UC Irvine
UC Irvine Previously Published Works

Title

Estimating sample sizes for predementia Alzheimer's trials based on the Alzheimer's Disease Neuroimaging Initiative

Permalink

https://escholarship.org/uc/item/3jg728hx

Journal

Neurobiology of Aging, 34(1)

ISSN

0197-4580

Authors

Grill, Joshua D

Di, Lijie

Lu, Po H

et al.

Publication Date

2013

DOI

10.1016/j.neurobiolaging.2012.03.006

Peer reviewed

Estimating sample sizes for pre-dementia Alzheimer's trials based on the Alzheimer's Disease Neuroimaging Initiative

Joshua D. Grill, PhD[1,*], **Lijie Di, MS**[1,3], **Po H. Lu, PsyD**[1], **Cathy Lee, PhD**[1], **John Ringman, MD, MS**[1], **Liana G. Apostolova, MD, MS**[1], **Nicole Chow**[1], **Omid Kohannim**[2], **Jeffrey L. Cummings, MD**[5], **Paul M. Thompson, PhD**[2,4], and **David Elashoff, PhD**[1,3] **for the Alzheimer's Disease Neuroimaging Initiative**[#]

[1]Mary S. Easton Center for Alzheimer's Disease Research, University of California, Los Angeles, David Geffen School of Medicine

[2]Laboratory of Neuro Imaging, Department of Neurology, University of California, Los Angeles, David Geffen School of Medicine

[3]Department of Medicine Statistics Core, University of California, Los Angeles, David Geffen School of Medicine

[4]Department of Psychiatry, University of California, Los Angeles, David Geffen School of Medicine

[5]Lou Ruvo Center for Brain Health, Cleveland Clinic, Las Vegas, NV

Abstract

This study modeled predementia Alzheimer's disease (AD) clinical trials. Longitudinal data from cognitively normal (CN) and mild cognitive impairment (MCI) participants in the AD Neuroimaging Initiative were used to calculate sample size requirements for trials using outcome measures including: the Clinical Dementia Rating scale sum of boxes (CDR-sb), Mini Mental Status Examination (MMSE), AD assessment scale-cognitive subscale with and without delayed recall, and the Rey Auditory Verbal Learning task (RAVLT). We examined the impact on sample sizes of enrichment for genetic and biomarker criteria, including cerebrospinal fluid protein and neuroimaging analyses. We observed little cognitive decline in the CN population at 36 months, regardless of the enrichment strategy. Nonetheless, in CN subjects, using RAVLT total as an outcome at 36 months required the fewest subjects across enrichment strategies, with apolipoprotein E genotype ε4 carrier status requiring the fewest (n=499 per arm to demonstrate a 25% reduction in disease progression). In MCI, enrichment reduced the required sample sizes for trials, relative to estimates based on all subjects. For MCI, the CDR-sb consistently required the smallest sample sizes. We conclude that predementia clinical trial conduct in AD is enhanced by the use of biomarker inclusion criteria.

*Corresponding author: Joshua Grill, PhD, UCLA Easton Alzheimer's Center, 10911 Weyburn Ave, Suite 200. Los Angeles, CA 90095, USA. Tel.: (310) 794-2511; fax: (310) 794-3148. jgrill@mednet.ucla.edu.

Keywords

Alzheimer's disease; clinical trials; mild cognitive impairment; preclinical; predementia; sample size; enrichment

---

Background

Studies of the biology of Alzheimer's disease (AD) have identified an array of targets for potential disease-modifying therapies [39] but clinical trials in patients with dementia have been unsuccessful so far [14,24,50,52,53]. Biological substrates of AD can be identified before patients become demented [46] and some AD biomarkers reach peak levels of abnormality prior to diagnosis [29,37]. Failed dementia trials may have intervened too late in the disease process to be effective [62].

Clinical trials of investigational drugs targeting AD biology can enroll patients earlier in the disease, before criteria for dementia are fulfilled. Primary prevention trials enroll volunteers with no clinical or biological signs of AD at baseline but require thousands of participants and take many years to complete, because only a fraction of participants will develop AD [15]. To date, few primary AD prevention trials have been conducted and no agent has been shown to delay or prevent dementia onset. Secondary prevention trials can enroll participants at increased risk for dementia, affording decreased sample sizes and trial lengths. Secondary prevention trials have included individuals with mild cognitive impairment (MCI), a clinical syndrome defined by memory impairment or other cognitive problems, when compared to age- and education-matched norms, in the absence of functional decline [49]. Even some trials enrolling MCI participants have encountered low rates of disease progression [21].

Biological markers of AD predict clinical progression and may be used to identify potential trial participants at greatest risk for dementia. Low levels of amyloid beta (Aβ) or elevated levels of total tau (tTau) or phosphorylated tau (pTau) in the cerebrospinal fluid (CSF; e.g. [40]); evidence of cerebral atrophy on magnetic resonance imaging (MRI; e.g. [6]); and brain glucose hypometabolism observed with fluorodeoxyglucose positron emission tomography (FDG PET) (e.g. [33]) identify MCI patients at increased and more immediate risk for AD dementia. Even in asymptomatic individuals, the presence of biological evidence of AD significantly increases the risk for future cognitive impairment and AD dementia [7,17,47].

Thus, it is likely that using AD biomarkers as enrollment criteria can reduce the number of participants needed and study duration for AD prevention trials. Using AD biomarkers as outcome measures in AD trials can similarly improve trial efficiency [9,27,28,31,35,57,59]. The US Food and Drug Administration (FDA), however, has not accepted any biomarker as a surrogate suitable for use as a primary outcome measure in AD trials. Moreover, FDA guidance outlines the use of clinical measures to achieve marketing approval [30,34]. Therefore, registration trials, even those conducted in very mild disease, continue to use the AD Assessment Scale-cognitive subscale (ADAS-cog) and other clinical scales as primary outcome measures.

The statistical power of predementia trials may be improved by population enrichment strategies using biomarkers. These trials might be able to employ a single primary outcome measure (rather than dual primary outcomes, as is the case in dementia trials [3]). Using the AD Neuroimaging Initiative (ADNI) dataset, we sought to identify the best enrichment strategies for predementia trials in relation to outcome measures to optimize statistical

power. We hypothesized that enriching cognitively normal (CN) and MCI trial populations through biomarker criteria would reduce required sample sizes.

Methods

ADNI

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.ucla.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the FDA, private pharmaceutical companies and nonprofit organizations, as a $60 million, five-year public- private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California – San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research, approximately 200 cognitively normal older individuals to be followed for three years, 400 people with MCI to be followed for three years and 200 people with early AD to be followed for two years." For up-to-date information, see www.adni-info.org.

The current analyses focused on the first iteration of ADNI, which enrolled a cohort of volunteers who were CN, MCI, and AD dementia at baseline. Clinical and biological data were collected, including magnetic resonance imaging volumetric measures, FDG PET, and CSF protein analysis. The current analyses focused on data from CN and MCI ADNI subjects. CN subjects had no subjective memory complaints at baseline. CN and MCI subjects scored between 24 and 30 on the Mini-Mental State Examinations (MMSE). CN subjects had a global Clinical Dementia Rating scale (CDR) of 0 at baseline. MCI subjects scored 0.5, with a required memory box score of 0.5 or higher at baseline. Subjects were also required to meet criteria for memory performance on the Wechsler Memory Scale-Revised Logical Memory II subscale: CN subjects   9 for 16 years or more of education,   5 for 8–15 years of education, and   3 for 0–7 years of education; MCI subjects   8 for 16 years or more of education,   4 for 8–15 years education, and   2 for 0–7 years of education. ADNI CN subjects could not have impairment in activities of daily living and MCI subjects could not meet criteria for dementia.

All ADNI participants had a modified Hachinski Scale score <4; a Geriatric Depression Scale (abbreviated 15-item version) score <6, were fluent in English or Spanish, had a suitable study partner who could accompany them to study visits, and lived at home. They had no significant neurologic or psychiatric disease, no history of alcohol or substance abuse, no clinically significant lab abnormalities in B12 level, rapid plasma reagin, or thyroid function tests, and no contraindication to neuroimaging. They did not take psychoactive drugs, including antidepressants with anticholinergic properties, or warfarin. They had not participated in a clinical trial of an investigational medication within one month of baseline or for the duration of their participation in ADNI and they were not involved in other studies that included neuropsychological testing that could interfere with the ADNI-related testing.

ADNI was designed to parallel AD clinical trials, employing a variety of psychometric outcome measures that are common in AD registration trials. We examined data collected from all CN and MCI participants at baseline and at 12, 24, and 36 months. We focused on outcome measures that included assessments of memory including the ADAS-cog with (ADAS12) and without (ADAS11) delayed recall component [51]; the MMSE [22]; the Rey Auditory Verbal Learning Task [63] total score summing the number recalled over the initial 5 learning trials (RAVLT total) and the recall score after a 30-minute delay (RAVLT delayed recall); and the CDR sum of the boxes score (CDR-sb)[45]. Longitudinal clinical data and baseline biomarker data were downloaded from the ADNI public database (http://www.loni.ucla.edu/ADNI/Data/) on May 2, 2011.

Enrichment strategies

Multiple strategies were used to limit the data used to calculate sample sizes. Enrichment was hypothesized to create a sample in which there would be a greater magnitude of cognitive decline over time and/or reduced variance, thus reducing the sample size necessary to detect a drug effect, at a specific level of statistical power. Each enrichment strategy was applied to both the CN and the MCI populations.

Apolipoprotein E (ApoE) ε4 carrier status—ApoE genotype is a well described genetic risk factor for AD [12]. ADNI ApoE genotyping was performed using blood samples, at the University of Pennsylvania AD Biomarker Laboratory. Subjects were divided into those who did and those who did not carry at least one ε4 allele.

CSF protein analysis—CSF collection and analysis have been described elsewhere [60]. CSF protein levels included here are $A\beta_{1-42}$ ($A\beta$), total Tau (tTau), and Tau phosphorylated at threonine 181 (pTau). In addition, the ratio of tTau/$A\beta$, and pTau/$A\beta$ were examined as enrichment criteria. Criteria for inclusion are described below.

Hippocampal volume—Hippocampal volumes were measured using a machine learning method based on adaptive boosting (AdaBoost), as described previously [43]. Briefly, this automated method performs brain MRI segmentation and quantifies hippocampal volume. It uses a pool of 14,000 features such as image intensity; tissue classification maps of gray matter, white matter, and CSF; and neighborhood based features from each voxel and designs an algorithm that can optimally segment the hippocampus (or another brain structure) from a limited region within brain MRIs standardized against a registered template. A weighted voting algorithm combines "weak learners" into a "strong learner." Prior work has shown this method consistently agrees with expert human rater tracings [42–44].

Lateral ventricle volume—Ventricular volumes were assessed using a semi-automated, multi-atlas segmentation technique that was developed at UCLA [10,11]. A small number (n=6) of lateral ventricles from the sample were manually traced and used to create ventricular models that could be converted into parametric surface atlases. Fluid registration of these atlases to every subject was performed, and an averaging technique combined the surface atlases for each image volume. The choice of the number of templates was empirically based on optimizing the False Discovery Rate. This technique distinguishes AD from normal controls and also demonstrates differences in ventricular volume based on ApoE carrier status [10].

Cerebral metabolism—Predefined ROI analysis of FDG PET cerebral metabolism was conducted as described previously [32]. Metabolic signal was intensity-normalized within subjects against the cerebellar *vermis* and *pons*. FDG uptake was extracted for left and right

temporal lobes as regions of interest. The average glucose uptake for this region across hemispheres was used to produce a single value for each subject at baseline. Only a subset (102 CN and 206 MCI) of participants in ADNI underwent FDG PET imaging.

Biological enrichment criteria cutoff points

To decide upon inclusion criteria for each enrichment strategy, we performed receiver operating characteristic (ROC) analyses, using data from the ADNI AD and CN populations. The inclusion criteria for MRI measures of volume (hippocampus and lateral ventricles) and FDG measures of metabolism were set to the threshold value for those measures that maximized the Youden index (the sensitivity plus the specificity minus 1 [66]) for discrimination between AD and CN groups. Previously determined cutoff CSF criteria based on ROC analyses of neuropathologically confirmed diagnoses of AD and normal controls were applied [60]. Specifically, the following criteria for inclusion were used: Aβ<192 pg/ mL; tTau>93 pg/mL; pTau>23pg/mL; ratio of tTau/Aβ>.39; ratio of pTau/Aβ>.1. These cutoffs have been used in prior ADNI analyses [48]. Baseline CSF samples were available from 200 MCI and 114 CN ADNI participants

Sample size calculations

We examined the mean decline in cognitive outcome measures at 12, 24, and 36 months. Participant data were included for all available outcome measures (i.e., missing values for an outcome of interest did not preclude inclusion for another). Sample size estimates under an assumption of normality and known variance were calculated from the equation:

$$\text{Sample size} = (z_{1-\beta} + z_{1-\alpha/2})^2 \times (2\sigma^2)/(\Delta\mu)^2$$

Here, $z_{1-\beta}$ = 0.842 to provide 80% power; $z_{1-\alpha/2}$ = 1.96 to test at the 5% level; $\Delta\mu$ is the mean change in score on the outcome of interest, relative to baseline, multiplied by the drug effect (0.25) to reflect the estimated mean difference between placebo group change scores and drug group change scores; and σ is the SD of the change scores in the groups (assuming SD is the same in treatment and placebo groups). This sample size equation is well described in the literature and has been used previously by others to estimate sample sizes in AD clinical trials [23,36]. We report sample sizes per trial arm, powered to detect a 25% drug effect (slowing of cognitive decline).

To assist in the comparison of sample size estimates, we calculated the 95% confidence intervals (CI) for the sample size. These confidence intervals were estimated by first calculating the 95% confidence interval for the effect size $\Delta\mu/\sigma$ through the noncentrality parameter *t* score [13]. These limits were then used in the equation above to calculate the 95% CI of the sample sizes. In cases where the confidence interval for the effect size crossed 0, the upper bound of the sample size CI is denoted as ∞. We also calculated 95% confidence intervals using bootstrap resampling, using 1000 iterations for each scenario. We found these confidence intervals to be on average 25% narrower than those calculated with our formula. We present only the more conservative estimates. Formal statistical comparisons of sample size outputs were not performed.

Findings

ADNI Subjects

Demographic summaries and baseline scores on clinical outcome measures for the included populations are found in Table 1. Data from the entire ADNI sample was used for this study.

Table 2 displays the number of ADNI participants who met criteria for the biomarker enrichment strategies at baseline. Twenty-seven percent of CN participants were ApoE ε4 carriers. Twenty-two percent of CN participants demonstrated hippocampal atrophy (<3070.5 mm^3) while 34% met criteria for lateral ventricular enlargement (28,216.5 mm^3). Of those who underwent LP, 38% met criteria for low CSF Aβ. Of those CN participants who had PET scans, 14.7% met criteria for temporal lobe hypometabolism (ratio of FDG uptake below 1.14).

More than half (53%) of the MCI population were ApoE ε4 carriers. Among biomarker enrichment strategies, CSF strategies included the most MCI patients (for example, 67.5% of participants who underwent LP met CSF Aβ criteria), while a smaller proportion of participants (37.9%) met FDG PET criteria for enrichment (Table 2).

Estimated trial sample sizes: CN population

At 12 months, the CN population demonstrated mean worsening only on the CDR-sb. At 24 months, mean decline was observed only on the CDR-sb and MMSE. Therefore, we focused on trial estimations for the CN population based on the 36-month longitudinal data. At 36-months mean worsening was observed for the MMSE, CDR-sb, RAVLT total and RAVLT delayed recall, but not the ADAS11 or ADAS12 (Table 3). Sample size calculations ranged from 1414 participants/arm for the RAVLT total to 50,790 participants/arm for the MMSE (Table 4).

With few exceptions, trial sample sizes based on enriched populations required fewer participants than did trials based on the entire CN population (Table 4). Whereas the CN population as a whole did not demonstrate a mean decline on the ADAS11 or ADAS12 at 36-months, enrichment for persons who met CSF Aβ, CSF pTau, CSF ratio of pTau/Aβ, FDG PET hypometabolism, and hippocampal volume criteria resulted in mean decline (and possible sample size calculations; Table 4) for these outcome measures. In each of these scenarios, trials using the ADAS12 required fewer participants than trials using the ADAS11. Enrichment for ApoE ε4 carriers, CSF tTau, and the ratio of tTau/Aβ resulted in decline on the ADAS12 but not the ADAS11. For six of the nine examined enrichment strategies, trials using the RAVLT total required the fewest participants. Among these, trials enriched for ApoE ε4 carriers required the fewest participants (n=499, CI: 243–1659). Trials enriched for FDG PET or hippocampal volume required the fewest participants when using the CDR-sb as an outcome. Trials enriched for lateral ventricle volume required the fewest participants when using the RAVLT delayed recall.

Estimated trial sample sizes: MCI population

Mean worsening was detected for the entire MCI population at all time points for each of the clinical outcome measures examined (see for example Table 5). Sample size requirements decreased substantially with increasing trial length (data not shown). We chose to focus on 24-month MCI trials (Tables 5 and 6), as this represents a likely scenario for these trials [1], but the results we present were similar at both 12 and 36 months (data not shown). Trials using the CDR-sb required fewer participants than trials using any other outcome measure. Trials using the ADAS12 required fewer participants than did trials using the ADAS11, or any other outcome beside the CDR-sb.

Within enrichment strategies, trials using the CDR-sb consistently produced the lowest required sample sizes (Table 6). The sole exception was trials enriched for FDG PET hypometabolism, which required the fewest subjects using the MMSE as an outcome (n=314, CI: 179–725). Enrichment using CSF criteria yielded the numerically lowest number of subjects required to detect changes in the CDR-sb for a 24-month trial. In every

scenario we examined, biomarker enrichment of the MCI population produced lower necessary sample sizes for each outcome measure.

Implications

Few studies have explored the ability to perform AD prevention trials in populations enrolled with no cognitive complaint. Individuals with no demonstrable cognitive abnormality who meet criteria for AD biomarkers may be defined as having "preclinical AD" [61] or as being "asymptomatic at risk for AD" [19]. Trials in persons who meet these criteria are being planned [2,3]. Similarly, a variety of working groups have proposed inclusion of "prodromal AD," MCI patients who meet AD biomarker criteria, in trials [2,5,20] and trials implementing these guidelines are now underway (www.clinicaltrials.gov).

How to best design successful predementia clinical trials is controversial and requires guidance from research studies. This project sought to identify optimal clinical outcome measures and biological enrichment strategies for use in AD trials enrolling asymptomatic or mildly symptomatic participants. We chose to examine continuous outcome measures, rather than "conversion" outcomes, because such measures are likely to provide greater sensitivity and therefore reduced sample sizes for predementia trials. We found that for most outcome measures, 12- and 24-month trials of cognitively normal participants are not realistic. Decline in outcome measure scores for this population at 36-months, when present, were small and trials to demonstrate a reduction in that decline required very large study populations. A relative exception to this was for trials using the RAVLT total score, which required 1404 participants per study arm. Even so, the CN population did not demonstrate a mean decline from baseline at 12 or 24 months for this outcome measure (data not shown) and longer-term follow-up and confirmation in independent samples of the decline in the RAVLT in CN participants is warranted.

When the CN population was enriched for persons meeting AD biomarker criteria at baseline, decline was observed for outcome measures that did not demonstrate detectable decline for the entire CN population, and the decline observed on the remaining outcome measures was increased in degree. The numeric reductions in the outputs of sample size calculations frequently exceeded 50% in some scenarios (Table 4). Thus, biomarker enrichment increases the efficiency of performing AD clinical trials in asymptomatic patients. The ideal means of enrichment, however, are not yet clear. Determining the specificities and sensitivities of methods to predict future cognitive decline for each biomarker criterion remains an important area of study. In the current exercise, for example, each of CSF Aβ, FDG PET hypometabolism, and hippocampal volume successfully reduced the necessary sample sizes of trials using the CDR-sb as an outcome. Alternatively, hippocampal volume and FDG-PET failed as enrichment strategies for trials using the RAVLT total, while enrichment for CSF Aβ still produced sample size requirements lower than that of the entire CN population. Determining what enrichment strategy is best for what outcome measure may depend on the specifics of the study population.

Not surprisingly, the overall MCI population produced more consistent decline on trial outcome measures, resulting in consistent calculation of trial sample sizes that were reduced, relative to the estimates based on the CN population. In the MCI population as a whole, the CDR-sb required substantially fewer participants than did all other clinical measures, in line with the observations of others [4,41]. Enrichment of the MCI population by any strategy reduced the needed sample sizes for all outcome measures, most likely resulting from the refinement of the total population to those who manifested prodromal AD. For the majority of enrichment strategies, the CDR-sb continued to require the fewest participants. Also

consistent was the lower required sample sizes for the ADAS12, relative to the ADAS11 [4,54]. The single exception to this, and to the substantially lower requirements for the CDR-sb than every other outcome measure by every other enrichment strategy, was in the setting of enrichment for FDG PET hypometabolism. When the MCI population was enriched for FDG PET hypometabolism, the ADAS11 required fewer participants than did the ADAS12 and the MMSE required fewer participants than did the CDR-sb (Table 6).

This study has limitations. It is derived entirely from a single data set, which has been used in a large number of studies with similar objectives [38,41,55,56,58] see also [8,65]. Within ADNI, subjects are well-educated, primarily Caucasian, and hold favorable attitudes toward research that may in part result from a high prevalence of a family history of AD. Further work modeling predementia trials based on alternate data sources is necessary. The conduct of ADNI-like studies on other continents may present such an opportunity.

We performed no formal comparisons of sample size outputs. Confidence intervals of the estimates are provided but, as has been seen in other studies [26,41], are wide. We also did not incorporate slope models into our study, focusing instead on change-from-baseline calculations. This decision was based on the on-going debate regarding the appropriate means of incorporating slope analyses into sample size estimates [18,55] and the fact that mean change from baseline is the general practice in AD registration trials. Our methods are also in contrast to the often-used practice of choosing a minimal clinically significant difference (for example, 2 points on the ADAS-cog) and powering a trial to detect such a difference at a given time point. Importantly, considering most of the trial scenarios in our results, the 25% drug effect would not achieve a clinically significant difference (for example, 25% of the mean decline for the ADAS11 at 24-months among the overall MCI population was 0.6 points). It is also true that the 25% drug effect would vary among the examined scenarios, as the overall rate of cognitive decline will vary among the different enriched populations. Thus, we do not propose that sample size decisions for predementia trials be based solely on the results of the current study. Rather, we believe that these results may be useful in considering predementia trial design choices, including primary and secondary outcome measures, enrichment strategies, and trial length.

Our analyses compared single biomarker modalities as enrichment strategies. Others have performed more integrated methods of enrichment as predictors of cognitive decline in the nondemented ADNI population. For example, McEvoy and colleagues showed that enriching for a composite measure of atrophy based on multiple brain regions resulted in a greater reduction in necessary MCI trial sample sizes (for trials using the CDR-SB or the ADAS11) than did genetic enrichment for ApoE genotype [41]. Similarly, Vemuri and colleagues showed that a composite index of structural abnormalities on MRI better predicted clinical progression on the CDR-SB than did clinical or CSF measures in MCI patients [64]. Using clinically available assays, Heister and colleagues demonstrated that combined use of hippocampal volume measures and psychometric testing better predicted conversion from MCI to dementia than did either volumetrics or cognitive testing alone or in combination with CSF measures [25].

Finally, our results consider only the number of participants that must complete an AD prevention trial. We did not examine the important variables of screen failures, participant recruitment, or patient attrition, all of which have significant impact on trial efficiency, perhaps especially in the setting of asymptomatic AD trials. As is seen in Table 2, each enrichment strategy is associated with a high screen failure rate, ranging from 80% for CSF tTau to 51% for the ratio of pTau/Aβ in the CN population. Thus, in the scenarios that we examined, the number of total CN or MCI participants that would need to be recruited to undergo biomarker testing is frequently quite high. For example, a trial using the RAVLT as

a single primary outcome enrolling CN participants who meet CSF Aβ criteria would need to enroll 5,450 participants to achieve the necessary sample sizes per study arm (not factoring in study attrition). Were the same study to use the CDR-sb as a single primary outcome, 8,550 participants would need to be screened. Though CSF criteria were frequently met by the CN ADNI population and have been shown to predict disease progression [16,58], asymptomatic participants often cite lumbar puncture as a barrier to trial participation and rated it as the diagnostic modality that they were least likely to be willing to endure in the setting of an AD prevention trial (unpublished results), suggesting that achieving recruitment goals in predementia trials may meet challenges. It is also unclear how participants will interpret the information of being eligible (or not) for preclinical AD trials. Thus, no matter what enrichment strategy might be chosen for use in preclinical trials, educational campaigns to facilitate recruitment and protect the welfare of participants will be critical to ensure the successful and ethical conduct of these trials.

In conclusion, our data fall short of suggesting a specific biomarker enrichment strategy as optimal for the design of preclinical AD trials. The CDR-sb was the seemingly ideal outcome measure when considering trials of MCI populations, whether they are enriched for AD biomarkers or not. The ideal outcome measures for trials of asymptomatic participants remain open to debate, though these results suggest that the RAVLT total score and CDR-sb may preferable to the ADAS11, ADAS12, MMSE, or RAVLT delayed recall. Replication of these results should be pursued in independent datasets and a variety of ongoing studies, including the next phase of ADNI, will contribute to the overall understanding of AD biomarkers and their utility in the setting of AD clinical trials.

Acknowledgments

References

1. Aisen PS. Alzheimer's disease therapeutic research: the path forward. Alzheimers Res Ther. 2009; 1(1):2. [PubMed: 19674435]

2. Aisen PS. Pre-dementia Alzheimer's trials: overview. J Nutr Health Aging. 2010; 14(4):294. [PubMed: 20305998]

3. 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; 76(3):280–6. [PubMed: 21178097]

4. Aisen PS, Petersen RC, Donohue MC, Gamst A, Raman R, Thomas RG, Walter S, Trojanowski JQ, Shaw LM, Beckett LA, Jack CR Jr, Jagust W, Toga AW, Saykin AJ, Morris JC, Green RC, Weiner

MW. Clinical Core of the Alzheimer's Disease Neuroimaging Initiative: progress and plans. Alzheimers Dement. 2010; 6(3):239–46. [PubMed: 20451872]

5. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. The diagnosis of mild cognitive impairment due to 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):270–9. [PubMed: 21514249]

6. Apostolova LG, Dutton RA, Dinov ID, Hayashi KM, Toga AW, Cummings JL, Thompson PM. Conversion of mild cognitive impairment to Alzheimer disease predicted by hippocampal atrophy maps. Arch Neurol. 2006; 63(5):693–9. [PubMed: 16682538]

7. Apostolova LG, Mosconi L, Thompson PM, Green AE, Hwang KS, Ramirez A, Mistur R, Tsui WH, de Leon MJ. Subregional hippocampal atrophy predicts Alzheimer's dementia in the cognitively normal. Neurobiol Aging. 2010; 31(7):1077–88. [PubMed: 18814937]

8. Ard MC, Edland SD. Power calculations for clinical trials in Alzheimer's disease. J Alzheimers Dis. 2011; 26 (Suppl 3):369–77. [PubMed: 21971476]

9. Chen K, Langbaum JB, Fleisher AS, Ayutyanont N, Reschke C, Lee W, Liu X, Bandy D, Alexander GE, Thompson PM, Foster NL, Harvey DJ, de Leon MJ, Koeppe RA, Jagust WJ, Weiner MW, Reiman EM. Twelve-month metabolic declines in probable Alzheimer's disease and amnestic mild cognitive impairment assessed using an empirically pre-defined statistical region-of-interest: findings from the Alzheimer's Disease Neuroimaging Initiative. Neuroimage. 2010; 51(2):654–64. [PubMed: 20202480]

10. Chou Y, Lepore N, Saharan P, Madsen S, Hua X, Jack C, Shaw L, Trojanowski J, Weiner M, Toga A, Thompson P. Initiative tAsDN. Ranking the Clinical and Pathological Correlates of Ventricular Expansion Mapped in 804 Alzheimer's Disease, MCI, and Normal Elderly Subjects. Neurobiology of Aging, Special Issue on ADNI. 2010; 31(8):1386–400.

11. Chou YY, Lepore N, de Zubicaray GI, Carmichael OT, Becker JT, Toga AW, Thompson PM. Automated ventricular mapping with multi-atlas fluid image alignment reveals genetic effects in Alzheimer's disease. Neuroimage. 2008; 40(2):615–30. [PubMed: 18222096]

12. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993; 261(5123):921–3. [PubMed: 8346443]

13. Cumming G, Finch C. A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions. Educational and Psychological Measurement. 2001; 61:43.

14. Cummings J. What can be inferred from the interruption of the semagacestat trial for treatment of Alzheimer's disease? Biol Psychiatry. 2010; 68(10):876–8. [PubMed: 21035622]

15. DeKosky ST. Maintaining adherence and retention in dementia prevention trials. Neurology. 2006; 67(9 Suppl 3):S14–6. [PubMed: 17101928]

16. Desikan RS, McEvoy LK, Thompson WK, Holland D, Roddey JC, Blennow K, Aisen PS, Brewer JB, Hyman BT, Dale AM. Amyloid-beta associated volume loss occurs only in the presence of phospho-tau. Ann Neurol. 2011; 70(4):657–61. [PubMed: 22002658]

17. Dickerson BC, Stoub TR, Shah RC, Sperling RA, Killiany RJ, Albert MS, Hyman BT, Blacker D, Detoledo-Morrell L. Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. Neurology. 2011; 76(16):1395–402. [PubMed: 21490323]

18. Donohue MC, Gamst AC, Aisen PS. Requiring an amyloid-beta1-42 biomarker for prodromal Alzheimer's disease or mild cognitive impairment does not lead to more efficient clinical trials. Alzheimers Dement. 2011; 7(2):245–6. author reply 7–9. [PubMed: 21414558]

19. Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, Delacourte A, Frisoni G, Fox NC, Galasko D, Gauthier S, Hampel H, Jicha GA, Meguro K, O'Brien J, Pasquier F, Robert P, Rossor M, Salloway S, Sarazin M, de Souza LC, Stern Y, Visser PJ, Scheltens P. Revising the definition of Alzheimer's disease: a new lexicon. Lancet Neurol. 2010; 9(11):1118–27. [PubMed: 20934914]

20. Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, Delacourte A, Galasko D, Gauthier S, Jicha G, Meguro K, O'Brien J, Pasquier F, Robert P, Rossor M, Salloway

S, Stern Y, Visser PJ, Scheltens P. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 2007; 6(8):734–46. [PubMed: 17616482]

21. Feldman HH, Ferris S, Winblad B, Sfikas N, Mancione L, He Y, Tekin S, Burns A, Cummings J, del Ser T, Inzitari D, Orgogozo JM, Sauer H, Scheltens P, Scarpini E, Herrmann N, Farlow M, Potkin S, Charles HC, Fox NC, Lane R. Effect of rivastigmine on delay to diagnosis of Alzheimer's disease from mild cognitive impairment: the InDDEx study. Lancet Neurol. 2007; 6(6):501–12. [PubMed: 17509485]

22. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12(3):189–98. [PubMed: 1202204]

23. Fox NC, Cousens S, Scahill R, Harvey RJ, Rossor MN. Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects. Arch Neurol. 2000; 57(3):339–44. [PubMed: 10714659]

24. Green RC, Schneider LS, Amato DA, Beelen AP, Wilcock G, Swabb EA, Zavitz KH. Effect of Tarenflurbil on Cognitive Decline and Activities of Daily Living in Patients With Mild Alzheimer Disease: A Randomized Controlled Trial. JAMA. 2009; 302(23):2557–64. [PubMed: 20009055]

25. Heister D, Brewer JB, Magda S, Blennow K, McEvoy LK. Predicting MCI outcome with clinically available MRI and CSF biomarkers. Neurology. 2011; 77(17):1619–28. [PubMed: 21998317]

26. Holland D, Brewer JB, Hagler DJ, Fennema-Notestine C, Dale AM. Subregional neuroanatomical change as a biomarker for Alzheimer's disease. Proc Natl Acad Sci U S A. 2009; 106(49):20954–9. [PubMed: 19996185]

27. Hua X, Lee S, Yanovsky I, Leow AD, Chou YY, Ho AJ, Gutman B, Toga AW, Jack CR Jr, Bernstein MA, Reiman EM, Harvey DJ, Kornak J, Schuff N, Alexander GE, Weiner MW, Thompson PM. Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: an ADNI study of 515 subjects. Neuroimage. 2009; 48(4):668–81. [PubMed: 19615450]

28. Jack CR Jr, Shiung MM, Gunter JL, O'Brien PC, Weigand SD, Knopman DS, Boeve BF, Ivnik RJ, Smith GE, Cha RH, Tangalos EG, Petersen RC. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. Neurology. 2004; 62(4):591–600. [PubMed: 14981176]

29. Jack CR Jr, Wiste HJ, Vemuri P, Weigand SD, Senjem ML, Zeng G, Bernstein MA, Gunter JL, Pankratz VS, Aisen PS, Weiner MW, Petersen RC, Shaw LM, Trojanowski JQ, Knopman DS. Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. Brain. 2010; 133(11):3336–48. [PubMed: 20935035]

30. Katz R. FDA: evidentiary standards for drug development and approval. NeuroRx. 2004; 1(3):307–16. [PubMed: 15717032]

31. Kohannim O, Hua X, Hibar DP, Lee S, Chou YY, Toga AW, Jack CR Jr, Weiner MW, Thompson PM. Boosting power for clinical trials using classifiers based on multiple biomarkers. Neurobiol Aging. 2010; 31(8):1429–42. [PubMed: 20541286]

32. Landau SM, Harvey D, Madison CM, Koeppe RA, Reiman EM, Foster NL, Weiner MW, Jagust WJ. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. Neurobiol Aging. 2009; 32(7):1207–18. [PubMed: 19660834]

33. Landau SM, Harvey D, Madison CM, Reiman EM, Foster NL, Aisen PS, Petersen RC, Shaw LM, Trojanowski JQ, Jack CR Jr, Weiner MW, Jagust WJ. Comparing predictors of conversion and decline in mild cognitive impairment. Neurology. 2010; 75(3):230–8. [PubMed: 20592257]

34. Leber P. Observations and suggestions on antidementia drug development. Alzheimer Dis Assoc Disord. 1996; 10 (Suppl 1):31–5. [PubMed: 8876787]

35. Leung KK, Barnes J, Ridgway GR, Bartlett JW, Clarkson MJ, Macdonald K, Schuff N, Fox NC, Ourselin S. Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. Neuroimage. 2010; 51(4):1345–59. [PubMed: 20230901]

36. Leung KK, Clarkson MJ, Bartlett JW, Clegg S, Jack CR Jr, Weiner MW, Fox NC, Ourselin S. Robust atrophy rate measurement in Alzheimer's disease using multi-site serial MRI: tissue-specific intensity normalization and parameter selection. Neuroimage. 2010; 50(2):516–23. [PubMed: 20034579]

37. Lo RY, Hubbard AE, Shaw LM, Trojanowski JQ, Petersen RC, Aisen PS, Weiner MW, Jagust WJ. Longitudinal Change of Biomarkers in Cognitive Decline. Arch Neurol. 2011; 68(10):1257–66. [PubMed: 21670386]

38. Lorenzi M, Donohue M, Paternico D, Scarpazza C, Ostrowitzki S, Blin O, Irving E, Frisoni GB. Enrichment through biomarkers in clinical trials of Alzheimer's drugs in patients with mild cognitive impairment. Neurobiol Aging. 2010; 31(8):1443–51. 51 e1. [PubMed: 20541287]

39. Mangialasche F, Solomon A, Winblad B, Mecocci P, Kivipelto M. Alzheimer's disease: clinical trials and drug development. Lancet Neurol. 2010; 9(7):702–16. [PubMed: 20610346]

40. Mattsson N, Zetterberg H, Hansson O, Andreasen N, Parnetti L, Jonsson M, Herukka SK, van der Flier WM, Blankenstein MA, Ewers M, Rich K, Kaiser E, Verbeek M, Tsolaki M, Mulugeta E, Rosen E, Aarsland D, Visser PJ, Schroder J, Marcusson J, de Leon M, Hampel H, Scheltens P, Pirttila T, Wallin A, Jonhagen ME, Minthon L, Winblad B, Blennow K. CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. JAMA. 2009; 302(4):385–93. [PubMed: 19622817]

41. 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; 24(3):269–77. [PubMed: 20683184]

42. Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, Parikshak N, Hua X, Toga AW, Jack CR Jr, Schuff N, Weiner MW, Thompson PM. Automated 3D mapping of hippocampal atrophy and its clinical correlates in 400 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. Hum Brain Mapp. 2009; 30(9):2766–88. [PubMed: 19172649]

43. Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, Parikshak N, Hua X, Toga AW, Jack CR Jr, Weiner MW, Thompson PM. Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls. Neuroimage. 2008; 43(1):59–68. [PubMed: 18675918]

44. Morra JH, Tu Z, Apostolova LG, Green AE, Toga AW, Thompson PM. Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation. IEEE Trans Med Imaging. 2010; 29(1):30–43. [PubMed: 19457748]

45. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology. 1993; 43(11):2412–4. [PubMed: 8232972]

46. Morris JC, Roe CM, Grant EA, Head D, Storandt M, Goate AM, Fagan AM, Holtzman DM, Mintun MA. Pittsburgh Compound B imaging and prediction of progression from cognitive normality to symptomatic Alzheimer disease. Arch Neurol. 2009; 66(12):1469–75. [PubMed: 20008650]

47. Mosconi L, Mistur R, Switalski R, Tsui WH, Glodzik L, Li Y, Pirraglia E, De Santi S, Reisberg B, Wisniewski T, de Leon MJ. FDG-PET changes in brain glucose metabolism from normal cognition to pathologically verified Alzheimer's disease. Eur J Nucl Med Mol Imaging. 2009; 36(5):811–22. [PubMed: 19142633]

48. Okonkwo OC, Alosco ML, Griffith HR, Mielke MM, Shaw LM, Trojanowski JQ, Tremont G. Cerebrospinal fluid abnormalities and rate of decline in everyday function across the dementia spectrum: normal aging, mild cognitive impairment, and Alzheimer disease. Arch Neurol. 2011; 67(6):688–96. [PubMed: 20558388]

49. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol. 1999; 56(3):303–8. [PubMed: 10190820]

50. Quinn JF, Raman R, Thomas RG, Yurko-Mauro K, Nelson EB, Van Dyck C, Galvin JE, Emond J, Jack CR Jr, Weiner M, Shinto L, Aisen PS. Docosahexaenoic acid supplementation and cognitive decline in Alzheimer disease: a randomized trial. JAMA. 2010; 304(17):1903–11. [PubMed: 21045096]

51. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry. 1984; 141(11):1356–64. [PubMed: 6496779]

52. Sabbagh MN. Drug development for Alzheimer's disease: Where are we now and where are we headed? Am J Geriatr Pharmacother. 2009; 7(3):167–85. [PubMed: 19616185]

53. Samson K. NerveCenter: Phase III Alzheimer trial halted: Search for therapeutic biomarkers continues. Ann Neurol. 2010; 68(4):A9–A12. [PubMed: 20976761]

54. Sano M, Raman R, Emond J, Thomas RG, Petersen R, Schneider LS, Aisen PS. Adding delayed recall to the Alzheimer Disease Assessment Scale is useful in studies of mild cognitive impairment but not Alzheimer disease. Alzheimer Dis Assoc Disord. 2011; 25(2):122–7. [PubMed: 20921876]

55. Schneider LS, Kennedy RE, Cutter GR. Requiring an amyloid-beta1-42 biomarker for prodromal Alzheimer's disease or mild cognitive impairment does not lead to more efficient clinical trials. Alzheimers Dement. 2010; 6(5):367–77. [PubMed: 20813339]

56. Schott JM, Bartlett JW, Barnes J, Leung KK, Ourselin S, Fox NC. Reduced sample sizes for atrophy outcomes in Alzheimer's disease trials: baseline adjustment. Neurobiol Aging. 2010; 31(8):1452–62. 62, e1–2. [PubMed: 20620665]

57. Schott JM, Bartlett JW, Barnes J, Leung KK, Ourselin S, Fox NC. Reduced sample sizes for atrophy outcomes in Alzheimer's disease trials: baseline adjustment. Neurobiol Aging. 31(8): 1452–62. 62, e1–2. [PubMed: 20620665]

58. Schott JM, Bartlett JW, Fox NC, Barnes J. Increased brain atrophy rates in cognitively normal older adults with low cerebrospinal fluid Abeta1-42. Ann Neurol. 2010; 68(6):825–34. [PubMed: 21181717]

59. Schuff N, Woerner N, Boreta L, Kornfield T, Shaw LM, Trojanowski JQ, Thompson PM, Jack CR Jr, Weiner MW. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain. 2009; 132(Pt 4):1067–77. [PubMed: 19251758]

60. 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, Trojanowski JQ. Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. Ann Neurol. 2009; 65(4):403–13. [PubMed: 19296504]

61. 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]

62. St George-Hyslop PH, Morris JC. Will anti-amyloid therapies work for Alzheimer's disease? Lancet. 2008; 372(9634):180–2. [PubMed: 18640437]

63. Taylor, E. Psychological appraisal of children with cerebral deficits. Cambridge, MA: Harvard University Press; 1959.

64. Vemuri P, Wiste HJ, Weigand SD, Shaw LM, Trojanowski JQ, Weiner MW, Knopman DS, Petersen RC, Jack CR Jr. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. Neurology. 2009; 73(4):294–301. [PubMed: 19636049]

65. 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. 2011

66. Youden W. Index for rating diagnostic tests. Cancer. 1950:32–5. [PubMed: 15405679]

Table 1

Demographic and cognitive characteristics of populations at baseline.

Characteristic	Normal	MCI
N	215	364
Age, mean ± SD, years	76.0 ± 5.1	74.9 ± 7.1
% Female	47.9%	36.3%
Race, n (%)		
Asian	3 (1.4%)	9 (2.5%)
African American	14 (6.5%)	12 (3.3%)
Caucasian	198 (92.1%)	343 (94.2%)
Ethnicity, n (%) Hispanic	2 (0.95%)	12 (3.3%)
Education, mean ± SD, years	16.1 ± 2.8	15.8 ± 3.0
ApoE genotype, % carriers	27.0	53.8
MMSE, mean ± SD	29.1 ± 1.0	27.0 ± 1.8
CDR global, mean ± SD	0.0 ± 0.0	0.5 ± 0.0
CDR-sb, mean ± SD	0.0 ± 0.1	1.6 ± 0.9
ADAS11, mean ± SD	6.1 ± 2.9	11.5 ± 4.5
ADAS12, mean ± SD	8.9 ± 4.1	17.6 ± 6.1
RAVLT total, mean ± SD	43.6 ± 9.2	30.9 ± 9.2
RAVLT delay, mean ± SD	7.5 ± 3.7	2.9 ± 3.3

Table 2

Number of ADNI participants who met biomarker enrichment criteria at baseline.

Enrichment strategy	CN participants meeting criteria, n (%)	MCI participants meeting criteria, n (%)
ApoE ε4 carrier status	58 (27.0)	194 (53.3)
CSF Aβ <192 pg/mL	40 (38.1)	135 (67.5)
CSF tTau>93 pg/mL	20 (19.0)	78 (39.0)
CSF pTau>23 pg/mL	39 (37.1)	127 (63.5)
CSF ratio tTau/Aβ>.39	38 (36.2)	126 (63.0)
CSF ratio pTau/Aβ>.1	51 (48.6)	142 (71.0)
FDG-PET <1.14	15 (14.7)	78 (37.9)
Hippocampal volume <3070.5 mm^3	48 (22.3)	179 (49.2)
Lateral ventricle volume 28,216.5 mm^3	74 (34.4)	180 (49.4)

Table 3

Mean changes ± SD at 36 months on clinical outcome measures of the ADNI CN population.

Enrichment strategy	ADAS11	ADAS12	CDR-SB	MMSE	RAVLT- total score	RAVLT- delayed recall
None	−0.47 ± 3.35	−0.17 ± 4.14	0.23 ± 0.82	−0.10 ± 1.38	−3.77 ± 8.91	−0.74 ± 3.49
ApoE e4 carrier status	−0.06 ± 3.92	0.47 ± 4.86	0.32 ± 0.97	−0.27 ± 1.51	−5.45 ± 7.65	−1.45 ± 3.22
CSF Aβ	0.10 ± 3.89	0.56 ± 5.20	0.44 ± 1.14	−0.15 ± 1.15	−3.88 ± 8.08	−1.0 ± 3.38
CSF tTau	−0.08 ± 4.44	0.27 ± 5.93	0.50 ± 1.46	−0.22 ± 1.26	−3.83 ± 6.91	−0.61 ± 3.57
CSF pTau	0.04 ± 3.83	0.71 ± 5.10	0.32 ± 1.08	−0.21 ± 1.20	−5.21 ± 7.77	−1.03 ± 3.38
CSF ratio tTau/Aβ	−0.41 ± 3.65	0.14 ± 4.86	0.42 ± 1.17	−0.13 ± 1.10	−4.44 ± 7.80	−0.56 ± 3.46
CSF ratio pTau/Aβ	0.12 ± 3.50	0.83 ± 4.74	0.36 ± 1.01	−0.21 ± 1.23	−5.09 ± 7.55	−0.98 ± 3.23
FDG-PET	1.06 ± 7.64	1.39 ± 9.32	1.20 ± 2.44	−0.50 ± 1.78	−4.33 ± 13.38	−0.75 ± 3.33
Hippocampal volume	0.40 ± 3.73	1.02 ± 4.17	0.57 ± 1.18	−0.36 ± 1.58	−4.21 ± 11.15	−1.12 ± 4.13
Lateral ventricle volume	−0.23 ± 3.18	−0.18 ± 3.69	0.19 ± 0.50	−0.06 ± 1.64	−4.36 ± 10.48	−1.32 ± 2.68

Table 4

Required sample sizes per arm for 36-month CN trials

Enrichment strategy	Needed sample size per group for each primary outcome measure, n (CI)					
	ADAS11	ADAS12	CDR-SB	MMSE	RAVLT- total score	RAVLT- delayed recall
None	No Decline	No Decline	3058 (1365-12,874)	50,790 (5722-∞)	1404 (777-3462)	5534 (1947-51,494)
ApoE e4 carrier status	No Decline	27,380 (1748-∞)	2372 (679-157,700)	8146 (1193-∞)	499 (243-1659)	1237 (461-11,215)
CSF Aβ	420,495 (1844-∞)	21,373 (1247-∞)	1710 (461-630,800)	14,447 (1143-∞)	1090 (358- 17,523)	2874 (617-∞)
CSF tTau	No Decline	117,234 (934-∞)	2136 (367-∞)	8114 (617-∞)	817 (234-100,928)	8559 (636-∞)
CSF pTau	2,200,678 (2060-∞)	13,037 (1096-∞)	2901 (617-∞)	8542 (971-∞)	559 (234-3001)	2708 (598-∞)
CSF ratio tTau/Aβ	No Decline	301,390 (1748-∞)	1955 (487-∞)	19,451 (1193-∞)	776 (286-6990)	9525 (971-∞)
CSF ratio pTau/Aβ	214,455 (2183-∞)	8128 (1096-∞)	2025 (580-157,700)	8618 (1143-∞)	552 (253-2183)	2752 (679-∞)
FDG PET	13,136 (515-∞)	11,308 (487-∞)	1039 (195-∞)	3198 (350-∞)	2397 (312-∞)	4964 (405-∞)
Hippocampal volume	21,359 (1502-∞)	4164 (835-∞)	1057 (385-9857)	4883 (899-∞)	1759 (530-70,089)	3427 (751-∞)
Lateral ventricle volume	No Decline	No Decline	1686 (598-17,523)	108,591 (2804-∞)	1453 (546-11,215)	1039 (437-5214)

Table 5

Mean changes ± SD on clinical outcome measures at 24 months of the ADNI MCI population.

Enrichment strategy	ADAS11	ADAS12	CDR-SB	MMSE	RAVLT- total score	RAVLT- delayed recall
None	2.81 ± 5.73	3.64 ± 6.76	1.46 ± 1.98	−1.72 ± 3.51	−3.04 ± 7.66	−0.65 ± 2.62
ApoE e4 carrier status	3.07 ± 5.84	4.01 ± 6.94	1.79 ± 2.04	−2.35 ± 3.68	−4.25 ± 7.38	−0.95 ± 2.51
CSF Aβ	3.73 ± 5.95	4.76 ± 6.95	1.84 ± 1.95	−2.23 ± 3.63	−3.94 ± 7.57	−0.89 ± 2.73
CSF tTau	4.51 ± 6.60	5.71 ± 7.45	1.94 ± 2.09	−2.67 ± 3.88	−4.27 ± 7.77	−1.14 ± 2.40
CSF pTau	3.62 ± 6.09	4.64 ± 7.05	1.84 ± 2.0	−2.27 ± 3.70	−4.10 ± 7.65	−0.88 ± 2.37
CSF ratio tTau/Aβ	3.69 ± 6.05	4.70 ± 7.0	1.92 ± 1.95	−2.38 ± 3.70	−3.97 ± 7.59	−0.95 ± 2.30
CSF ratio pTau/Aβ	3.58 ± 5.96	4.58 ± 6.92	1.77 ± 1.98	−2.16 ± 3.60	−3.98 ± 7.52	−0.81 ± 2.35
FDG-PET	3.60 ± 4.29	4.49 ± 5.40	1.75 ± 2.0	−2.92 ± 3.26	−4.61 ± 7.27	−0.70 ± 2.41
Hippocampal volume	2.88 ± 4.99	3.79 ± 5.81	1.78 ± 1.95	−2.25 ± 3.46	−4.37 ± 7.22	−0.76 ± 2.10
Lateral ventricle volume	3.44 ± 5.60	4.45 ± 6.67	1.72 ± 2.12	−2.14 ± 3.98	−3.88 ± 7.70	−0.76 ± 2.44

Table 6

Required sample sizes per arm for 24-month MCI trials

Enrichment strategy	Needed sample size per group for each primary outcome measure, n (CI)					
	ADAS11	ADAS12	CDR-SB	MMSE	RAVLT- total score	RAVLT- delayed recall
None	1046 (679-1844)	868 (580-1431)	458 (334-679)	1044 (679-1844)	1600 (971-3219)	4126 (1947-14,931)
ApoE e4 carrier status	908 (366-1947)	753 (341-1502)	329 (241-636)	616 (312-1143)	758 (341-1502)	1747 (468-5214)
CSF Aβ	639 (376-1365)	536 (326-1096)	282 (188-474)	667 (385-1431)	919 (501-2317)	1794 (805-7788)
CSF tTau	537 (280-1431)	428 (234-1051)	292 (173-617)	529 (280-1431)	832 (385-3001)	1125 (474-5214)
CSF pTau	714 (405-1659)	579 (342-1247)	296 (280-515)	665 (376-1502)	875 (474-2317)	1835 (777-8731)
CSF ratio tTau/Aβ	676 (385-1502)	557 (326-1193)	258 (173-437)	609 (358-1304)	917 (487-2465)	1457 (679-5214)
CSF ratio pTau/Aβ	696 (404-1502)	574 (342-1192)	313 (209-530)	702 (405-1577)	898 (487-2183)	2089 (899-9857)
FDG PET	357 (198-866)	363 (202-866)	325 (185-751)	314 (179-725)	626 (305-1947)	2977 (835-280,356)
Hippocampal volume	754 (449-1577)	588 (367-1143)	300 (205-487)	598 (367-1143)	687 (415-1431)	1891 (899-6990)
Lateral ventricle volume	666 (395-1365)	565 (350-1096)	381 (253-657)	869 (501-1947)	991 (546-2317)	2605 (1844-12,874)