

Featured Articles

Requiring an amyloid- β_{1-42} biomarker for prodromal Alzheimer's disease or mild cognitive impairment does not lead to more efficient clinical trials

Lon S. Schneider^{a,*}, Richard E. Kennedy^b, Gary R. Cutter^b; and the Alzheimer's Disease Neuroimaging Initiative

^aDepartments of Psychiatry and Neurology, Keck School of Medicine, and Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

^bSchool of Public Health, University of Alabama, Birmingham, AL, USA

Abstract

Background: Low cerebrospinal fluid (CSF) amyloid- β_{1-42} concentration and high total-tau/A β_{1-42} ratio have been recommended to support the diagnosis of prodromal Alzheimer's disease (AD) in patients with amnesic mild cognitive impairment (aMCI) and also to select patients for clinical trials (Shaw et al, *Ann Neurol* 2009;65:403–13; Dubois et al, *Lancet Neurol* 2007;6:734–46).

Methods: We tested this recommendation with clinical trials simulations using patients from the Alzheimer Disease Neuroimaging Initiative who fulfilled the following entry criteria: (1) aMCI, (2) aMCI with CSF A $\beta_{1-42} \leq 192$ mg/mL, (3) and aMCI with total-tau/A $\beta_{1-42} > 0.39$. For each criterion, we randomly resampled the database obtaining samples for 1000 trials for each trial scenario, planning for 1 or 2 year trials with samples from 50 to 400 patients per treatment or placebo group, with up to 40% dropouts, outcomes after using the AD assessment scale-cognitive subscale and clinical dementia rating scale with effect sizes ranging from 0.15 to 0.75, and calculated statistical power.

Findings: Approximately 70% to 74% of aMCI patients with CSF measures met biomarker criteria. The addition of the low A β_{1-42} or high tau/A β_{1-42} requirement resulted in minimal or no increase in the power of the trials compared with enrolling aMCI without requiring the biomarker criteria. Slightly larger mean differences between the placebo and treatment groups fulfilling biomarker criteria were offset by increased outcome variability within the groups.

Interpretations: Although patients with aMCI or patients with prodromal AD meeting CSF biomarkers criteria were slightly more cognitively impaired and showed greater decline than patients with aMCI diagnosed without considering the biomarkers, the requirement of biomarker-positive patients would most likely not result in more efficient clinical trials, and trials would take longer because fewer patients would be available. A CSF A β_{1-42} marker, however, could be useful as an explanatory variable or covariate when warranted by the action of a drug.

© 2010 The Alzheimer's Association. All rights reserved.

Keywords:

Alzheimer disease; Mild cognitive impairment; Biomarkers; Clinical trials; Simulations; Amyloid-beta protein; Alzheimer's disease neuroimaging initiative (ADNI); Alzheimer's disease assessment scale; Clinical dementia rating

1. Background

Using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), researchers concluded that low cerebro-

spinal fluid (CSF) amyloid- β_{1-42} concentrations or high total-tau protein to A β_{1-42} ratios in patients diagnosed with amnesic mild cognitive impairment (aMCI) predicted progression to Alzheimer's dementia (e.g., 89% with high t-tau/A β_{1-42} ratio within 1 year) [1]. Those researchers suggested that biomarkers could be used as clinical trials entry requirements to improve the efficiency and reduce the

Study registration: ClinicalTrials.gov Identifier: NCT00106899.

*Corresponding author. Tel.: 323-442-7600; Fax: 323-442-7601.

E-mail address: lschneid@usc.edu

sample sizes of trials. Some others have recommended CSF $A\beta_{1-42}$ markers for this purpose as well [2–5], and at least one pharmaceutical manufacturer required similar biomarker criteria to support a prodromal Alzheimer's disease (AD) diagnosis in a targeted design clinical trial [6].

Research criteria for prodromal or early AD, that is, AD before the onset of dementia, have been proposed on the basis of the fulfillment of the criteria for the equivalent of aMCI by a patient, with an episodic memory deficit that fails to improve on cueing, and a biomarker associated with AD that could be identified and measured by either brain imaging or CSF $A\beta_{1-42}$ or tau protein assays [2]. Moreover, ad hoc groups have recommended that future clinical trials for prodromal AD could be made more efficient by introducing the requirement for a CSF $A\beta_{1-42}$ biomarker [1,3]. Relying on the ADNI database [5,7], researchers calculated that, to demonstrate a 40% reduction in progression on clinical ratings, with 80% power, an alpha error $P \leq .05$, and a 2-year drop-out rate of <40%, about 100 or 150 patients would be required for one or another primary outcome per drug and placebo group for those who are selected by using the biomarker criteria as compared with twice as many without the biomarker criteria.

We empirically tested the potential efficiency of these recommendations by statistically simulating a range of clinical trials scenarios with aMCI patients with or without biomarker inclusion criteria using the same database on which the CSF biomarker recommendations were based [1].

2. Methods

2.1. ADNI study overview and participants

ADNI is a natural history, nontreatment, observational study aimed at setting standards for brain imaging studies and biomarkers for diagnosis and treatment trials [7]. A total of 59 sites, which were mostly academic, recruited 188 participants with mild AD (i.e., mini-mental state examination [MMSE] scores ranging from 21 to 26), 405 with mild cognitive impairment (MCI) (MMSE ranging from 24 to 30), and 229 with no cognitive impairment, who were followed up with regular assessments [5,7].

The MCI inclusion criteria, detailed in previously published data [5,7], are identical to criteria for the MCI of the amnesic-type used in previous MCI clinical trials on cholinesterase inhibitors [8,9], which required a clinical dementia rating (CDR) [10] score of 0.5 with the memory box scored at 0.5 or greater, and delayed recall from the logical memory II subscale of the Wechsler memory scale—revised [11] to be ≤ 8 for 16 years of education, ≤ 4 for 8 to 15 years, or ≤ 2 for 0 to 7 years. Patients were required to be largely intact with regard to general cognition and functional performance, and could not qualify for a dementia diagnosis. As in most current clinical trials, participants could continue using marketed anti-dementia drugs if they had been on stable doses for at least 4 weeks before entry [7].

The main imaging and biomarkers include brain magnetic resonance imaging, positron-emission tomography, and CSF $A\beta$ and tau protein concentrations [5,7]. The main clinical ratings reflected the following clinical trials outcomes: the Alzheimer's Disease Assessment Scale—cognitive subscale (ADAS-cog) [12], CDR [10], MMSE [13], and functional activities questionnaire [14]. Clinical assessments were carried out at 6-month intervals over the first 2 years.

The ADAS-cog [11] evaluates memory, reasoning, language, orientation, praxis, language, and word finding difficulty, and is scored from 0 to 70 errors. The CDR [10] is used to rate impairment (from 0 = not impaired to 3 = severely impaired) in each of the following six categories: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care; and are summed into the CDR sum of the boxes score (CDR-sb) as a severity measure ranging from 0 to 18.

2.2. Simulation methods

Simulations were conducted under a detailed protocol [15] to reflect typical clinical trials for an experimental drug for aMCI or early AD with one treatment and placebo group, 1:1 allocation ratio, and parameters selected to be consistent with previously published trials [8,9]. For each trial scenario, a separate set of patients was constructed by randomly choosing from the ADNI dataset with replacement, that is, patients from the dataset could be present in the simulated groups for more than one occasion. Sample sizes of 50, 100, 200, and 400 per group were used; 12- and 24-month-long trials were considered; the ADAS-cog and CDR-sb were the primary outcomes. The placebo group outcome was the score for the patient at the specified time point in the ADNI database. For the treatment group, a range of effect sizes ranging from 0.15 to 0.75 with 0.10 increments were used to compute an expected treatment effect (or slowing down of decline) reflecting very small to moderately large effect sizes [16]. For each patient, an individual treatment effect was randomly generated from a χ^2 distribution with a mean equal to the expected treatment effect to allow for a more realistic distribution of declines over time, where a few patients may fail or worsen more markedly than would be predicted by a normal distribution. The individual treatment effect was shifted by subtracting two times the expected treatment effect, then adding the resultant to the patient's score at the specified time point in the database. Therefore, even when a patient was reused in the analysis, the actual value used would be modified by this randomly selected amount in the treatment arm. In the placebo arm, use of the same patient would lead to a slight underestimation of the variance, thereby slightly improving the statistical power examined later in the text. Dropout rates of 20% and 40% in both the treatment and placebo groups were incorporated into the scenarios.

2.3. Selection criteria

Patients were selected for the samples in a way similar to when they apply for clinical trials using the following three sets of inclusion criteria: (1) aMCI diagnosis at screening with the logical memory II delayed memory deficit defined previously; (2) aMCI with CSF $A\beta_{1-42} \leq 192$ pg/mL; and (3) aMCI with $t\text{-tau}/A\beta_{1-42} > 0.39$. The latter two criteria were specifically recommended as aforementioned [1], fulfill newly proposed research criteria for AD [2], and are consistent with criteria for a commercial prodromal AD clinical trial [6].

2.4. Statistical analysis

The primary analyses were conducted using a mixed effects linear model (covariance pattern model), which adjusts for missing data to test for differences between the baseline and endpoint values for the treatment and placebo groups. The mixed effects model minimizes bias and better controls for type I error in the presence of missing data [17]. For each set of inclusion criteria, a full model was constructed with group effect, visit effect, and group-by-visit interactions, with age and gender as covariates, and a reduced model with visit, age, and gender effects. A compound symmetric covariance structure was used to model the correlation between visits for each participant. Parameters were estimated using maximum likelihood. Probability values for the group (treatment) effect were calculated using twice the difference in the logarithmic likelihood of the full and reduced model, which follows a χ -squared distribution with the appropriate degrees of freedom. Secondary analyses examined last ob-

servation carried forward samples to impute missing values and complete cases using the nonparametric Wilcoxon test to detect any differences between treatment and placebo groups as a result of the skewed distributions of the outcomes. For all analyses, the missing data pattern present in the ADNI database was used to realistically simulate drop-outs; observations were missing in simulated datasets in cases where they were originally missing in the ADNI database.

One thousand simulations were carried out for each scenario so that estimates of power could be obtained up to three digits. Power was calculated as the proportion of 1000 simulated trials per trial scenario having a P value $\leq .05$. Analyses were performed using version 2.10.1 of the R programming environment [18]. Mixed model analyses were performed using version 3.1-89 of the nlme package for R [19]. The database was downloaded on December 7, 2009 from <http://www.loni.ucla.edu/twiki/bin/view/ADNI/ADNIClinicalFAQ>.

3. Results

3.1. Patient characteristics

The three sets of inclusion criteria were similar with respect to demographic and clinical characteristics, predominantly, patients were Caucasian, male, married, highly educated, and 57.7% of those whose history was ascertained had a family history of dementia (Table 1).

Overall, 54.0% of the patients were *APOE* $\epsilon 4$ genotype carriers (one or two alleles) and 64% to 66% of the biomarker-positive patients were *APOE* $\epsilon 4$ genotype

Table 1
Clinical characteristics and ratings among groups based on different selection criteria

	aMCI	aMCI + low $A\beta_{1-42}$	aMCI + high $t\text{-tau}/A\beta_{1-42}$	<i>P</i> value
N	400	148	137	
Age (years), mean (SD)	74.92 (7.41)	74.66 (7.09)	74.66 (7.45)	.84
Gender, male %	64.5%	64.9%	62.8%	.92
Race, Caucasian %	93.5%	96.6%	96.4%	.93
Marital status, married %	80.3%	84.5%	81.8%	.95
Education, college %	64.3%	62.8%	62.0%	.92
<i>APOE</i> $\epsilon 4$ genotype %	54.0%	64.2%	66.0%	.17
MMSE, screening, mean (SD)	27.01 (1.78)	26.79 (1.79)	26.83 (1.82)	.32
CDR-sb, screening, mean (SD)	1.61 (0.88)	1.65 (0.91)	1.63 (0.89)	.95
CDR-sb, 6 mo, mean (SD)	1.97 (1.37)	2.09 (1.26)	2.06 (1.23)	.30
CDR-sb, 12 mo, mean (SD)	2.27 (1.52)	2.51 (1.39)	2.51 (1.42)	.03
CDR-sb, 24 mo, mean (SD)	3.06 (2.23)	3.44 (2.14)	3.49 (2.15)	.03
ADAS-cog, baseline, mean (SD)	11.56 (4.42)	12.25 (4.54)	12.41 (4.55)	.07
ADAS-cog, 6 mo, mean (SD)	12.32 (5.51)	13.39 (5.31)	13.42 (5.25)	.02
ADAS-cog, 12 mo, mean (SD)	12.55 (6.19)	13.34 (5.93)	13.59 (5.92)	.06
ADAS-cog, 24 mo, mean (SD)	14.12 (7.43)	15.76 (7.08)	15.85 (7.12)	.01
Dementia, 24 mo, mean (SD)	28.5%	35.8%	38.0%	.23

Abbreviations: aMCI = amnesic mild cognitive impairment; low $A\beta_{1-42}$ = cerebrospinal fluid levels < 192 ng/mL; high $t\text{-tau}/A\beta_{1-42}$ = CSF ratio > 0.39 ; *APOE* $\epsilon 4$ genotype = % with 1 or 2 alleles; MMSE = Mini-mental State Examination; CDR-sb = clinical dementia rating-sum of boxes score; ADAS-cog = AD Assessment Scale-cognitive.

NOTE. The groups with CSF measures are subsets of the amnesic MCI group and are defined by having a low CSF $A\beta_{1-42}$ concentration or high $t\text{-tau}/A\beta_{1-42}$ ratio in the 199 patients who received CSF analyses.

carriers. In all, 44.0% used cholinesterase inhibitors (84.9% donepezil), 9% used cholinesterase inhibitors and memantine, 2.4% used only memantine, and 53.5% used neither. Median duration of previous use of cholinesterase inhibitors was 0.97 years and for memantine it was 0.88 years [20].

Over 95% of the population was classified by investigators as MCI due to AD; 400 of 402 patients (99.5%) fulfilled criteria for aMCI. $A\beta_{1-42}$ and total tau levels were determined in 199 and 196 patients, respectively, from CSF samples taken at enrollment. Mean and median $A\beta_{1-42}$ levels were 163.6 ng/mL (SD = 54.8) and 146 ng/mL, respectively; 148 (74.4%) patients had CSF $A\beta_{1-42}$ levels of <192 pg/mL. Mean and median total tau levels were 103.6 ng/mL (SD = 60.8) and 87 ng/mL, respectively; 137 (69.9%) patients had t-tau/ $A\beta_{1-42}$ of >0.39.

There were statistical trends for the biomarker-positive groups to have slightly worse rating scale scores. Among

the aMCI patients, 95.5%, 90.0%, 80.6%, and 72.1% had outcomes available at 6-, 12-, 18- and 24-months, respectively; and 28.5% of them progressed to dementia at 24 months as compared with 35.8% of low $A\beta_{1-42}$ and 38.8% of high t-tau/ $A\beta_{1-42}$ patients ($P = .23$).

3.2. Outcomes

Patients showed considerable heterogeneity in their clinical course and within each diagnostic group. Plots depicting change in ADAS-cog and CDR-sb over time for random samples of 25 participants in the aMCI and low $A\beta_{1-42}$ diagnostic groups are available online in [supplemental figures](#).

Power calculations for the mixed model analyses for the three groups are shown across a range of effect and sample sizes providing for 20% and 40% dropouts (Figs. 1–4).

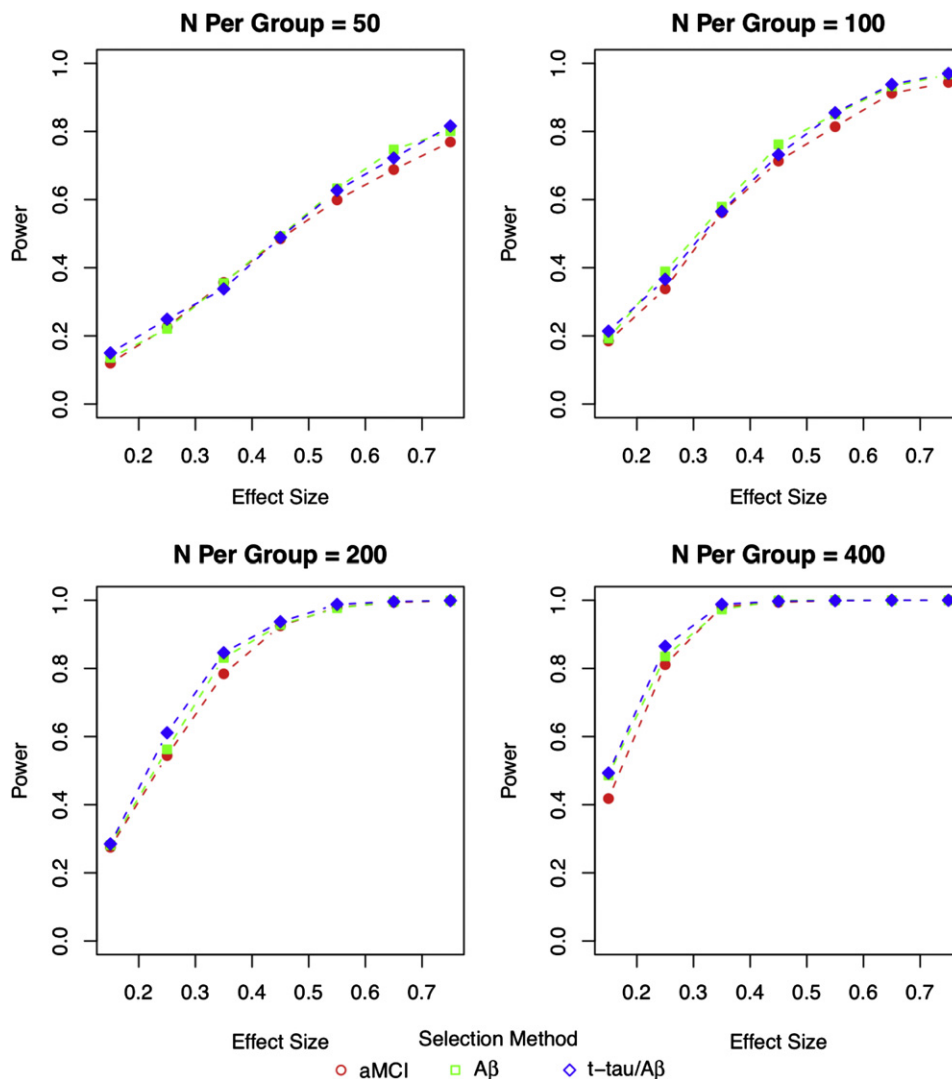


Fig. 1. Power for Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog) outcomes in 24-month-long trials with 20% dropouts. Power calculations for the ADAS-cog in patients with amnesic mild cognitive impairment (aMCI), with aMCI plus low cerebrospinal fluid (CSF) $A\beta_{1-42}$, and aMCI plus high CSF t-tau/ $A\beta_{1-42}$ by effect size and sample size. Addition of a biomarker selection criterion did not appreciably increase power under any of the scenarios. Simulation parameters included $\alpha = 0.05$, χ^2 random errors, and 20% dropouts with mixed model analysis for participants with missing data.

Power increased with increasing effect size and sample size, but there was generally little difference in power across the three inclusion criteria. For both the ADAS-cog and the CDR-sb, the requirement for biomarker criteria typically resulted in 2% to 5% increase in power, 7% in a few scenarios (Tables 2, 3; Figs. 1–4). Although there were greater mean differences between placebo and treatment groups with the biomarker criteria, there were also greater increases in variability which tended to reduce the effect sizes.

Analyses of 12-month trials using the aforementioned scenarios also did not show a meaningful difference in power for the outcomes across the selection criteria. For both the ADAS-cog and the CDR-sb, aMCI alone typically resulted in 2% to 4% increase in power as compared with the additional requirement for biomarker criteria. Secondary analyses of last observation carried forward and complete cases samples showed similar, very small differences between

diagnostic groups (Results are available online in [supplemental figures and tables](#)).

4. Discussion

ADNI was designed to provide information for future clinical trials and it is ideal for evaluating the benefits of the usage of CSF biomarkers [1,3,5,7]. The assessment of expert-proposed targeted trials designs for AD and the performance of $A\beta_{1-42}$ diagnostic or predictive biomarkers under experimental, clinical trials conditions have not been done previously. The results in this study provide an empirical estimation of the distribution and accuracy of clinical outcomes and potential biases for future AD trials that would use $A\beta_{1-42}$ biomarkers or a prodromal AD diagnosis as entry criteria [15]. The low $A\beta_{1-42}$ and the high t-tau/ $A\beta_{1-42}$ criteria when added to an aMCI diagnosis did not meaningfully

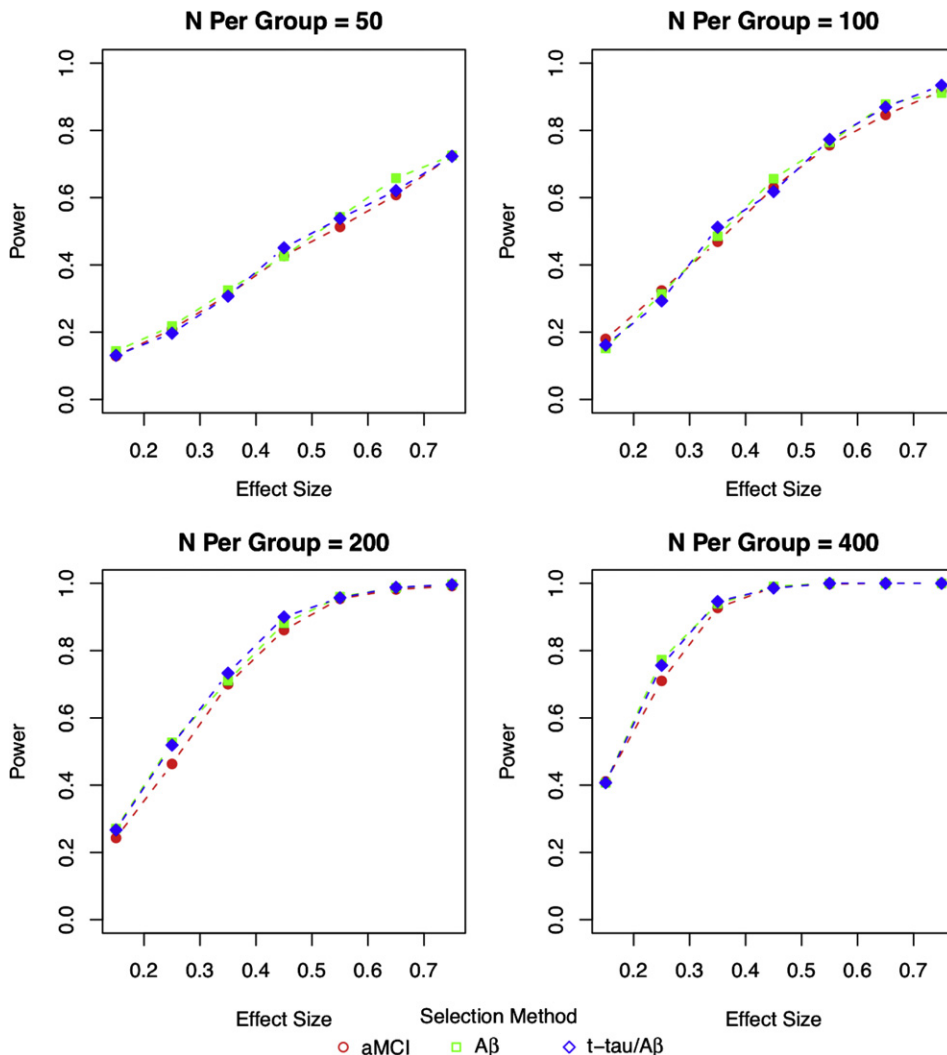


Fig. 2. Power for ADAS-cog outcomes in 24-month-long trials with 40% dropouts. Power calculations for the ADAS-cog in patients with aMCI, with aMCI plus low CSF $A\beta_{1-42}$, and aMCI plus high CSF t-tau/ $A\beta_{1-42}$ by effect size and sample size. Addition of a CSF biomarker selection criterion did not appreciably increase power under any of the scenarios. Simulation parameters included $\alpha = 0.05$, χ^2 random errors, and 40% dropouts with mixed model analysis for participants with missing data.

affect the efficiency of the trials as compared with the aMCI diagnosis alone.

In the more plausible trials scenarios of small effect sizes of 0.35 or less, 40% dropouts over 2 years, and 200 to 400 patients per group, the gain in power was typically 4% or less with either clinical outcome. This small gain must be weighed against the additional efforts of obtaining CSF, analyzing it, and excluding a proportion of aMCI patients. At least 26% of the ADNI aMCI patients who had lumbar punctures would not fulfill the biomarker criteria, increasing cost and time for recruitment by about one-third in exchange for very little or no gain in statistical power.

The considerable heterogeneity among biomarker-positive participants is a likely explanation for our results. Despite greater clinical worsening of about 0.8 ADAS-cog and 0.4 CDR-sb points over 2 years in the biomarker-positive groups as compared with the overall aMCI group

without regard to biomarkers, the standard deviations of the outcomes were larger, decreasing the power to detect treatment differences, that is, the within-group effect sizes were about the same. The use of these biomarker criteria for a targeted clinical trial may select from the extremes of the distribution, where increased within-group variability may offset any increase in mean difference between groups.

A notable difference between the ADAS-cog and CDR-sb outcomes was that the within-group effect sizes (i.e., mean change/SD of the change) were generally larger in the case of the CDR-sb. However, this did not translate to more efficient trials using the CDR-sb in preference to the ADAS-cog outcome in terms of treatment effects, power, and required sample sizes.

Longitudinal studies [21–24], including ADNI [1,25], demonstrate that CSF $A\beta_{1-42}$ and t-tau concentrations predict clinical progression in MCI patients; and subgroups

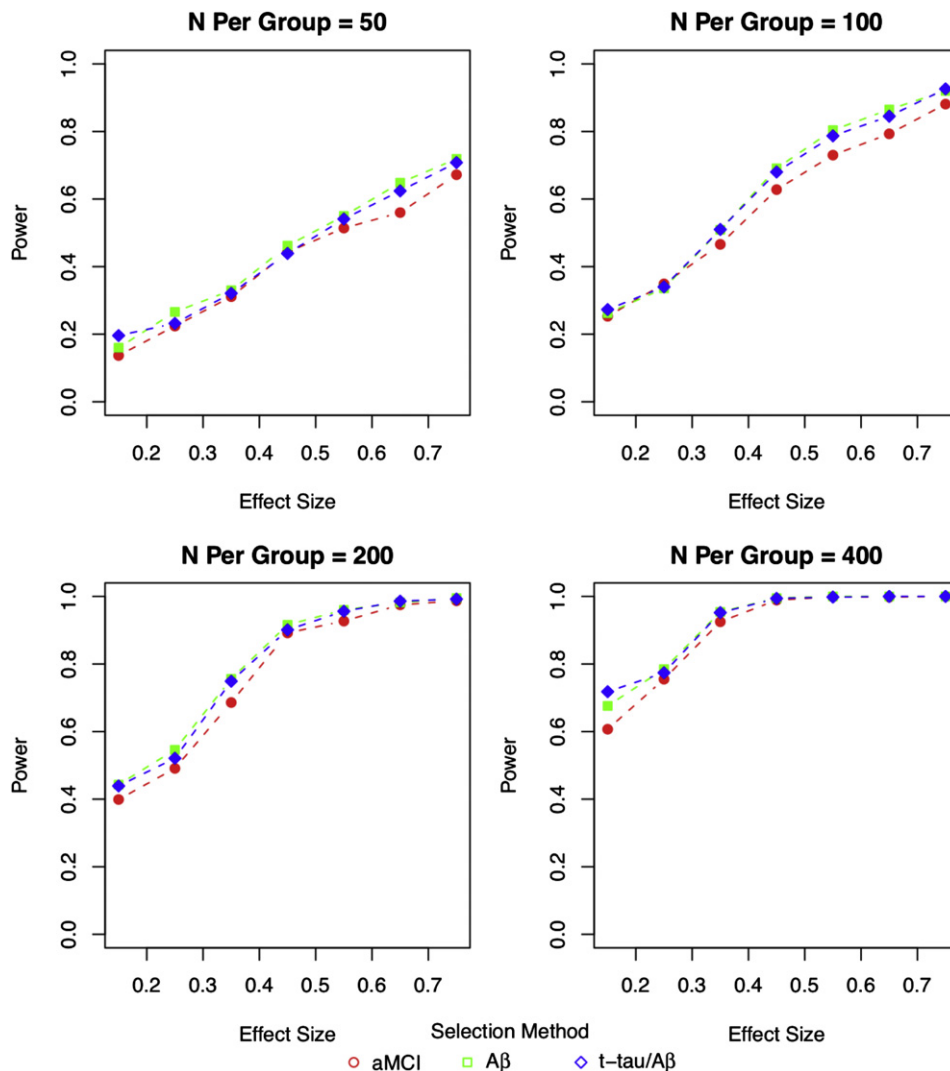


Fig 3. Power for clinical dementia rating sum of the boxes score (CDR-sb) outcomes in 24-month-long trials with 20% dropouts. Power calculations for the CDR-sb in patients with aMCI, aMCI plus low CSF $A\beta_{1-42}$, and aMCI plus high CSF t-tau/ $A\beta_{1-42}$ by effect size and sample size. Addition of the biomarkers resulted in very small increases in statistical power, mainly for smaller samples at large effect sizes. Simulation parameters included $\alpha = 0.05$, χ^2 random errors, and 20% dropouts with mixed model analysis for participants with missing data.

that progress to dementia at differential rates defined by CSF biomarkers can be identified [1,21,23–25]. Although a consistent finding [26], it has also been consistently observed that either the memory impairment or the CSF abnormalities provided approximately equal predicted clinical declines without differential sensitivity [24,26]. Similarly, in our analyses the biomarker-positive groups showed only fractional differences on mean baseline and changed scores as compared with the aMCI group selected without consideration of CSF biomarkers. Therefore, it appears that positive CSF biomarkers, when obtained in a clinical research environment after an aMCI diagnosis is made—and therefore perhaps in clinical practice as well—may mainly identify more advanced aMCI or prodromal AD; and if so, then cognitive severity appears to be the more pragmatic predictor of decline [24,26]. Further evidence for this is that the 148 aMCI

patients with low concentration values of CSF $A\beta_{1-42}$ scored significantly worse on screening ADAS-cog, CDR-sb, logical memory-delayed, and functional activities than the 51 patients with high CSF $A\beta_{1-42}$ (data not shown).

These results have substantial implications for clinical trials planning and interpretation. Assumptions that low CSF $A\beta_{1-42}$ or high t-tau/ $A\beta_{1-42}$ are more relevant selection criteria for clinical trials are based on views that they aid diagnosis and index greater brain $A\beta$ load and neurodegeneration [27]. However, it is not known whether such biomarker-positive patients would be more likely to respond to an experimental drug or whether a therapeutic effect will be detected more readily. The opposite could be true and targeted design trials that select only patients with $A\beta_{1-42}$ biomarkers may inadvertently select those who are less likely to benefit because they are too advanced. In fact, the use of CSF

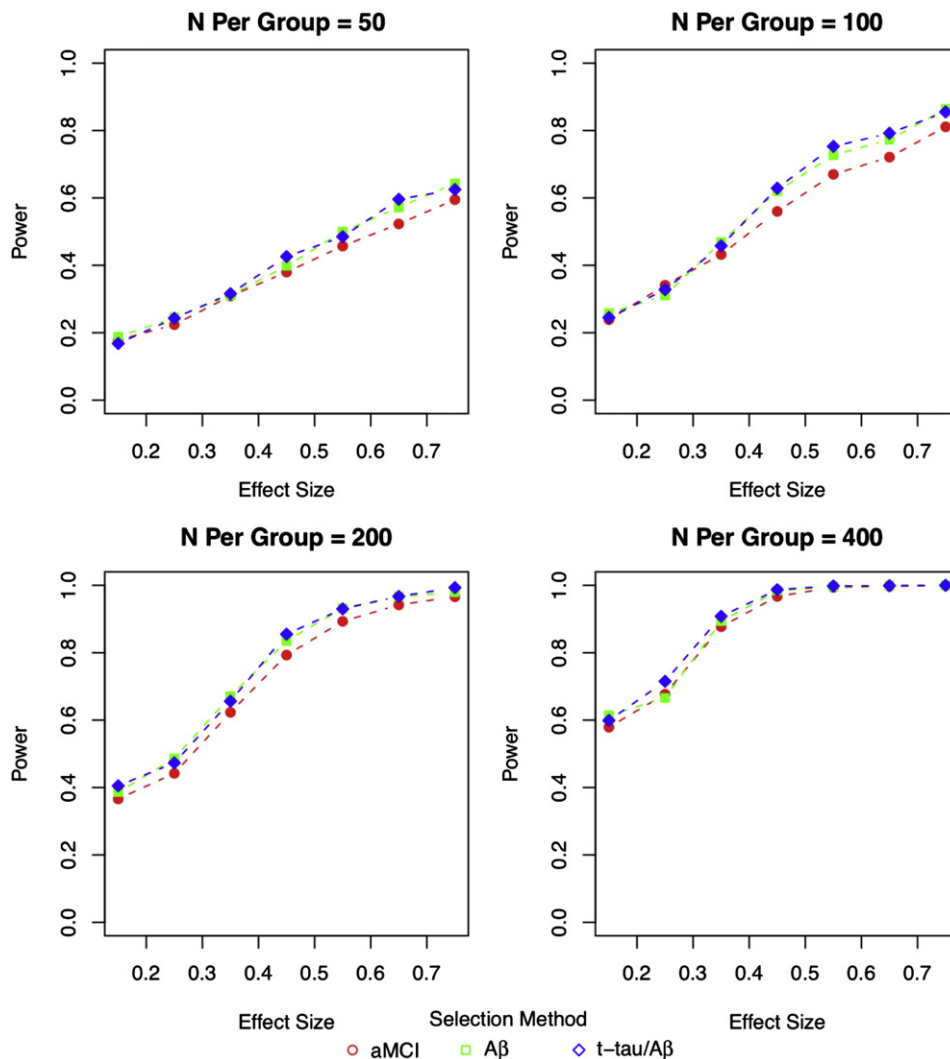


Fig. 4. Power for CDR-sb outcomes in 24-month-long trials with 40% dropouts. Power calculations for the CDR-sb in patients with aMCI, aMCI plus low CSF $A\beta_{1-42}$, and aMCI plus high CSF t-tau/ $A\beta_{1-42}$ by effect size and sample size. Addition of biomarkers resulted in only small increases in statistical power, mainly with smaller sample sizes and large effect sizes. Simulation parameters included $\alpha = 0.05$ with 24-month-long trials, χ^2 random errors, and 40% dropouts with mixed model analysis for participants with missing data.

$A\beta_{1-42}$ biomarkers after a clinical aMCI diagnosis is made may not achieve the desired goal of identifying prodromal AD patients early enough in their illness course for a disease-modifying drug to show an effect.

Moreover, the efficiency of a targeted clinical trial design where the premise is that the low CSF $A\beta_{1-42}$ patients will both deteriorate more and be particularly responsive to treatment, depends on the effectiveness of the drug in both the biomarker-positive and -negative groups, the proportion of biomarker-positive patients in the sample, and the accuracy of the assay [28]. When a small proportion of available patients are biomarker-positive and the drug has little benefit for biomarker-negative patients, then in such cases choosing only biomarker-positive patients would indeed require fewer patients than a standard clinical trial design [28]. In this study, 70% to 74% of aMCI patients were biomarker-positive, potentially limiting the usefulness of CSF $A\beta_{1-42}$ for screening, and there was no meaningful effect on statistical power. For more efficient trials based on preferentially selecting biomarker-positive patients, the treatment in question must be substantially more effective in that group as

compared with the biomarker-negative group. It is important to identify and validate biomarkers for diagnosis and prediction of both disease progression and treatment response when designing targeted clinical trials [4]; however, CSF $A\beta_{1-42}$ biomarkers may be differentially informative at different stages [4,24].

The results also demonstrate differences between modeling and simulations in estimating power for clinical trials. Typically, parameters for power calculations are obtained using summary statistics from reference groups and, assuming a range of effect sizes, corresponding sample sizes are calculated. This approach depends on the critical assumptions that the reference group adequately represents the characteristics of the planned trial sample and the summary statistics capture the heterogeneity among the trial participants. However, heterogeneity in the pattern of outcomes may be unrecognized using summary statistics, particularly, when the model requires scores to change linearly over time, and could explain why we observed no significant increase in statistical power in biomarker-positive patients, whereas others calculated greater power for the same sample sizes

Table 2
Power for ADAS-cog outcomes in 24-month-long trials with 20% and 40% dropouts

N per group	Dropout %	Effect size	Selection method	Treatment group mean	Placebo group mean	Treatment group SD	Placebo group SD	Power mixed model
100	20	0.35	aMCI	0.88	2.86	5.92	5.62	0.56
100	20	0.35	$A\beta$	1.66	3.71	6.18	5.85	0.58
100	20	0.35	t-tau/ $A\beta$	1.58	3.66	6.27	5.92	0.57
100	20	0.45	aMCI	0.33	2.85	6.03	5.61	0.71
100	20	0.45	$A\beta$	1.04	3.73	6.25	5.88	0.76
100	20	0.45	t-tau/ $A\beta$	0.99	3.65	6.41	5.94	0.73
200	20	0.25	aMCI	0.85	2.84	5.93	5.62	0.54
200	20	0.25	$A\beta$	1.66	3.72	6.22	5.88	0.56
200	20	0.25	t-tau/ $A\beta$	1.55	3.67	6.27	5.96	0.61
200	20	0.35	aMCI	0.32	2.85	6.08	5.65	0.78
200	20	0.35	$A\beta$	1.05	3.71	6.28	5.86	0.83
200	20	0.35	t-tau/ $A\beta$	0.96	3.64	6.40	5.95	0.85
200	40	0.35	aMCI	0.89	2.85	5.97	5.65	0.70
200	40	0.35	$A\beta$	1.65	3.68	6.18	5.86	0.71
200	40	0.35	t-tau/ $A\beta$	1.57	3.65	6.30	5.95	0.73
200	40	0.45	aMCI	0.32	2.87	6.10	5.65	0.86
200	40	0.45	$A\beta$	1.06	3.70	6.34	5.87	0.88
200	40	0.45	t-tau/ $A\beta$	0.93	3.68	6.36	5.99	0.90
400	20	0.25	aMCI	1.45	2.86	5.92	5.63	0.81
400	20	0.25	$A\beta$	2.23	3.70	6.15	5.88	0.84
400	20	0.25	t-tau/ $A\beta$	2.17	3.68	6.23	5.98	0.87
400	40	0.25	aMCI	0.86	2.85	6.00	5.66	0.71
400	40	0.25	$A\beta$	1.67	3.70	6.27	5.89	0.77
400	40	0.25	t-tau/ $A\beta$	1.54	3.68	6.32	6.00	0.76
400	40	0.35	aMCI	1.46	2.86	5.92	5.63	0.93
400	40	0.35	$A\beta$	2.25	3.73	6.14	5.88	0.94
400	40	0.35	t-tau/ $A\beta$	2.16	3.67	6.23	6.00	0.95

NOTE. To ensure an approximate power of 80% to 90% for the mixed model analysis, simulations show that for small effects of 0.25, typical to that of cholinesterase inhibitors, somewhat fewer than 400 patients per group are needed with a dropout rate of 20%, and for medium size effects of 0.45, somewhat greater than 100 per group are needed with a dropout rate of 20%. Requiring low $A\beta_{1-42}$ biomarker (“ $A\beta$ ”) or high total tau to $A\beta_{1-42}$ ratio (“t-tau/ $A\beta$ ”) in the selection criteria resulted in very small increases in statistical power; these participants showed greater placebo decline but also increased variability, that is, standard deviation of change. Simulation parameters included $\alpha = 0.05$, effect sizes of 0.15 to 0.75 with χ^2 random errors, and 20% and 40% dropouts with mixed model analysis for participants with missing data.

by using summary data. Therefore, the heterogeneity resulting from the sampling process of the simulations better anticipates the heterogeneity that would be observed in a prospective trial.

One limitation to making inferences from these results is that, although ADNI was meant to inform clinical trials methods, it is not itself a randomized trial. Patients volunteered for a study without planned treatment intervention in which lumbar puncture was optional and ratings were not done under the double-blinded conditions of a randomized, controlled trial. Investigators could obtain knowledge related to *APOE* $\epsilon 4$ genotype, clinical characteristics, test performance, course, severity, and medication use, which could have in turn influenced their diagnosis, clinical ratings, and performing lumbar punctures. Another potential limitation is that the substantial majority who underwent CSF examinations had low $A\beta_{1-42}$ concentrations and high t-tau/ $A\beta_{1-42}$ ratios, and, although consistent with a European MCI sample [24], may not represent samples from broader

communities or nonacademic clinics. Finally, although the use of cholinesterase inhibitors is allowed in all long-term AD clinical trials [29], the nearly half the patients using the drugs were slightly more impaired and declined more as compared with those not using them and this may have affected illness course [20]. The random treatment allocations and thousands of simulations ensured that results were not biased in this respect; however, future simulation studies and clinical trials might consider the potential effects of marketed medications on both internal and external validity.

In summary, selecting aMCI or prodromal AD patients for a clinical trial on the basis of CSF $A\beta_{1-42}$ biomarker criteria will most likely identify relatively more severe patients and not enhance the statistical power of the trials. In the absence of a strong scientific rationale, it may be more practical and clinically relevant to not have $A\beta_{1-42}$ CSF biomarkers as a criterion for trials entry in this setting and to restrict their use as explanatory or stratification variables when there are reasons to do so.

Table 3
Power for CDR-sb outcomes in 24-month-long trials with 20% and 40% dropouts

N per group	Dropout %	Effect size	Selection method	Treatment group mean	Placebo group mean	Treatment group SD	Placebo group SD	Power mixed model
100	20	0.35	aMCI	0.90	1.48	2.21	1.97	0.47
100	20	0.35	$A\beta$	1.22	1.83	2.22	1.94	0.51
100	20	0.35	t-tau/ $A\beta$	1.30	1.91	2.23	1.92	0.51
100	20	0.45	aMCI	0.74	1.48	2.22	1.97	0.63
100	20	0.45	$A\beta$	1.05	1.83	2.25	1.93	0.69
100	20	0.45	t-tau/ $A\beta$	1.13	1.91	2.26	1.93	0.68
200	20	0.35	aMCI	0.91	1.48	2.22	1.97	0.69
200	20	0.35	$A\beta$	1.22	1.83	2.23	1.94	0.76
200	20	0.35	t-tau/ $A\beta$	1.30	1.90	2.23	1.93	0.75
200	20	0.45	aMCI	0.73	1.48	2.22	1.98	0.89
200	20	0.45	$A\beta$	1.04	1.83	2.26	1.94	0.92
200	20	0.45	t-tau/ $A\beta$	1.11	1.91	2.25	1.93	0.90
200	40	0.35	aMCI	0.91	1.48	2.22	1.98	0.62
200	40	0.35	$A\beta$	1.22	1.83	2.23	1.93	0.67
200	40	0.35	t-tau/ $A\beta$	1.30	1.90	2.22	1.93	0.66
200	40	0.45	aMCI	0.73	1.48	2.20	1.96	0.79
200	40	0.45	$A\beta$	1.03	1.83	2.24	1.94	0.84
200	40	0.45	t-tau/ $A\beta$	1.11	1.92	2.25	1.93	0.86
400	20	0.25	aMCI	1.05	1.48	2.16	1.98	0.76
400	20	0.25	$A\beta$	1.39	1.83	2.15	1.94	0.79
400	20	0.25	t-tau/ $A\beta$	1.46	1.90	2.15	1.94	0.77
400	20	0.35	aMCI	0.91	1.48	2.24	1.98	0.93
400	20	0.35	$A\beta$	1.22	1.83	2.23	1.95	0.95
400	20	0.35	t-tau/ $A\beta$	1.30	1.90	2.23	1.93	0.95
400	40	0.25	aMCI	1.05	1.48	2.16	1.98	0.68
400	40	0.25	$A\beta$	1.39	1.83	2.16	1.93	0.67
400	40	0.25	t-tau/ $A\beta$	1.46	1.91	2.15	1.94	0.72
400	40	0.35	aMCI	0.91	1.48	2.23	1.99	0.88
400	40	0.35	$A\beta$	1.22	1.83	2.23	1.94	0.89
400	40	0.35	t-tau/ $A\beta$	1.29	1.91	2.23	1.93	0.91

NOTE. To ensure an approximate power of 80% to 90% for the mixed model analysis, simulations show that for small effects of 0.25, somewhat more than 400 patients per group are needed with a dropout rate of 20%, and for medium size effects of 0.45, somewhat less than 200 per group are needed with a dropout rate of 20%. Requiring low amyloid- β_{1-42} biomarker (“ $A\beta$ ”) or high t-tau/ $A\beta_{1-42}$ (“t-tau/ $A\beta$ ”) as selection criteria resulted in very small increases in statistical power. Gain in power was less prominent as total power increased. Simulation parameters included $\alpha = 0.05$, effect sizes of 0.15 to 0.75 with χ^2 random errors, and 20% to 40% dropouts analyzed with mixed model analysis for participants with missing data.

Acknowledgment

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI, NIA U01 AG024904) database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or the writing of this report. A complete list of investigators in ADNI is listed at: http://www.loni.ucla.edu/ADNI/About/About_InvestigatorsTable.shtml.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (NIH U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514; the Dana Foundation; the USC Alzheimer's Disease Research Center NIH P50 AG05142; and NIH T32 HL072757.

Contributors: All authors participated in the design of the study, statistical analysis, writing and editing of the manuscript, and all reviewed and approved the final draft.

Conflict of interest statement: [During the 36-month window period before submission] LSS reports being an editor on the Cochrane Collaboration Dementia and Cognitive Improvement Group, which oversees systematic reviews of drugs for cognitive impairment and dementia; receiving a grant from the Alzheimer's Association for a registry for dementia and cognitive impairment trials; receiving grant or research support from Baxter, Elan Pharmaceuticals, Johnson & Johnson, Eli Lilly, Myriad, Novartis, and Pfizer; and having served as a consultant for or receiving consulting fees from Abbott Laboratories, AC Immune, Allergan, Alion, Alzheimer Drug Discovery Foundation, AstraZeneca, Bristol-Myers Squibb, Elan, Eli Lilly, Exonhit, Forest, GlaxoSmithKline, Ipsen, Johnson & Johnson, Lundbeck, Myriad, Medavante, Medivation, Merck, Novartis, Pfizer, Roche, Sanofi-Aventis, Schering-Plough, Servier, Toyama, and Transition Therapeutics.

REK declares that he has no conflict of interest. GRC reports having served on data and safety monitoring committees for AntiSense, Sanofi-Aventis, Bayhill, BioMS, Daichi-Sankyo, Eli Lilly, GlaxoSmithKline, Genmab, Medivation, Ono, PTC Therapeutics, Vivus, NHLBI, NINDS, and the NMSS; and having served as a consultant to or receiving consulting fees from Alexion, Accentia, Bayer, Barofold, Biogen-Idec, CibaVision, Enzo, Eisai, Genentech, Millenium, Novartis, Consortium of Multiple Sclerosis Centers, Peptimmune, Klein-Buendel Incorporated, Incyte, Somnus, Teva, Visioneering Technologies.

References

- [1] Shaw LM, Vanderstichele H, Knopik-Czajka M, Clark CM, Aisen PS, Petersen RC, et al. Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol* 2009; 65:403–13.
- [2] Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 2007; 6:734–46.
- [3] Jack CR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, Petersen RC, Trojanowski JQ. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol* 2010;9:119–28.
- [4] Blennow K, Hampel H, Weiner M, Zetterberg H. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol* 2010; 6:131–44.
- [5] Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* 2010;74:201–9.
- [6] A Multicenter, Double Blind, Placebo-Controlled, Safety and Tolerability Study of BMS-708163 in Patients With Prodromal Alzheimer's Disease. Available at: <http://clinicaltrials.gov/ct2/show/NCT00890890?term=Alzheimer&lead=Bristol&rank=2>. Accessed April 15, 2010.
- [7] Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, Trojanowski JQ, Toga AW, Beckett L. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimer Dement* 2005;1:55–66.
- [8] Petersen RC, Thomas RG, Grundman M, Bennett D, Doody R, Ferris S, et al. Vitamin E and donepezil for the treatment of mild cognitive impairment. *N Engl J Med* 2005;352:2379–88.
- [9] Doody RS, Ferris SH, Salloway S, Sun Y, Goldman R, Watkins WE, Xu Y, Murthy AK. Donepezil treatment of patients with MCI: a 48-week randomized, placebo-controlled trial. *Neurology* 2009;72:1555–61.
- [10] Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993;43:2412–4.
- [11] Wechsler DA. Wechsler Memory Scale—Revised. New York, NY: Psychological Corporation; 1987.
- [12] Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, et al. 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.
- [13] 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;12:129–138.
- [14] Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol* 1982;37:323–9.
- [15] Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. *Stat Med* 2006;25:4279–92.

- [16] Cohen J. Statistical power analysis. *Curr Dir Psychol Sci* 1992; 1:98–101.
- [17] Siddiqui O, Hung H, O'Neill R. MMRM vs. LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets. *J Biopharm Stat* 2009;19:227–46.
- [18] R Development Core Team (2009). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: <http://www.Rproject.org>.
- [19] Pinheiro JC, Bates DM. Mixed Effects Models in S and S-Plus. New York, NY: Springer; 2000.
- [20] Schneider LS, Insel P, Weiner M. Cholinesterase inhibitors and memantine use by patients in the Alzheimer Disease Neuroimaging Initiative. *Arch Neurol* (in press).
- [21] Fagan AM, Roe CM, Xiong C, Mintun MA, Morris JC, Holtzman DM. Cerebrospinal fluid tau/beta-amyloid(42) ratio as a prediction of cognitive decline in nondemented older adults. *Arch Neurol* 2007;64:343–9.
- [22] Mattsson N, Zetterberg H, Hansson O, Andreasen N, Parnetti L, Jonsson M, et al. CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. *JAMA* 2009;302:385–93.
- [23] Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *Lancet Neurol* 2006;5:228–34.
- [24] Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund L-O, Freund-Levi Y, et al. Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: a prospective cohort study. *Lancet Neurol* 2009;8:619–27.
- [25] De Meyer G, Shapiro F, Vanderstichele H, Vanmechelen E, Engelborgh B, De Deyn PP, et al. Diagnosis-independent Alzheimer disease biomarker signature in cognitively normal elderly people. *Arch Neurol* 2010;67:949–56.
- [26] Schmand B, Huizenga HM, Van Gool WA. Meta-analysis of CSF and MRI biomarkers for detecting preclinical Alzheimer's disease. *Psychol Med* 2009;29:1–11.
- [27] Fagan AM, Mintun MA, Mach RH, Lee SY, Dence CS, Shah AR, et al. Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Abeta42 in humans. *Ann Neurol* 2006;59:512–9.
- [28] Maitournam A, Simon R. On the efficiency of targeted clinical trials. *Stat Med* 2005;24:329–39.
- [29] Schneider LS, Sano M. Current Alzheimer disease clinical trials: methods and placebo outcomes. *Alzheimer Dement* 2009;5:388–97.

Did you know?

The screenshot displays the homepage of the journal *Alzheimer's & Dementia*. At the top, there is a search bar with a dropdown menu set to 'This Periodical' and a search button. Below the search bar, there are links for 'Advanced Search - MEDLINE - My Recent Searches - My Saved Searches - Search Tips'. The main content area features a 'Current Issue' section for November 2009, Vol. 5, No. 6, with a 'Now Included on MEDLINE' badge. A 'Featured Articles' section lists several research topics, including cognitive performance in African Americans, cerebral blood flow in vascular dementia, and amyloid plaques. The left sidebar contains navigation links such as 'JOURNAL HOME', 'CURRENT ISSUE', 'BROWSE ALL ISSUES', 'ARTICLES IN PRESS', 'SEARCH THIS JOURNAL', 'JOURNAL INFORMATION', 'ADVERTISING INFORMATION', and 'ALZHEIMER'S ASSOCIATION'. At the bottom, there is a 'JOURNAL ACCESS' section with a 'Full-text articles are available from July 2005 to the present' notice and a 'LINKS OF INTEREST' section with a 'Journal of the American Medical Directors Association (JAMDA)' link.

You can search **Alzheimer's & Dementia** and 400 top medical and health sciences journals online, including **MEDLINE**.

Visit www.alzheimersanddementia.org today!