subsidized housing, churches, and social service agencies to ensure a range of socioeconomic status, race, and ethnicity. A requirement for study entry is that participants understood what was involved in the study in order to sign an Informed Consent and agreed to donate their brains, spinal cords, nerves and muscles at the time of death. [20] A total of N=1513 elderly individuals who were either cognitively normal or with MCI or AD at baseline were available for our analyses as of September 21, 2012. 1037 were cognitively normal, 402 had MCI and 74 had AD at baseline. Baseline characteristics of the participants are shown in Table 1.

Standard Protocol Approval and Patient Consents

Written informed consent was obtained from all study participants. The study was approved by the Institutional Review Boards of Rush University Medical Center and the Washington University School of Medicine.

Cognitive Assessment and Clinical Diagnosis

Cognition was assessed annually with a comprehensive battery testing cognitive domains commonly affected by aging and AD: episodic memory, semantic memory, working memory, perceptual speed, and visual–spatial ability. Details of the cognitive function tests have been reported previously. [21–22] In brief, 20 cognitive tests were administered annually. Seven were episodic memory measures: Word List Memory, Recall, and Recognition [23] and immediate and delayed recall of Story A from Logical Memory of the Wechsler Memory Scale–Revised [24] and of the East Boston Story [25–26] . Semantic memory was assessed with a

15-item version [21] of the Boston Naming Test, [27] Verbal Fluency, [23,26] and a 15-item version [26] of the National Adult Reading Test . [28] Working memory tests included Digit Span Forward and Digit Span Backward [24] and digit ordering . [26,29] Four measures of perceptual speed were administered: the oral version of the Symbol Digit Modalities Test , [30] Number Comparison , [26,31] and two measures from a modified version [22] of the Stroop Neuropsychological Screening Test: [32] number of color names correctly read in 30s minus the number of errors and number of colors correctly named in 30s minus the number of errors. Visuospatial ability was assessed with a 15-item version of Judgment of Line Orientation [33] and a 16-item version of Standard Progressive Matrices . [34] In addition, the Mini-Mental State Examination (MMSE [35]) also was used to serve as a brief measure of cognitive function. To minimize floor and ceiling artifacts and other sources of measurement error, a global composite measure of cognition that was previously reported [21–22] and based on all 20 tests including MMSE and measures of episodic memory, semantic memory, working memory, perceptual speed, and visuospatial ability was created by converting raw scores to z scores, using the baseline mean and SD across the entire cohort, and averaging the z scores. Further information about the individual tests and the derivation of the composite measure was described elsewhere. [21–22] Cognitive testing was scored by computer and reviewed by a neuropsychologist to determine cognitive impairment. Participants were then evaluated by a clinician for a medical history and a neurologic examination to diagnose AD using National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria. [36] The diagnosis of dementia required a history of cognitive decline and impairment in at least 2 cognitive domains, and MCI (all types) required the presence of cognitive impairment in the absence of dementia. [20]

Types of Item Scores

Individual items were scored as binary for most cognitive tests except for five (Color Name, Logical Memory IA and IIA, Verbal Fluency, and Words Correctly Read) which could not be reasonably decomposed into meaningful binary items and therefore were treated as single-item tests with the original count score. Table 2 presents the number of individual items as well as baseline summary statistics for each of the 20 cognitive tests (including MMSE). A few items from several tests were not originally scored as binary (e.g., one item in MMSE has an original score of 0 to 5), such items were treated as binary in the item analyses using the perfect score vs. others. Sensitivity analyses were also done with other possible cutoffs for these items.

Other Covariates

Demographics such as age, sex and years of education were recorded at the study entry. *APOE* genotyping was done and individuals were dichotomized into those with at least 1 or more copies of the E4 allele (E4 positive) vs. those without an E4 copy (E4 negative).

Statistical Analysis

A two-step procedure was implemented to select the most informative items with binary score (i.e., endorsement and non-endorsement of the item, oriented in the way that non-

endorsement always indicates problems with cognition) from the entire cognitive battery for better tracking the disease progression and designing future RCTs at the early stages. The first step examined the longitudinal changes of scores for each individual item and tested whether the item score changes over time were associated with the onset of MCI. Specifically, for each individual, we first computed the age at onset of MCI. If an individual never developed MCI or AD during the entire follow-up, the age at onset was considered as right-censored. If a subject was already classified as MCI or AD at the baseline, the age at onset of MCI was considered interval censored between 0 and the age at baseline. For each individual and each binary item, we then computed the age of non-endorsement defined as the age when the item was incorrectly endorsed. Because of the fluctuation of item level scores over time, the item-specific age of non-endorsement for each individual also was considered interval censored with the left side of the interval as the first age in the follow-up when the item was endorsed incorrectly and the right side as the age at the first occurrence of non-endorsement after which the item remained incorrectly endorsed over time. [37] If an individual already incorrectly endorsed the item at baseline, the left side of the age of non-endorsement was defined as 0. If an individual never incorrectly endorsed the item during the entire follow-up, the age of non-endorsement was considered as right-censored at the age of last assessment. For each item, the item-specific age of non-endorsement and the age at onset of MCI were then correlated across the entire cohort including subjects who were normal, MCI, or AD at baseline. The correlation was estimated by a Kendall's coefficient of concordance through a bivariate smooth estimate of the joint density on the logarithms of the two time scales that was obtained using a mixture of Gaussian densities fixed on a grid with weights determined by a penalized likelihood approach. [37–38] Items with a significant correlation ($p<0.05$) were identified as informative for tracking early disease progression.

Five tests in the cognitive battery could not be reasonably decomposed as binary items because their scores are counts that can be any non-negative integers: Color Name, Logical Memory IA and IIA, Verbal Fluency (with 2 items, Animal and Fruit), and Words Correctly Read. Preliminary analyses similar to those used for the binary items were conducted using different cutoffs, suggesting that all five tests have reasonable sensitivity and specificity for tracking early disease changes. Thus, they (a total of 6 items) were always included as informative items, using their original scores (not the dichotomized scores).

Because for each item, the correlation between the age of non-endorsement and the age at onset of MCI does not indicate whether the age of non-endorsement occurred earlier or later than the age at onset of MCI, the second step of our item selection procedure compared the center of the interval censored age of non-endorsement and age at onset of MCI, and identified the items whose age of non-endorsement occurred later than the age at onset MCI for at least more than half (e.g., >51%) of individuals in the entire cohort. This subset of the informative items was used for tracking the disease progression and designing RCTs on individuals who were already MCI at baseline.

Power Analysis

In traditional prevention trials, it is common to randomize high-risk subjects to active drug or placebo and analyze time to the onset of the disease as the primary efficacy endpoint with

a standard Cox proportional hazards model (PHM). [39] Recently, for modern RCTs either on 'preclinical AD' or on MCI, it has been suggested that the PHM approach is subject to a loss of power to detect a treatment effect, in comparison to the approach with a linear mixed model for repeated measures (MMRM [40]) on a well defined cognitive composite. [41] Based on the informative items identified, three cognitive composite scores were computed similar to the way the global cognitive composite score was defined (see Cognitive Assessment and Clinical Diagnosis): one using all informative items, another using the subset of the informative items whose age of non-endorsement occurred later than the age at onset of MCI, and the other using the items that were not identified as informative (i.e., uninformative). For each composite, a mixed model for repeated measures (MMRM [40]) was then implemented to estimate the mean 2-year, 3-year, and 4-year change from the baseline as well as the relevant variance and covariance parameters from subjects who were considered 'preclinical AD' (in the absence of biomarker data, operationally defined as those who were cognitively normal at baseline but with at least 1 copy of APOE E4 allele) or MCI at baseline. To examine how the newly formed cognitive composites with informative items influence the sample sizes required for adequately powering future RCTs at the early stages, these estimates were used to further estimate the sample sizes for future RCTs on 'preclinical AD' or MCI with either 2-year, 3-year, or 4-year annual follow-ups using a standard normal test. [42] A power of 80% was assumed for all power analyses. For comparison purpose, similar power analyses were also conducted using the same subjects sample but the global cognitive composite derived from the entire cognitive battery as well as the cognitive composite derived from items not selected as informative. All statistical analyses were implemented in SAS. [43]

Results

A total of 226 items (including 220 items with binary data and 6 items from 5 tests with count data) from the cognitive battery were analyzed. Of the 1037 cognitively normal individuals, 404 developed MCI (all types) or AD during up to 12 years of follow-up. Individual items with very limited longitudinal data (i.e., those with less than 30% of the total annual assessments across the entire cohort for a total of 34 items) were excluded from the item analyses. The first step of the two-step procedure identified a total of 181 binary items whose age of non-endorsement is significantly correlated with the age at onset of MCI, resulting in a total of 187 informative items including the 6 items from 5 tests with count data. At the second step, a subset of 62 items were further identified for tracking disease progression on individuals who had MCI at baseline because their age of non-endorsement occurred later than the age at onset of MCI for at least more than half (e.g., >51%) of individuals in the entire cohort. The last column of Table 2 presents the number of informative items identified from each of the 20 cognitive tests in the battery. Table 2A in the Appendix lists all individual items that were found to be informative for tracking the early disease progression.

Using a consistent way of forming cognitive composites (i.e., averaging the z scores across multiple tests obtained by using the baseline mean and SD of each test), four cognitive composites were computed using the following: the 187 informative items as identified above, the 62 informative items whose age of non-endorsement occurred later than the age

at onset of MCI, all 20 cognitive tests in the battery, and the items not identified as informative by our analyses. Table 3 presents the estimated mean change and associated standard error (SE) of the cognitive composites from baseline among subjects with 'preclinical AD' at baseline, as well as among those with MCI at baseline. Results in Table 3 suggest that, for both 'preclinical AD' and MCI, the rate of change on the cognitive composite from the informative items is larger than that on the cognitive composite based on the entire battery, which in turn is larger than that on the cognitive composite based on items not identified as informative.

To assess the ability of the items identified as informative for tracking early disease progression to improve the design of modern RCTs on early stage AD, we considered two types of future two-arm RCT to test the cognitive efficacy of a novel therapeutic compound against a placebo on individuals who were 'preclinical AD' or MCI at baseline. The sample size ratio of the RCTs is assumed 1:1 between the two arms. The longitudinal follow-ups are assumed to be annual with a range of 2 to 4 years. The effect size (ES) of the novel treatment is assumed as a percentage of improvement on the change from baseline as compared to the placebo, the latter of which was estimated by a MMRM assuming a covariance structure of compound symmetry. We used change on the cognitive composite from the 187 informative items as the primary efficacy endpoint for the RCT on 'preclinical AD', and change on the cognitive composite from the subset of 62 informative items as the primary efficacy endpoint for the RCT on MCI because these items' age of non-endorsement occurred later than the age at onset of MCI. As indicated in Table 4, for the trial on 'preclinical AD' with a 2-years or 3-years annual follow-up, the use of 187 informative items provides smaller sample sizes than the use of entire cognitive battery, resulting in a reduction of sample size from 2% to 10% across a wide range of effect sizes. On the other hand, the sample size of the trial using items not identified as informative is at least 6 times of that using the informative items. For example, with a reasonable effect size of 40%, the RCT of 2-year follow-up can be adequately powered with a total of 3484 subjects using the informative items and 3887 subjects using the entire battery, and 22159 subjects using the items not identified as informative. For the RCT on 'preclinical AD' with a 4-years follow-up using informative items, the sample size is slightly lower than that of using the entire battery, but is only about one seventh of that using items not identified as informative. Very similar observations can be made for the RCTs on MCI. In comparison to the entire battery, the 62 informative items provide a sample size reduction of 11% to 30% in a RCT on MCI with a 2 or 3-years of follow-up, and only a slight increase (about 1%) in a RCT on MCI with a 4-years follow-up. On the other hand, the items not identified as informative require sample sizes that are 4 to 7 times that of the informative items.

Discussion

Individual items are the foundation for cognitive outcome measures used in RCTs on AD, and their test scores are inherently noisy when used to longitudinally track cognitive progression, especially at the early stages of disease. To our best knowledge, our item level longitudinal analyses represent the first comprehensive effort to associate the onset of early cognitive symptoms (i.e., MCI) with longitudinal change in item level scores from a comprehensive cognitive battery. We found that, out of a total of 226 items from a large

cognitive battery administered longitudinally in MAP on a large sample size of 1513 individuals, the longitudinal item score changes were associated with the onset of MCI for 187 items over an annual follow-up of up to 13 years. Of these, the item score changes for 62 items (i.e., from endorsement to non-endorsement of items) occurred after the onset of MCI. A total of 39 items (i.e., 17.26%) were found to be uninformative for tracking the cognitive progression at early disease stage, i.e., the longitudinal changes of item scores were not associated with the onset of MCI.

Although the conventional test scores from cognitive batteries used in many AD studies have been very successful in cross sectionally discriminating fully developed symptomatic AD from normal aging, they are less satisfactory in longitudinally tracking the early changes of AD when individuals are at the early or preclinical stage of the disease. In fact, when some of current cognitive tests are administered to individuals in the preclinical or early stage of the disease, the resulting data are subject to enormous ceiling and floor effects. Because cognitive items and tests with significant ceiling and floor effects have limited use in tracking longitudinal changes, they are unlikely to be correlated with the early disease progression, and therefore have limited power to predict whether or at what time point an individual will develop subtle sign of early changes which will eventually lead to the onset of MCI and AD.

There is currently a major conundrum in the search of effective treatments of AD. On the one hand, accumulating research evidence indicates that neurodegenerative processes associated with AD begin years prior to the symptomatic onset of AD, [4–6] suggesting that the optimum time window for treatment interventions is when the disease is clinically at the early prodromal stage or even the latent or preclinical stage. On the other hand, the lack of detection of progression by cognitive tests routinely used in current AD research makes the sample size for clinical trials on early disease stage or preclinical AD a formidable task to achieve. This bottleneck is primarily due to the lack of cognitive measures that can reliably detect the earliest possible cognitive changes of early stage AD in the presence of high inherent inter-individual variability during the progression of the disease. [44] This question is challenging because, by definition, longitudinal cognitive changes have to be subtle during the early stages of the disease.

In comparison to the entire battery, we found that the 187 informative items (in the form of a standard composite) provide smaller or comparable sample sizes to adequately power future RCTs on individuals with 'preclinical AD', operationally defined as, in the absence of biomarker data in this analysis, those who were cognitively normal at baseline but with at least one allele of APOE4. For future RCTs on MCI, a subset of 62 informative items whose item score changes occurred later than the onset of MCI also provides sample sizes smaller than or comparable to those from the entire battery. Importantly, we found several folds of increased sample sizes to power future RCTs on either 'preclinical AD' or MCI when items not identified as informative by our analyses were used. These results have 3 major implications in designing future RCTs on early stage AD and in tracking early disease progression. First, they suggest that the commonly used cognitive batteries with years of longitudinal data remain the most important pilot data to design future RCTs on early stage AD, as a majority of the items were informative to early disease progression, consistent with

several reports (albeit cross-sectional). [45–49] Second, given the current challenges facing modern RCTs on early stage AD in terms of choosing appropriate cognitive outcomes and determining the adequate sample sizes, our results suggest the feasibility of using subsets of informative items from an existing cognitive battery as the cognitive efficacy outcome in powering future RCTs on 'preclinical AD' or MCI. Third, when it comes to tracking early disease progression, data collected on a large number of uninformative items may represent less than optimal use of precious research resources as well as an increased burden to research participants. In fact, as demonstrated by our results, the uninformative items in the battery dramatically reduce the power of the informative items in tracking early progression and designing RCTs at the early stage, partly due to the decreased rate of change and the inflated variance because of the contamination from the items not informative to early disease changes.

It is important to point out that the items that were not identified as informative (i.e., uninformative items) by our analyses do not automatically become invalid items from the cognitive battery. These items all have face validity and may be useful in tracking the disease progression at other stages. Even at the early disease stage, their composite score also shows some cognitive decline on subjects with MCI or ‘preclinical AD’ as presented in Table 3, albeit at a much lesser degree when compared to the informative items. Our results should therefore not be interpreted as against the use of these items in AD research. For example, the parent MAP study that provided the data performs annual clinical evaluations on participants until death; thus, items that are sensitive to change among persons with moderate to severe dementia also are needed.

Our study has many strengths. The participants were community-dwelling and examined with annual home visits. Thus, many of the biases inherent in getting persons to be evaluated in a clinic setting are reduced. The overall participation rate exceeded 90% over the entire length of follow-up, reducing bias that results from attrition. The large pool of items from 20 cognitive performance tests is among the largest item pools currently available in community-based prospective cohort studies. This allowed us to identify a large number of items of potential utility in RCTs on preclinical AD and MCI.

Our study also has limitations. The study cohort is selected, the generalizability of the findings needs to be established through independent studies, especially those with a population-based longitudinal design. Further, although it is the most cost effective to utilize an existing longitudinal cognitive database to select most informative items to track early disease progression, neuropsychological theory-based development of outcome measures on prospectively designed longitudinal studies is needed to fully establish the validity and psychometric properties of cognitive outcomes that can serve as the primary efficacy endpoint of future RCTs on early stage AD. Next, our analyses were based on the cognitive data already collected according to a well established protocol. [20] Therefore, the effect of order, presentation, and possibly interference [50–51] of the cognitive testing on our findings can not be adequately addressed. In addition, our analytic approach was based on the technique of survival analysis, which implicitly assumed that everyone will develop MCI/AD if he or she lives long enough. Whereas this assumption may not be entirely unreasonable, its impact on the analysis results warrants further investigation. Finally,

biomarker data would be needed to more accurately define 'preclinical AD', and our results need to be further validated when biomarker data are available.

Acknowledgments

The authors thank RUSH MAP Clinical Core for subject assessments. This study was supported by National Institute on Aging (NIA) grant R01 AG034119 for Chengjie Xiong. This study was also partly supported by the NIA grant P50 AG05681, P01 AG03991, P01AG26276, and U01 AG032438 for Chengjie Xiong and John Morris. The data set used in this manuscript can be requested to RUSH MAP and is available from the corresponding author after appropriate approval.

References

1. Bateman RJ, Xiong C, Benzinger T, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. The New England Journal of Medicine. 2012; 367(9):795–804. [PubMed: 22784036]
2. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dement. 2011; 7:280–92.
3. Katzman R. Editorial: The prevalence and malignancy of Alzheimer disease. A major killer. Arch Neurol. 1976; 33:217–218. [PubMed: 1259639]
4. Price JL, McKeel DW, Buckles VD, Roe CM, Xiong C, Grundman M, Hansen LA, Petersen RC, Parisi JE, Dickson DW, Smith CD, Davis DG, Schmitt FA, Markesbery WR, Kaye J, Kurlan R, Hulette C, Kurland BF, Higdon R, Kukull W, Morris JC. Neuropathology of nondemented aging: presumptive evidence for preclinical Alzheimer disease. Neurobiol Aging. 2009; 30:1026–1036. [PubMed: 19376612]
5. Bennett DA, Schneider JA, Arvanitakis Z, Kelly JF, Aggarwal NT, Shah RC, Wilson RS. Neuropathology of older persons without cognitive impairment from two community-based studies. Neurology. 2006; 66:837–1844.
6. Morris JC, Price JL. Pathologic correlates of nondemented aging, mild cognitive impairment, and early stage Alzheimer's disease. J Mol Neurosci. 2001; 17:101–118. [PubMed: 11816784]
7. Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B. Current concepts in mild cognitive impairment. Arch Neurol. 2001; 58:1985–1992. [PubMed: 11735772]
8. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. Neurology. 1993; 43:2412–2414. [PubMed: 8232972]
9. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR Jr, 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-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011; 7(3):280–92. [PubMed: 21514248]
10. Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, Sano M, Bieliauskas L, Geldmacher D, Clark C, Thal LJ. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord. 1997; 11 (Suppl 2):S13–21. [PubMed: 9236948]
11. Ringman JM, Grill J, Rodriguez-Agudelo Y, Chavez M, Xiong C. Prevention trials in persons at-risk for dominantly-inherited Alzheimer's disease: opportunities and challenges. Alzheimer's & Dementia. 2009; 5:166–171.
12. DeKosky ST, Fitzpatrick A, Ives DG, Saxton J, Williamson J, Lopez OL, Burke G, Fried L, Kuller LH, Robbins J, Tracy R, Woolard N, Dunn L, Kronmal R, Nahin R, Furberg C. GEMS Investigators. The Ginkgo Evaluation of Memory (GEM) study: design and baseline data of a

randomized trial of Ginkgo biloba extract in prevention of dementia. Contemp Clin Trials. 2006; 27:238–253. [PubMed: 16627007]

13. Launer LJ. Prevention of AD: the which, when, and on whom? Alzheimer Dis Assoc Disord. 2006; 20 (3 Suppl 2):S75–78. [PubMed: 16917200]
14. Hampel H, Frank R, Broich K, Teipel SJ, Katz RG, Hardy J, Herholz K, Bokde ALW, Jessen F, Hoessler YC, Sanhai WR, Zetterberg H, Woodcock J, Blennow K. Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. Nature Reviews|Drug Discovery. 2010; 9:560–574.
15. Fagan A, Roe C, Xiong C, Mintun M, Morris JC, Holtzman D. Cerebrospinal fluid tau/ A 42 ratio as a prediction of cognitive decline in nondemented older adults. Arch Neurol. 2007; 64:343–349. [PubMed: 17210801]
16. Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, Lewczuk P, Dean R, Siemers E, Potter W, Lee VM-Y, Trojanowski JQ. the Alzheimer's Disease Neuroimaging Initiative. Cerebrospinal fluid biomarker signature in Alzheimer's Disease Neuroimaging Initiative subjects. Ann Neurol. 2009; 65:403–413. [PubMed: 19296504]
17. Mintun MA, LaRossa GN, Sheline YI, Dence CS, Lee SY, Mach RH, Klunk WE, Mathis CA, DeKosky ST, Morris JC. [^{11}C] PIB in a nondemented population: Potential antecedent marker of Alzheimer disease. Neurology. 2006; 67:446–452. [PubMed: 16894106]
18. Chen K, Reiman EM, Alexander GE, Bandy D, Renaut R, Crum WR, Fox NC, Rossor MN. An automated algorithm for the computation of brain volume change from sequential MRIs using an iterative principal component analysis and its evaluation for the assessment of whole-brain atrophy rates in patients with probable Alzheimer's disease. NeuroImage. 2004; 22:134–143. [PubMed: 15110003]
19. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM338287.pdf
20. 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; 9:646–663. [PubMed: 22471867]
21. Wilson RS, Barnes LL, Bennett DA. Assessment of lifetime participation in cognitively stimulating activities. Journal of Clinical and Experimental Neuropsychology. 2003; 25:634–642. [PubMed: 12815501]
22. Wilson RS, Barnes LL, Krueger KR, Hoganson G, Bienias JL, Bennett DA. Early and late life cognitive activity and cognitive systems in old age. Journal of the International Neuropsychological Society. 2005; 11:400–407. [PubMed: 16209420]
23. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology. 1989; 39:1159–1165. [PubMed: 2771064]
24. Wechsler, D. Wechsler Memory Scale–Revised manual. New York: Psychological Corporation; 1987.
25. Albert MS, Smith LA, Scherr PA, et al. Use of brief cognitive tests to identify individuals in the community with clinically diagnosed Alzheimer's disease. Int J Neurosci. 1991; 57:167–178. [PubMed: 1938160]
26. 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. Psychology and Aging. 2002; 17:179–193. [PubMed: 12061405]
27. Kaplan, EF.; Goodglass, H.; Weintraub, S. The Boston Naming Test. Lea & Febiger; Philadelphia: 1983.
28. Nelson, HE. National Adult Reading Test (NART): test manual. Windsor, UK: NFER Nelson; 1982.
29. Cooper JA, Sagar HJ. Incidental and intentional recall in Parkinson's disease: an account based on diminished attentional resources. J Clin Exp Neuropsychol. 1993; 15:713–731. [PubMed: 8276931]

30. Smith, A. Symbol Digit Modalities Test manual–revised. Los Angeles: Western Psychological; 1982.
31. Ekstrom, RB.; French, JW.; Harman, HH.; Dermen, D. Manual for kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service; 1976.
32. Trenerry, MR.; Crosson, B.; DeBoe, J.; Leber, WR. Psychological Assessment Resources. Odessa, FL: 1989. The Stroop Neuropsychological Screening Test.
33. Benton AL, Varney N, et al. Visuospatial judgment: a clinical test. Arch Neurol. 1978; 35:364–367. [PubMed: 655909]
34. Raven, JC.; Court, JH.; Raven, J. Raven manual: Section 3. 1992. Oxford: Oxford Psychologists Press; 1992. Standard progressive matrices.
35. Folstein MF, Folstein SE, McHugh PR. “Mini-Mental State”: a practical method for grading the mental state of patients for the clinician. J Psychiatr Res. 1975; 12:189–198. [PubMed: 1202204]
36. 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; 34:939–944. [PubMed: 6610841]
37. Bogaerts, Kris; Lesaffre, Emmanuel. Estimating local and global measures of association for bivariate interval censored data with a smooth estimate of the density. Statist Med. 2008; 27:5941–5955.
38. Bogaerts, Kris; Lesaffre, Emmanuel. Modeling the association of bivariate interval-censored data using the copula approach. Statist Med. 2008; 27:6379–6392.
39. Lawless, JF. Statistical models and methods for lifetime data. Wiley; New York: 1982.
40. Diggle, PJ.; Heagerty, P.; Liang, K-Y.; Zeger, SL. Analysis of Longitudinal Data. 2. New York: Oxford University Press; 2002.
41. Donohue MC, Gamst AC, Thomas RG, Xu R, Beckett L, Petersen RC, Weiner MW, Aisen P. the Alzheimer's Disease Neuroimaging Initiative. The relative efficiency of time-to-threshold and rate of change in longitudinal data. Contemp Clin Trials. 2011; 32(5):685–693.10.1016/j.cct.2011.04.007 [PubMed: 21554992]
42. Xiong, C.; Zhu, K.; Yu, K.; Miller, JP. Statistical Modeling in Biomedical Research: Longitudinal Data Analysis. In: Rao, CR.; Philip Miller, J.; Rao, DC., editors. Handbook of Statistics: Epidemiology and Medical Statistics. Amsterdam: Elsevier B.V; 2008.
43. SAS Institute, Inc. User’s Guide. Cary, NC: 2013. SAS/STAT® 13.1.
44. Stern Y, Jacobs DM. Preliminary findings from the Predictors Study: utility of clinical signs for predicting disease course. Alzheimer’s Disease and Associated Disorders. 1995; 9(1):S14–S18.
45. Grober E, Buschke H, Crystal H, Bang S, Dresner R. Screening for dementia by memory testing. Neurology. 1988; 38:900–903. [PubMed: 3368071]
46. Backman L, Jones S, Berger AK, et al. Cognitive impairment in preclinical Alzheimer’s disease: a meta-analysis. Neuropsychology. 2005; 19(4):520–531. [PubMed: 16060827]
47. Hall CB, Derby C, LeValley A, Katz MJ, Verghese J, Lipton RB. Education delays accelerated decline on a memory test in persons who develop dementia. Neurology. 2007; 69:1657–1664. [PubMed: 17954781]
48. Hall CB, Lipton RB, Sliwinski M, Katz MJ, Derby CA, Verghese J. Cognitive activities delay onset of memory decline in persons who develop dementia. Neurology. 2009; 73:356–361. [PubMed: 19652139]
49. Twamley EW, Legendre Ropacki SA, Bondi MW. Neuropsychological and neuroimaging changes in preclinical Alzheimer’s disease. J Int Neuropsychol Soc. 2006; 12(5):707–735. [PubMed: 16961952]
50. Rowe G, Hasher L, Turcotte J. Interference, aging, and visuospatial working memory: the role of similarity. Neuropsychology. 2010; 24(6):804–7. [PubMed: 21038966]
51. Beglinger LJ, Gaydos B, Tangphao-Daniels O, Duff K, Kareken DA, Crawford J, Fastenau PS, Siemers ER. Practice effects and the use of alternate forms in serial neuropsychological testing. Archives of Clinical Neuropsychology. 2005; 20:517–529. [PubMed: 15896564]

Appendix

Table 2A

Informative items for designing RCTs on early AD (187 on 'Preclinical AD', and 62 on MCI) and their description

Test name	List of items	Ways items are administered	Items for RCTs on MCI (Y=Yes)
Boston Naming Test	bed	Participants are shown pictures of certain objects. Then they are requested to name the objects.	Y
	camel		Y
	canoe		Y
	domino		N
	flower		Y
	funnel		N
	hammock		Y
	harmon		Y
	house		Y
	mask		Y
	tongs		N
	volcano		N
	whistle		Y
Color Naming	cname	Number of colors read correctly in 30 secs.	Y
Delayed Story Recall	Injuries	A three sentence story is read to the participants. Then they are requested to recall the story after a distractor-filled delay of approximately 3 minutes. Item names are key words of the story.	N
	Everyone		N
	Well		N
	Three		N
	Children		Y
	House		N
	Fire		Y
	Fireman		N
	Climb		N
	Children		N
	Rescured		N
	Minor		N
Digit Ordering	DigitOrder-41	A series of numbers are read aloud to the participants. One series at a time. After each series, participants are requested to repeat the series starting with the smallest number and going to the largest number.	Y
	DigitOrder-98		Y
	DigitOrder-104		Y
	DigitOrder-263		Y
	DigitOrder-2413		N
	DigitOrder-4216		N
	DigitOrder-37570		N
	DigitOrder-79210		N

Test name	List of items	Ways items are administered	Items for RCTs on MCI (Y=Yes)
Digits Backward	DigitBack-38	A series of number sequences of increasing length are read out to the participants. Participants are requested to repeat the numbers backwards.	Y
	DigitBack-493		N
	DigitBack-526		N
	DigitBack-3814		N
	DigitBack-1795		N
	DigitBack-62972		N
	DigitBack-48527		N
Digits Forward	DigitFor-8396	A series of number sequences of increasing length are read out to the participants. Participants are requested to repeat the numbers forwards.	Y
	DigitFor-36925		N
	DigitFor-69471		N
	DigitFor-918427		N
	DigitFor-635482		N
	DigitFor-2814975		N
Immediate Story Recall	Injuries	A three sentence story is read to the participants. Then they are requested to recall the story immediately. Item names are key words of the story.	N
	Everyone		N
	Well		N
	Three		Y
	Children		Y
	House		N
	Fire		Y
	Fireman		Y
	Climb		N
	Children		Y
	Rescured		Y
	Minor		N
Line Orientation	line10a	Each item requires the participants to estimate the angle subtended by two lines in a match-to-sample format.	Y
	line10b		N
	line11a		Y
	line11b		N
	line12a		N
	line12b		N
	line13a		N
	line13b		N
	line14a		N
	line14b		N
	line15b		N
	line1a		N
	line2a		N

Test name	List of items	Ways items are administered	Items for RCTs on MCI (Y=Yes)
	line2b		N
	line4a		N
	line4b		N
	line5a		Y
	line5b		N
	line6b		N
	line7b		N
	line8a		N
	line8b		N
	line9b		N
Logical Memory IA		A brief story is read to the participants. Then they are asked to retell it from memory immediately	Y
Logical Memory IIA		A brief story is read to the participants. Then they are asked to retell it from	Y
MMSE30	apple	Repeat the 1st word of 3 words read before-- apple	Y
	folds	Fold a piece of paper in half	Y
	paper	Put a piece of paper in right hand	N
	penny	Repeat the 3rd word of 3 words read before-- penny	Y
	places	Place a piece of paper on lap	N
	StreetName	Name the street number of this place	N
	StreetNumber	Name the street name of this place	N
	WORLD-backward	Spell the word 'world' backward	N
	apple-recall	Recall the first word of 3 words given previously -- apple	N
	table-recall	Recall the second word of 3 words given previously -- table	N
	penny-recall	Recall the third word of 3 words given previously -- apple	N
	RepeatPhrase	Repeat a phrase	N
	ReadWords	Read the words shown on a card	Y
	WriteSentence	Write any complete sentence	Y
	Year	Name the current year	Y
	Copy	Copy the drawing on a piece of paper	N
	Season	Name the current season	N
	Day	Name the day of the week	Y
	Month	Name the current month	Y
	State	Name the State of this place	Y
	Country	Name this country	Y
	City	Name the city of this place	Y
	Room	Name the room of this place	Y

Test name	List of items	Ways items are administered	Items for RCTs on MCI (Y=Yes)
	table	Repeat the 2nd word of 3 words read before-- table	Y
Number Comparison	Comparison-3	Participants are presented with 48 pairs of numbers. Some of the numbers are exactly the same while others do not match. The participants are asked to identify pairs as “same” or “different”.	N
	Comparison-4		N
Progressive Matrices	Pattern-a11	Participants are shown a series of pattern and asked to identify the pattern below which would complete the pattern on top.	N
	Pattern-a2		Y
	Pattern-a5		Y
	Pattern-a6		Y
	Pattern-a7		N
	Pattern-a8		N
	Pattern-b10		N
	Pattern-b2		N
	Pattern-b3		N
	Pattern-b4		N
	Pattern-b5		N
	Pattern-b6		N
	Pattern-b8		N
Reading Test	Ache	Participants are shown a series of words and asked to pronounce these words the best they can.	Y
	Placebo		N
	Façade		N
	Impugn		N
	Blatant		N
	Reify		N
	Topiary		N
	Naïve		N
	Recipe		Y
	Heir		Y
	Indict		N
	Debt		Y
	Sieve		N
	Corps		N
Symbol Digit Modalities	Symbol-1	Participants are shown a series of symbol. Each symbol corresponds to a number from 1 to 9. They are asked to call out the numbers that match the symbols shown to them one at a time.	N
	Symbol-2		N
	Symbol-3		N
	Symbol-4		N
Verbal Fluency	animal	Participants are asked to generate exemplars from the category in successive 1 minute trials	Y
	fruit	Participants are asked to generate exemplars from the category in successive 1 minute trials.	Y

Test name	List of items	Ways items are administered	Items for RCTs on MCI (Y=Yes)
Word List Memory	wordt1_1-butter	A 10-word list is presented, three times (total of 30 words), with three immediate recall trials and delayed tests of recall and recognition.	N
	wordt1_2-Arm		N
	wordt1_3-Shore		N
	wordt1_7-Pole		N
	wordt1_9-Grass		N
	wordt2_2-Cabin		N
	wordt2_3-Butter		N
	wordt2_4-Shore		N
	wordt2_5-Engine		N
	wordt2_6-Arm		N
	wordt2_7-Queen		N
	wordt2_8-Letter		N
	wordt2_9-Pole		N
	wordt3_1-Queen		N
	wordt3_2-Grass		N
	wordt3_3-Arm		N
	wordt3_4-Cabin		N
	wordt3_5-Pole		N
	wordt3_6-Shore		N
	wordt3_7-Butter		N
	wordt3_8-Engine		N
	wordt3_9-Ticket		N
	wordt3_x-Letter		N
Word List Recall	recall_1-Butter	Participants are asked to read a list of ten words one at a time. Few minutes later they are asked to identify as many words as they can recall.	N
	recall_2-Arm		N
	recall_3-Shore		N
	recall_4-Letter		N
	recall_5-Queen		N
	recall_6-Cabin		N
	recall_7-Pole		N
	recall_8-Ticket		N
	recall_9-Grass		N
	recall_x-Engine		N
Word List Recognition	wordrec1-LETTER	Participants are shown ten sets of four words, one set at a time, and asked to select the words from each set that (s)he was shown previously.	Y
	wordrec2-POLE		Y
	wordrec3-ENGINE		Y
	wordrec4-ARM		Y
	wordrec5-QUEEN		Y

Test name	List of items	Ways items are administered	Items for RCTs on MCI (Y=Yes)
	wordrec6-CABIN		Y
	wordrec7-TICKET		Y
	wordrec8-BUTTER		Y
	wordrec9-GRASS		Y
	wordrecx-SHORE		Y
Word Corrected Read	WordRead	Words correctly read. Part of Stroop Neuropsychological Screening Test.	Y

Table 1

Baseline characteristics of the sample (Total n = 1513)

	Normal, n=1037	**MCI, n=402**	**AD, n=74**
Age (mean, SD)	78.81 (7.47)	81.91 (7.44)	84.86 (6.07)
Gender (% of female)	75.70	69.90	54.05
Education (y, mean, SD)	14.45 (3.28)	14.43 (3.02)	14.09 (3.92)
Race: % for Caucasian	93.83	90.80	94.59
% for African American	5.40	7.96	5.41
% of others	0.68	0.75	0
APOE4 positive (%)	17.65	27.61	32.43
MMSE (mean, SD)	28.41 (1.70)	26.6 (2.5)	18.33 (6.73)
Global cognition (mean, SD)	0.26 (0.46)	−0.41 (0.49)	−1.51 (0.74)

Table 2

Item characteristics of the cognitive tests at baseline and the number of informative items identified (SD=Standard Deviation, Inf.=Informative)

Test name	Number of items	Score types	Total score: Mean (SD)			Number Inf. Items
			Normal	MCI	AD	
color name	1	count	19.61 (7.24)	14.74 (7.49)	6.83 (6.95)	1
Logical Memory IIA	1	count	10.34 (3.83)	5.35 (4.04)	1.53 (2.23)	1
Boston Naming Test	15	binary	14.11 (1.63)	13.85 (6.62)	17.8 (26.03)	13
Verbal Fluency	2	count	35.32 (8.38)	27.66 (8.33)	17.03 (9.35)	2
Digits Backward	12	binary	6.48 (2.11)	5.39 (1.88)	4.86 (3.87)	7
Digits Forward	12	binary	8.45 (2.02)	7.77 (2.11)	7.16 (3.38)	6
Digit Ordering	10	binary	7.29 (1.45)	6.34 (1.71)	3.32 (2.55)	8
Delayed Story Recall	12	binary	9.64 (2.95)	7.59 (3.85)	10.04 (21.12)	12
Immediate Story Recall	12	binary	9.91 (1.69)	8.57 (3.25)	12.15 (20.36)	12
Line Orientation	30	binary	25.11 (3.32)	22.84 (4.22)	19.63 (6.70)	23
MMSE	27	binary	25.60 (2.17)	23.88 (3.35)	18.93 (19.84)	24
Number Comparison	4	binary	1.81 (0.96)	1.44 (0.96)	0.86 (0.75)	2
Progressive Matrices	16	binary	11.74 (2.60)	9.9 (2.88)	8.36 (2.67)	13
Reading Test	15	binary	12.65 (3.42)	12.32 (7.40)	22.8 (33.89)	14
Symbol Digit Modalities	5	binary	3.09 (1.23)	2.18 (1.32)	0.70 (0.94)	4
Word List Memory	30	binary	18.31 (7.93)	14.55 (13.62)	28.59 (66.46)	23
Word List Recall	10	binary	6.03 (3.00)	3.59 (7.04)	7.70 (22.72)	10
Word List Recognition	10	binary	9.84 (0.49)	8.67 (4.56)	13.25 (22.94)	10
Logical Memory IA	1	count	11.95 (3.84)	7.70 (4.12)	3.19 (3.25)	1
Words correctly read	1	count	51.13 (13.02)	45.8 (14.15)	30.75 (16.07)	1

Table 3

Estimated mean change (standard error, SE) of cognitive composites from baseline on individuals with MCI and 'Preclinical AD'

Types of items used in the composite	Clinical stage	Years of follow-up (sample size)	Mean change(SE) from baseline	Ratio of Mean/SE
Informative Items	Preclinical AD	2y (N=139)	−0.085 (0.030)	−2.833
		3y (N=121)	−0.093 (0.031)	−3.000
		4y (N=102)	−0.197 (0.036)	−5.472
	MCI	2y (N=275)	−0.261 (0.038)	−6.868
		3y (N=212)	−0.416 (0.048)	−8.667
		4y (N=174)	−0.558 (0.060)	−9.300
Uninformative Items	Preclinical AD	2y (N=139)	−0.034 (0.030)	−1.133
		3y (N=121)	−0.034 (0.035)	−0.971
		4y (N=102)	−0.083 (0.040)	−2.075
	MCI	2y (N=275)	−0.072 (0.028)	−2.571
		3y (N=212)	−0.144 (0.040)	−3.600
		4y (N=174)	−0.206 (0.045)	−4.578
All Tests	Preclinical AD	2y (N=139)	−0.063 (0.024)	−2.625
		3y (N=121)	−0.082 (0.028)	−2.929
		4y (N=102)	−0.182 (0.033)	−5.515
	MCI	2y (N=275)	−0.129 (0.022)	−5.863
		3y (N=212)	−0.244 (0.030)	−8.133
		4y (N=174)	−0.330 (0.036)	−9.167

Table 4

Sample sizes of future RCTs on 'Preclinical AD' and MCI (Inf=cognitive composite using informative items alone, UnInf=cognitive composite using uninformative items, Glob=cognitive composite using all 20 tests)

		2 Years follow-up			3 Years follow-up			4 Years follow-up		
Clinical Stages	**ES (%)**	**Inf**	**Glob**	**UnInf**	**Inf**	**Glob**	**UnInf**	**Inf**	**Glob**	**UnInf**
Pre-clinical AD	20	13933	15545	88635	10632	10896	102055	2706	2658	18637
	25	8917	9949	56727	6805	6973	65316	1732	1701	11928
	30	6193	6909	39394	4726	4843	45358	1203	1182	8283
	35	4550	5076	28943	3472	3558	33325	884	868	6086
	40	3484	3887	22159	2658	2724	25514	677	665	4660
	45	2753	3071	17509	2101	2153	20160	535	525	3682
	50	2230	2488	14182	1702	1744	16329	433	426	2982
	55	1843	2056	11721	1406	1441	13495	358	352	2465
	60	1549	1728	9849	1182	1211	11340	301	296	2071
	65	1320	1472	8392	1007	1032	9663	257	252	1765
	70	1138	1269	7236	868	890	8332	221	217	1522
	75	991	1106	6303	757	775	7258	193	189	1326
	80	871	972	5540	665	681	6379	170	167	1165
MCI:	20	4594	6522	32030	2173	2437	12610	1581	1560	6393
	25	2940	4174	20499	1391	1560	8071	1012	999	4091
	30	2042	2899	14236	966	1083	5605	703	694	2841
	35	1500	2130	10459	710	796	4118	517	510	2088
	40	1149	1631	8008	544	610	3153	396	390	1599
	45	908	1289	6327	430	482	2491	313	309	1263
	50	735	1044	5125	348	390	2018	253	250	1023
	55	608	863	4236	288	323	1668	210	207	846
	60	511	725	3559	242	271	1402	176	174	711
	65	435	618	3033	206	231	1194	150	148	606
	70	375	533	2615	178	199	1030	130	128	522
	75	327	464	2278	155	174	897	113	111	455

Clinical Stages		2 Years follow-up			3 Years follow-up			4 Years follow-up		
	ES (%)	Inf	Glob	UnInf	Inf	Glob	UnInf	Inf	Glob	UnInf
	80	288	408	2002	136	153	789	99	98	400