Neurobiology of Aging 48 (2016) 172–181

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Cognitive and functional changes associated with Aβ pathology and the progression to mild cognitive impairment

Philip S. Insel [a,b,c,*], Michael C. Donohue [d], R. Scott Mackin [b,e], Paul S. Aisen [d], Oskar Hansson [a,f], Michael W. Weiner [b,c], Niklas Mattsson [a,f,g], and the Alzheimer's Disease Neuroimaging Initiative

[a] Clinical Memory Research Unit, Faculty of Medicine, Lund University, Lund, Sweden
[b] Center for Imaging of Neurodegenerative Diseases, Department of Veterans Affairs Medical Center, San Francisco, CA, USA
[c] Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA, USA
[d] Department of Neurology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA
[e] Department of Psychiatry, University of California, San Francisco, CA, USA
[f] Memory Clinic, Skåne University Hospital, Lund University, Lund, Sweden
[g] Department of Neurology, Skåne University Hospital, Lund University, Lund, Sweden

ARTICLE INFO

Article history:
Received 5 December 2015
Received in revised form 12 August 2016
Accepted 15 August 2016
Available online 26 August 2016

Keywords:
β-amyloid
Clinical trials
Cognition
Function
Composite
Mild cognitive impairment

ABSTRACT

Cognitively-normal people with evidence of β-amyloid (Aβ) pathology and subtle cognitive dysfunction are believed to be at high risk for progression to mild cognitive impairment due to Alzheimer's disease (AD). Clinical trials in later stages of AD typically include a coprimary endpoint to demonstrate efficacy on both cognitive and functional assessments. Recent trials focus on cognitively-normal people, but functional decline has not been explored for trial designs in this group. The goal of this study was therefore to characterize cognitive and functional decline in (1) cognitively-normal people converting to mild cognitive impairment (MCI) and (2) cognitively-normal β-amyloid-positive (Aβ+) people. Specifically, we sought to identify and compare the cognitive and functional assessments and their weighted combinations that maximize the longitudinal decline specific to these 2 groups. We studied 68 people who converted from normal cognition to MCI and 70 nonconverters, as well as 137 Aβ+ and 210 β-amyloid-negative cognitively-normal people. We used bootstrap aggregation and cross-validated mixed-models to estimate the distribution of weights applied to cognitive and functional outcomes to form composites. We also evaluated best subset optimization. Using optimized composites, we estimated statistical power for a variety of clinical trial scenarios. Overall, 55.4% of cognitively-normal to MCI converters were Aβ+. Large gains in power estimates were obtained when requiring participants to have both subtle cognitive dysfunction and Aβ pathology compared with requiring Aβ pathology alone. Additional power resulted when including functional as well as cognitive outcomes as part of the composite. Composites formed by applying equal weights to all measures provided the highest estimates of cross-validated power, although similar to both continuous weight optimization and best subset optimization. Using a composite to detect a 30% slowing of decline, 80% power was obtained for predicted Aβ+ converters with 375 completers/arm for a 30-month trial using a combination of cognitive/functional measures. In the Aβ+ group, power to approach levels suitable for a phase III clinical trial would require considerably larger sample sizes. Composites incorporating both cognitive and functional measures may substantially increase the power of a trial in a preclinical (Aβ+) AD population with subtle evidence of cognitive dysfunction.

1. Introduction

Accumulating evidence from Alzheimer's disease (AD) biomarker studies suggests β-amyloid (Aβ) deposition may occur decades before the diagnosis of clinical dementia (Morris, 2005). Anti-Aβ treatments are thought to have a higher likelihood of slowing progression if administered at the earliest signs of the

* Corresponding author at: Center for Imaging of Neurodegenerative Diseases, Department of Veterans Affairs Medical Center, San Francisco, CA 94121, USA. Tel.: +1 858 652 8480; fax: 415-221-4810.
E-mail address: philipinsel@gmail.com (P.S. Insel).

0197-4580/$ – see front matter © 2016 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neurobiolaging.2016.08.017

pathological cascade, before substantial neurodegeneration and other downstream effects of Aβ deposition (Sperling et al., 2011b). Classification of Alzheimer's disease into progressive stages has helped to organize the current thinking about the emergence of subtle clinical symptoms and the development of cognitive and functional impairment during the continuum of disease progression (Sperling et al., 2011a). The initial stages of preclinical AD are defined by amyloidosis and neurodegeneration. The final preclinical stage also includes some evidence of subtle cognitive dysfunction, although below levels of cognitive and functional impairment required to meet criteria for mild cognitive impairment (MCI) due to AD. As the disease progresses into MCI and dementia, cognitive and functional deficits may be observed. Identifying the biomarkers and clinical assessments that can predict and monitor the progression from the early stages of AD to more advanced disease will help to elucidate the disease process and inform clinical trial design (Insel et al., 2015). Here we sought to determine the optimal combination of cognitive and functional measures to track disease progression in cognitively-normal people progressing to MCI, and of Aβ-positive (Aβ+) cognitively-normal people. Composite endpoints comprising both cognitive and functional measures are currently being used in clinical trials of MCI populations (Ard et al., 2015; Raghavan et al., 2013; Wang et al., 2016). Here we consider the inclusion of functional measures in the endpoint for clinical trials in preclinical AD.

The Food and Drug Administration (FDA) recently offered draft guidance to update their recommendations on primary endpoint selection in clinical trials for early-stage AD (US Dept of Health and Human Services, 2015). With the focus of recent clinical trials on treatment in these earlier stages of AD, including prodromal AD and preclinical AD, the FDA recognized the difficulty in demonstrating drug efficacy using prior guidelines developed for trials with subjects in the dementia stage of AD (Kozauer et al., 2013; McKhann et al., 2011). Trial design in later stages of AD has typically included a coprimary endpoint to demonstrate efficacy on both a cognitive and a functional assessment. However, the assessment tools used in these trials have not been validated in earlier stage subjects (Snyder et al., 2014), leading the FDA to consider the use of a single primary composite endpoint that captures both cognitive and functional decline, in trials of prodromal AD subjects. Preclinical AD subjects are, by definition, cognitively-normal and should not have any functional impairment due to cognitive dysfunction. We hypothesize that as the target population progresses on the continuum of decline, assessing functional changes may take a more central role in demonstrating a drug effect to be clinically meaningful. However, the feasibility and value of assessing functional decline as part of a trial endpoint in a preclinical population are unknown.

Since the FDA guidance, several cognitive composites have been developed to capture the decline specific to subjects with preclinical AD, but no attempts have been made to develop combined cognitive and functional composites. The Alzheimer's Prevention Initiative (API) has developed cognitive composites using Presenilin 1 E280A mutation carriers (Ayutyanont et al., 2014) and also cognitively-normal elders who converted to MCI or AD (Langbaum et al., 2014). A third cognitive composite, to be used as the primary endpoint in the A4 trial (Sperling et al., 2014), was developed to capture decline in Aβ+ cognitively-normal elders (Donohue et al., 2014), and selected individual components based on a literature review. Functional assessments were not evaluated in the API or the A4 composites.

The aim of this study was to identify and compare the cognitive and/or functional assessments and their weighted combinations that maximize the longitudinal decline specific to (1) cognitively-normal to MCI converters (cCN); and (2) cognitively-normal Aβ+ subjects. Conversion status is not known at the beginning of the study, and thus, power estimates based on subjects' true conversion status would not be useful to inform a clinical trial. Therefore, to reflect a realistic modern trial scenario, subjects who were both predicted to convert using information available at baseline and were also Aβ+ (pcCN), were used to estimate clinical trial power. Using the battery of assessments from the Alzheimer's Disease Neuroimaging Initiative (ADNI), we sought to characterize the importance of each cognitive and functional assessment in our 3 groups (cCN, pcCN, and Aβ+) as well as provide cross-validated estimates of power when using the composites in clinical trial scenarios.

2. Material and methods

2.1. Participants

Data were obtained from the ADNI database (adni.loni.usc.edu). ADNI is the result of efforts of many coinvestigators, and participants have been recruited from over 50 sites across the United States and Canada (see www.adni-info.org). The population in this study included ADNI-1 and ADNI-2 participants enrolled into the cognitively-normal or subjective memory complaint cohorts, were tested for cerebrospinal fluid (CSF) biomarkers or ^{18}F-florbetapir positron emission tomography (PET), and were followed longitudinally for neuropsychological testing.

2.2. Cerebrospinal fluid biomarker concentrations

Each CSF sample was collected by lumbar puncture and shipped on dry ice to the ADNI Biomarker Core laboratory at the University of Pennsylvania Medical Center for long-term storage at −80 °C. CSF $A\beta_{42}$ was measured using the multiplex xMAP Luminex platform (Luminex Corp, Austin, TX, USA) with the research use only INNOBIA AlzBio3 kit (Fujirebio/Innogenetics, Ghent, Belgium) (Olsson et al., 2005; Shaw et al., 2009).

2.3. Florbetapir PET

Methods to acquire and process ADNI florbetapir PET image data were described previously (Landau et al., 2012). Full details of acquisition and analysis can be found at http://adni.loni.usc.edu/methods/.

2.4. Cognitive and functional outcomes

Cognitive measures assessed were the Mini–Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale–cognitive subscale, 13-item version (ADAS13), immediate and delayed memory recall from the Wechsler Memory Scale (iMemory, dMemory), immediate and delayed Rey Auditory Verbal Learning Test (iAVLT, dAVLT), Trail Making Test parts A and B (Trails A & B), Boston Naming Test, and Category Fluency. The Clinical Dementia Rating Sum of Boxes (CDR-SB) was also assessed, which includes both cognitive and functional items, and finally the Functional Assessment Questionnaire (FAQ), which is purely a functional assessment (Kaplan et al., 1982; Morris, 1993; Pfeiffer et al., 1982; Reitan, 1958; Rey, 1964; Rosen et al., 1984; Wechsler, 1987).

2.5. Statistical analysis

This study included 3 main sets of analyses. The first was a comparison of normal participants who converted to a diagnosis of MCI (cCN) during 7 years of follow-up versus stable cognitively-normal (sCN) participants, during the same period. Follow-up on

converters continued beyond diagnosis of MCI. The second analysis was a comparison of participants predicted to convert to MCI (pcCN, only including Aβ+ subjects) versus the β-amyloid-negative (Aβ−) participants from the sCN group. The third analysis was a comparison of Aβ+ versus Aβ− participants, irrespective of conversion information. There was considerable overlap among these groups. The cCN or sCN participants that also had Aβ information (n = 56 from the cCN group, and n = 57 from the sCN group) were also included in either the Aβ− or Aβ+ groups. This is described further in the results section. All pcCN participants were included in the Aβ+ group.

Aβ+ was defined as florbetapir PET SUVR >1.10 at any point during follow-up (Landau et al., 2012). Subjects without florbetapir PET were considered Aβ+ if CSF $A\beta_{42}$ < 192 ng/L (Shaw et al., 2009).

In each of the 3 groups, we compared 2 types of optimization: the first allowed continuous weights for each component while the second was more constrained, allowing only combinations of components with 0 or 1 weights (0 = exclusion, 1 = inclusion), providing the best subset of components. For the continuous weight optimization, in each group, composite weights were estimated via bootstrap resampling and cross-validation to find the set of weights that maximized the separation of the groups over the first 48 months of follow-up. Spline knots for models limited to 48-month follow-up were placed at 12, 24, and 36 months, post baseline. The median weight from this distribution for each outcome was used to form the composite to be evaluated for trial power. Details of each step are described in the following.

Longitudinal cognitive and functional measures were modeled using linear mixed-effects regression with a random intercept and slope and an unstructured covariance matrix for the random effects. Variance components were estimated conditional on converter (or amyloid) status. All models included age (years), education (years), gender, group, time since initial visit, and the interaction between group and time, as predictors. To capture departures from linearity in the trajectory of cognition and function, continuous time from the baseline test was parameterized using a 3-knot restricted cubic spline, with knots placed at 1, 3, and 5 years, post baseline. Differences in group trajectories were tested using interactions between the 2 parameters for time resulting from the spline and the group factor, group × (β_{time1} + β_{time2}). Likelihood ratio tests were used to test the significance of the interaction for longitudinal differences and Wald tests on the main effect for group were used to test for baseline differences (Chambers and Hastie, 1992). The *p*-values were 2-sided and adjusted for multiplicity using a Hochberg correction (Hochberg, 1988). The *p*-values < 0.05 were considered significant and the *p*-values < 0.10 were considered marginally significant.

Within each analysis, we aimed to identify the composite weights that maximized the separation of the trajectories of the groups over time. We evaluated 2 types of composites: one that considered only the 10 cognitive components and another that also included the 2 measures with functional assessments. To form the composites, z-scores (mean centered and scaled to the standard deviation of all baseline and longitudinal scores) of each component were weighted and summed. For the continuous weight optimization, weights for each component could be anywhere on the interval [0, 1]. Numerical optimization was used to search the space of candidate weights to maximize the separation of the groups. Bound constrained optimization (Byrd et al., 1995) was used to maximize the likelihood ratio test for group trajectory differences.

The large number of cognitive and functional components considered and the space of possible weight combinations increases the risk of overfitting. To minimize overfitting, weights were estimated and evaluated using bootstrap aggregation and 10-fold cross-validation (Breiman, 1996). Folds were balanced on group status and cognitive and functional measures. In each training set, 100 bootstrap samples were used to estimate a distribution of optimal composite weights. The median weight for each component from this distribution was then used to form the composite to be evaluated in the test set. The steps of the analysis are shown in a flowchart in Supplementary Fig. 1. The resulting estimates of longitudinal change and variance in the test sets were averaged and used to estimate power for hypothetical clinical trials, as described in the following. For the best subset optimization, the number of components that maximized the cross-validated likelihood ratio test was used in the final composite.

To determine the value of the composites derived by this analysis, we used the cross-validated estimates of change and variance to simulate hypothetical clinical trial scenarios with a proportional treatment effect over time in the active group. Averaging over test set estimates for change from baseline to 18, 24, 30, and 36 months and the estimates of the residual error, subject-specific intercepts and slopes, and the correlation between the intercepts and slopes, we estimated the power to detect a 30% decrease in the difference between the change in the groups. Sampling from the aforementioned estimates and assuming a range of sample sizes, we simulated 1000 longitudinal clinical trials for each sample size, composite type (cognition and function, cognition only, best subset, and for comparison, flat weights across all 12 components), and group. Power was estimated as the proportion of significant *p*-values for the difference in change from baseline at the final visit between the drug and placebo groups, using a mixed-model repeated-measures design (Siddiqui et al., 2009).

The pcCN subjects (restricted to Aβ+ subjects) were identified using baseline cognitive/functional assessments, and demographic and *APOE* information, with a random forest model (Breiman, 2001). Using cross-validated estimates of the probability of conversion, we repeated all steps of the analysis described previously to estimate power for a clinical trial based on participants who were both Aβ+ and predicted to convert (pcCN), to make our results applicable to trials requiring Aβ+ for inclusion. Three-fold cross-validation was used for the pcCN analysis because of the reduced sample size.

The association between groups within each cohort and missing data was modeled using generalized mixed-effects regression with a binomial indicator for a missing visit. All analyses were done in R version 3.1.1 (The R Foundation for Statistical Computing, Vienna, Austria; www.r-project.org).

3. Results

3.1. Cohort characteristics

In the ADNI data set, 68 subjects converted to MCI during 7 years of follow-up while 70 subjects remained cognitively-normal throughout the same period. cCN subjects were older and had more *APOE* ε4 carriers compared to sCN (Table 1). There were no significant differences in gender or education. As described in the Section 2.5, we also identified a group of cognitively-normal subjects who were predicted to convert to MCI (pcCN, only including Aβ+ subjects). Characteristics of the pcCN group are shown in Supplementary Table 1.

One hundred thirty-seven Aβ+ subjects and 210 Aβ− subjects were included in the analysis. Aβ+ subjects were older, less educated, and had more *APOE* ε4 carriers (Table 1). There was no difference in gender. A Kaplan-Meier plot showing the distribution of conversion times for the cCN, Aβ−, and Aβ+ groups is shown in Supplementary Fig. 2.

Table 1
Baseline characteristics

Variable	Converters (N = 68), mean (SD)	Nonconverters (N = 70), mean (SD)	*p*
Age	76.5 (5.55)	74.9 (4.11)	0.014
Gender, female N (%)	28 (41.2)	37 (52.9)	0.178
Education	16.0 (2.67)	16.4 (2.75)	0.255
APOE ε4 N (%)	26 (38.2)	16 (22.9)	0.064
Aβ positivity N (%) (available for 56 cCN and 57 sCN)	31 (55.4)	18 (31.6)	0.014
	Aβ+ (N = 137), mean (SD)	Aβ− (N = 210), mean (SD)	*p*
Age	75.6 (5.09)	73.5 (5.91)	<0.001
Gender, female N (%)	74 (54.0)	98 (46.7)	0.189
Education	16.0 (2.71)	16.6 (2.65)	0.044
APOE ε4 N (%)	51 (37.2)	40 (19.0)	<0.001

Key: Aβ+, β-amyloid-positive; Aβ−, β-amyloid-negative; SD, standard deviation.

Of the 68 cCN participants, 56 had Aβ information: 31 (55.4%) were Aβ+ and 25 (44.6%) were Aβ−. Of the 70 sCN participants, 57 had Aβ information: 18 (31.6%) were Aβ+ and 39 (68.4%) were Aβ−.

3.2. Baseline cognitive/functional differences

When baseline cognitive/functional measures were compared in cCN versus sCN, cCN subjects performed worse on all 12 outcomes. Results with multiple comparison corrections are shown on the top left of Table 2. There were fewer differences on baseline cognitive/functional measures in Aβ+ versus Aβ− participants (top right of Table 2).

3.3. Longitudinal change

cCN subjects worsened significantly faster on 10 of the 12 cognitive and functional outcomes compared to sCN subjects, with the exception of the Boston Naming Test and Trails A over 7 years of follow-up (Fig. 1, Table 2, Supplementary Fig. 3). The largest effect size was in the CDR-SB, and the largest effect sizes among measures without functional items were in the iAVLT and the ADAS13. Longitudinal trajectories of the pcCN group are shown in Supplementary Fig. 4.

Aβ+ subjects worsened significantly faster on 6 of the 12 outcomes compared to Aβ− subjects (Fig. 1, Table 2). The largest effect size was in the ADAS13.

cCN subjects were more likely than sCN subjects to be missing data during the course of the 7 years of follow-up (log OR = 0.82, standard error = 0.15, $p < 0.001$). However, sCN subjects were selected to have a minimum follow-up time of 7 years. Aβ-positivity was not associated with increased missingness (log OR = −0.04, standard error = 0.27, $p = 0.87$).

3.4. Composite weight distributions

The distributions from 1000 bootstrap samples of the composite weights that maximized the separation of the groups are shown in Supplementary Fig. 5. Composite weights were estimated separately for the 3 groups (cCN, pcCN, and Aβ+).

The largest contributing outcomes in the composite for cCN versus sCN were the 2 delayed memory recall measures (dMemory, dAVLT), CDR-SB, and the MMSE (top left of Supplementary Fig. 5). Outcomes with smaller, although nonzero, positive median weights, included Category Fluency, iMemory, Trails A, and the Boston Naming Test. When the functional measures were excluded, the delayed memory recall measures and MMSE remained the largest weighted outcomes and ADAS13 became more heavily weighted.

Table 2
Baseline and longitudinal differences

Baseline	Baseline difference, converters versus nonconverters (N = 138)		Baseline difference, Aβ+ versus Aβ− (N = 347)	
	$Z_{converter} - Z_{nonconverter}$ (SE)	*p*	$Z_{Aβ+} - Z_{Aβ−}$ (SE)	*p*
MMSE	−0.06 (0.13)	0.641	0.06 (0.09)	0.997
ADAS13	0.58 (0.11)	<0.001	0.02 (0.08)	0.997
dMemory	−0.44 (0.13)	0.01	−0.09 (0.09)	0.997
iMemory	−0.42 (0.13)	0.012	−0.05 (0.09)	0.997
dAVLT	−0.46 (0.14)	0.01	−0.03 (0.10)	0.997
iAVLT	−0.48 (0.12)	0.001	0.03 (0.09)	0.997
Trails A	0.37 (0.14)	0.063	0.23 (0.10)	0.298
Trails B	0.32 (0.12)	0.063	0.28 (0.10)	0.062
Boston Naming Test	−0.32 (0.13)	0.084	−0.06 (0.09)	0.997
Category Fluency	−0.30 (0.16)	0.191	0.06 (0.10)	0.997
CDR-SB	0.09 (0.06)	0.299	0.0002 (0.04)	0.997
FAQ	0.19 (0.09)	0.156	−0.06 (0.06)	0.997
Longitudinal change	**Converters versus nonconverters (N = 138, N Obs = 871)**		**Aβ+ versus Aβ− (N = 347, N Obs = 1441)**	
	χ^2	*p*	χ^2	*p*
MMSE	15.03	0.002	25.96	<0.001
ADAS13	29.25	<0.001	38.95	<0.001
dMemory	24.73	<0.001	11.92	0.018
iMemory	18.47	<0.001	11.99	0.018
dAVLT	26.75	<0.001	5.48	0.194
iAVLT	32.26	<0.001	9.44	0.052
Trails A	6.83	0.066	1.85	0.396
Trails B	21.55	<0.001	3.38	0.369
Boston Naming Test	3.82	0.148	9.13	0.052
Category Fluency	15.16	0.002	5.79	0.194
CDR-SB	51.61	<0.001	17.64	0.001
FAQ	51.22	<0.001	34.18	<0.001

Key: Aβ+, β-amyloid-positive; Aβ−, β-amyloid-negative; ADAS13, Alzheimer's Disease Assessment Scale—cognitive subscale, 13-item version; iAVLT, immediate Rey Auditory Verbal Learning Test; CDR-SB, Clinical Dementia Rating Sum of Boxes; dAVLT, delayed Rey Auditory Verbal Learning Test; dMemory, delayed memory recall from the Wechsler Memory Scale; FAQ, Functional Assessment Questionnaire; iMemory, immediate memory recall from the Wechsler Memory Scale; MMSE, Mini–Mental State Examination; SE, standard error.

Composite weights that maximized the separation of pcCN and sCN subjects were also estimated. Using baseline information including demographics, *APOE* ε4 status, and cognitive/functional variables that were not heavily weighted in the true converter composite (ADAS13, Trails A & B, FAQ, Boston Naming Test, iAVLT, and iMemory), composite weights were estimated based on 32 Aβ+ pcCN and 31 Aβ− sCN participants. In reality, these 32 pcCN participants consisted of 25 converters and 7 nonconverters, resulting in a 78% positive predicted value from the model estimates. pcCN subjects were older, less educated, had more *APOE* ε4 allele carriers, and had lower cognitive scores at baseline compared with sCN subjects, similar to cCN subjects (Supplementary Table 1). We then estimated composite weights for this cohort. These weights are shown in the middle row of Supplementary Fig. 5. Similar to the cCN composite, the main outcomes for the pcCN composite were dMemory, CDR-SB, and MMSE, but in contrast, included the Boston Naming Test. When functional measures were excluded, the ADAS13 carried more weight. Note that the pcCN were Aβ+ by design because we aimed to make our results applicable to a trial requiring Aβ+ for inclusion.

The composites for Aβ+ versus Aβ− were heavily weighted by ADAS13, FAQ, and MMSE (bottom left Supplementary Fig. 5). When functional measures were excluded, ADAS13 and MMSE dominated the composites.

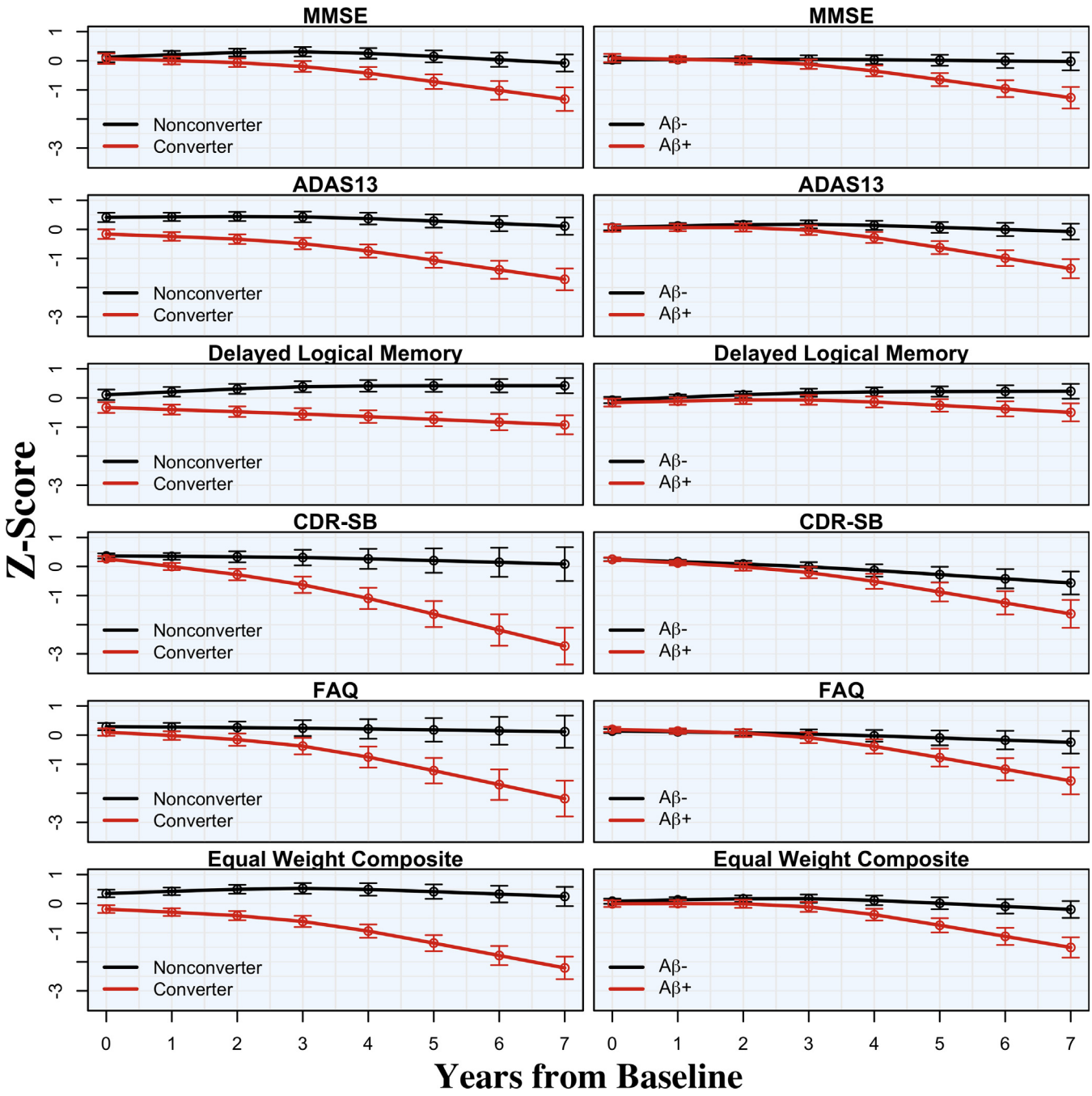

Fig. 1. Longitudinal plots of cognitive and functional assessments of converters versus nonconverters on the left and Aβ+ versus Aβ− on the right. Z-scores of each assessment are plotted from baseline through 7 years of follow-up. Abbreviations: Aβ+, β-amyloid-positive; Aβ−, β-amyloid-negative; ADAS13, Alzheimer's Disease Assessment Scale—cognitive subscale, 13-item version; CDR-SB, Clinical Dementia Rating Sum of Boxes; FAQ, Functional Assessment Questionnaire; MMSE, Mini−Mental State Examination.

3.5. Best subset components

The best subset results were similar to the continuous optimization results. For the cCN versus sCN comparison, 5 components provided the optimal cross-validated composite, with the MMSE, dMemory, dAVLT, CDR-SB, and Category Fluency selected in nearly all cross-validation folds. For the pcCN versus sCN comparison, 7 components were selected, including the MMSE, dMemory, dAVLT, CDR-SB, Category Fluency, and iMemory in nearly all folds and occasionally either ADAS13 or Trails A. For the Aβ+ versus Aβ− comparison, 3 components were selected—the MMSE, ADAS13, and FAQ. The power for these composites is described in the following.

3.6. Power

We estimated the power to detect a 30% slowing of decline using the average out-of-sample estimates of change and variance for each composite and group, over a range of sample sizes. The composite with flat weights across all measures was the best performing composite, attaining 80% power with 375 completers/arm in a hypothetical 30-month trial. Eighty percent power was attained with 450 completers per arm using the optimized cognitive/functional composite in a hypothetical 30-month clinical trial. Sixty-five percent of power was obtained with 500 completers per arm over a 30-month trial, using a

composite with cognition only. We also compared flat weight and optimized cognitive/functional composites in 48-month trials for Aβ+ pcCN subjects. They performed similarly (Supplementary Fig. 6).

The power estimates for the Aβ+ subjects are shown in the lower portion of Fig. 2. With similar sample sizes as the comparisons mentioned previously, power estimates never exceeded 40% with any type of composite.

3.7. *Effect sizes and variance components*

In Fig. 3, the magnitude of change, the within- and between-subject standard deviations, and effect sizes are plotted against the number of components included in the best subset composites, for cCN versus sCN and Aβ+ versus Aβ−. For both groups, the magnitude of change and both types of SD decrease with an increasing number of components included in the composites.

4. Discussion

The main findings of this study were as follows: (1) including participants with Aβ pathology as well as subtle cognitive dysfunction, predictive of progression to MCI, resulted in large gains in power estimates compared to participants with Aβ pathology alone; (2) further gains in power were obtained by including measures with functional items in the composite; (3) composites formed by applying equal weights to all 12 measures provided the highest estimates of cross-validated power, although similar to continuous weight optimization and best subset optimization; (4) as the number of components in the composite increased, the magnitude of change decreased, but both the within-subject and between-subject variance decreased, leading to an increase in effect size; (5) in cCN and pcCN participants, the composite measures selected via optimization were both delayed memory recall assessments, CDR-SB, MMSE, Category Fluency, and immediate memory recall; in Aβ+ participants, ADAS13, MMSE, and FAQ were

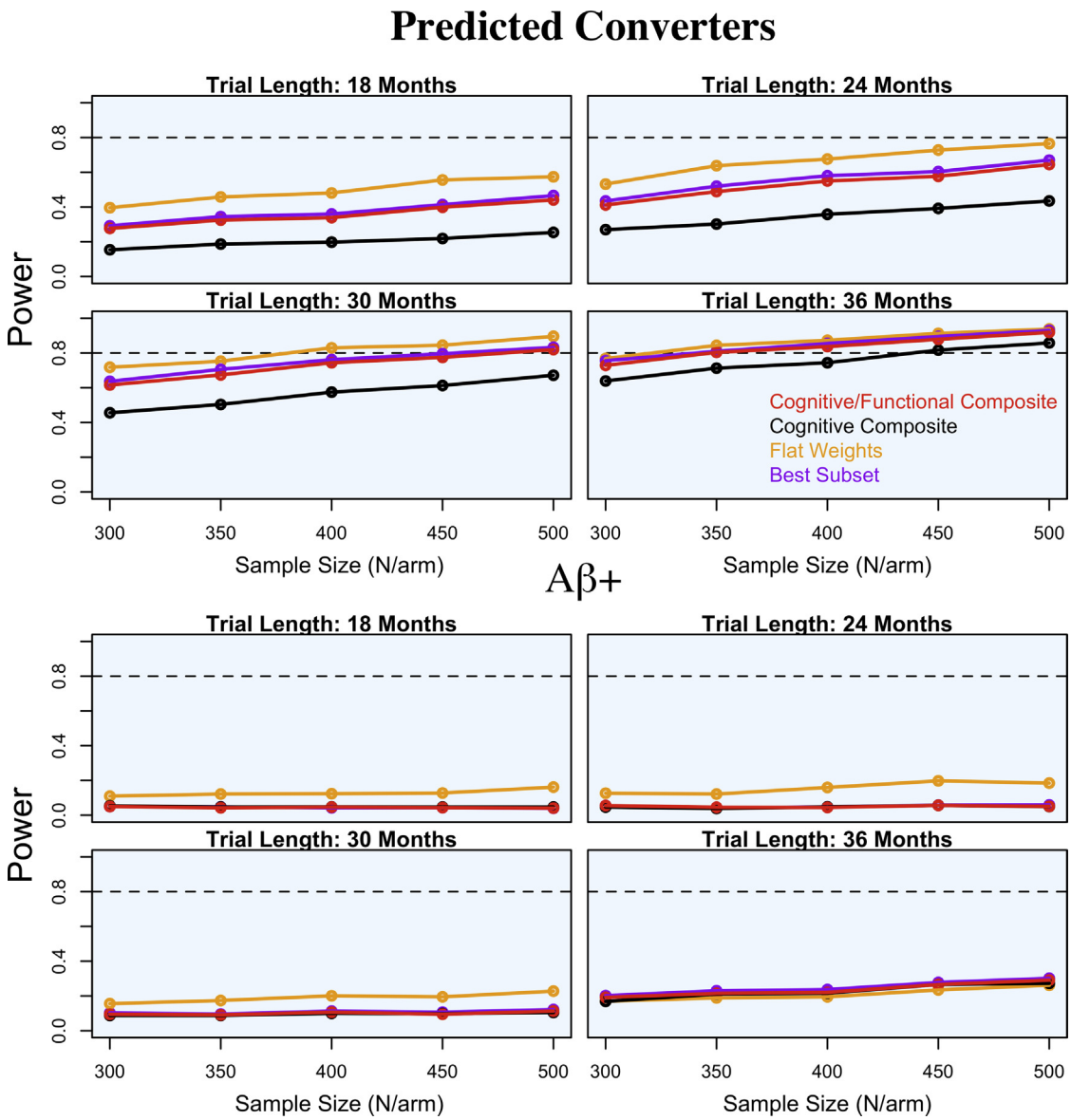

Fig. 2. Plots of power estimates at different sample sizes of completers per arm. The top 2 rows show power estimates for the predicted converters for trials ranging from 18 to 36 months with 300–500 subjects per arm for the 4 types of endpoints. The bottom 2 rows show power estimates for Aβ+ subjects over the same length trials and sample sizes for the 4 types of composites. Abbreviation: Aβ+, β-amyloid-positive.

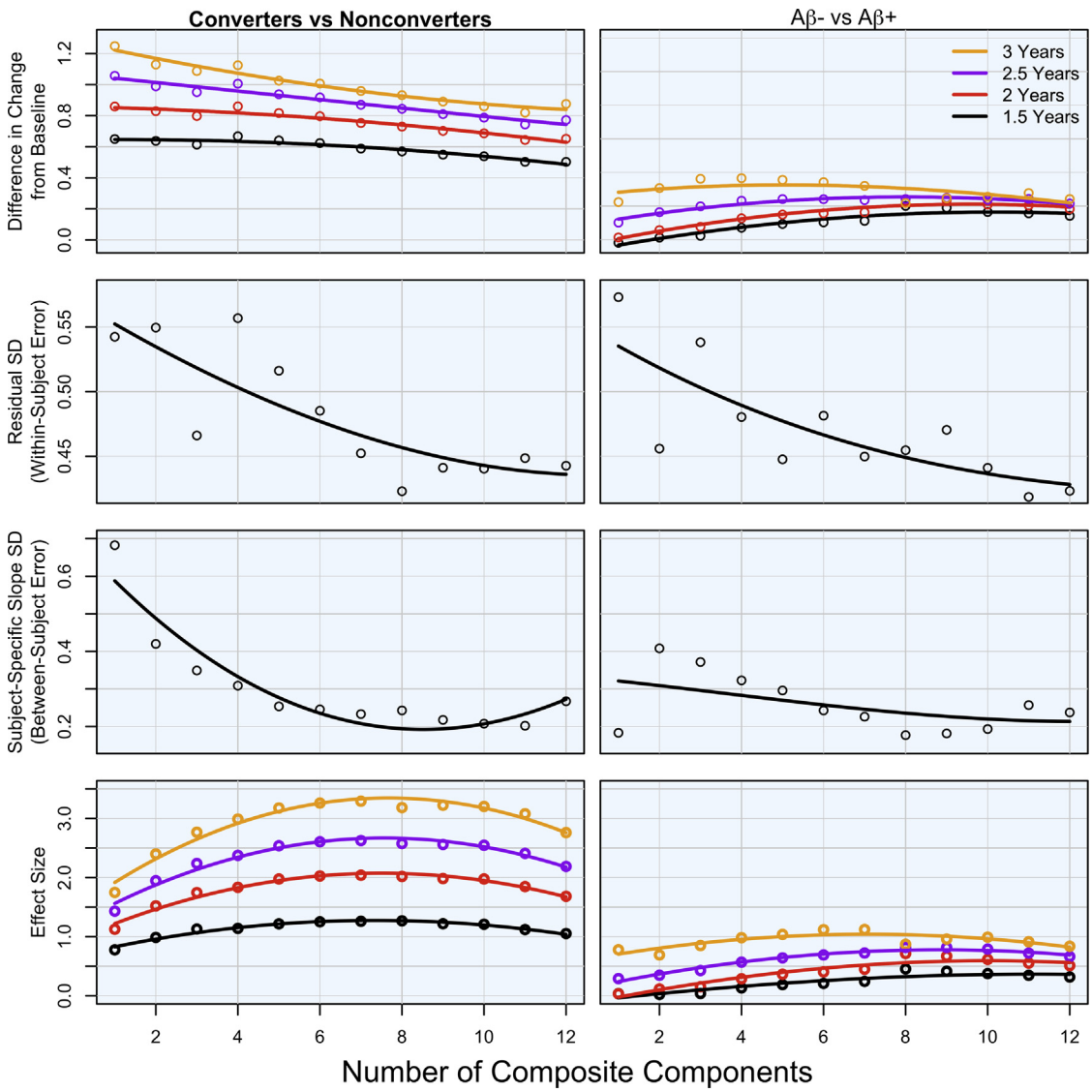

Fig. 3. Differences in the magnitude of change between groups are plotted against the number of components included in the composite, in the top row. The best single component is furthest left on the x-axis, followed by best 2-component combination, all with equal weights. The second and third rows show how the within- and between-subject variance changes as the number of components increases for each analysis group. The bottom rows show how the effect size changes with increasing number of components. Abbreviations: Aβ+, β-amyloid-positive; Aβ−, β-amyloid-negative; SD, standard deviation.

selected, however, these composites did not outperform the equal weight composites when cross-validated in either group; and (6) only 55.4% of cCN subjects were Aβ+, which explains part of the difference between our analysis of cCN and Aβ+ subjects, and points to the importance of non-Aβ–mediated processes to explain development of cognitive and functional decline.

4.1. Power increase with predicted converters

Substantial increases in power estimates result when including pcCN subjects compared to Aβ+ subjects, in all clinical trial simulations (Fig. 2). This might be expected when considering Fig. 1 and Table 2, where decline is limited both in magnitude and number of outcomes in the Aβ+ subjects compared to cCN subjects, especially over the first 36 months. In contrast, the cCN subjects have already diverged from sCN subjects on several measures at baseline and continue to separate on delayed memory recall, global cognitive, and functional outcomes. Lower cognitive scores and continued decline in both the pcCN and cCN groups indicate that these participants are likely already in a later stage of disease at baseline compared with Aβ+ participants. The lower power estimates using ADNI Aβ+ subjects are consistent with estimates reported in the analysis of 2 Aβ+ cohorts for the A4 composite (Donohue et al., 2014). Substantially shallower decline was observed in the cognitively-normal ADNI Aβ+ subjects compared to the cognitively-normal Aβ+ subjects from the Australian Imaging, Biomarkers & Lifestyle Flagship Study of Ageing (AIBL, Ellis et al., 2009; Donohue et al., 2014). The sharper decline seen in the Aβ+ subjects in the AIBL cohort may be due to subtle cognitive dysfunction at baseline including a 0.4 point lower average MMSE score, 0.5 point lower delayed memory delayed recall score, as well as a 20% increase in *APOE* ε4 allele carriers, compared with the ADNI Aβ+ subjects. These subtle differences in baseline cognition and the increased proportion of *APOE* ε4 carriers may account for the differences in power estimates, which are closer to the estimates of the pcCN cohort observed in this analysis. Taken together, these results point to the importance of assessing other baseline characteristics besides Aβ status when selecting preclinical populations

for trial enrichment. This should come as no surprise given the vast literature on the variability of the clinical impact of Aβ pathology in elderly people, where a similar degree of Aβ pathology may be seen in people who are cognitively normal, slightly impaired or fully demented. This variability likely stems from individual differences in cognitive and brain reserve mechanisms, differences in the presence and spread of important copathologies (including tau pathology), and differences in the time that the individual has been exposed to Aβ pathology before testing. Additional sources of variation regarding the effect of Aβ pathology on cognition in cognitively-normal cohorts include biomarker modality (PET vs. CSF) and also choice of threshold for Aβ-positivity (Insel et al., 2014; Mattsson et al., 2014, 2015). The impact of this on the power of clinical trials, as found in our results, is in agreement with a previously proposed staging of preclinical AD (Sperling et al., 2011b), where subjects with a combination of positive AD biomarkers (including Aβ biomarkers) and subtle cognitive dysfunction are thought to be at much higher risk for future clinical deterioration compared to subjects with positive Aβ biomarkers alone (Vos et al., 2013).

4.2. Additional power increase with functional components

Including the CDR-SB and FAQ in either the optimized composite or the equal weight composite resulted in an additional increase in power over the cognitive composite in the pcCN subjects, reaching 80% power with 375 completers per arm for a 30-month trial (Fig. 2). Including measures with functional items provided moderate improvements in power for the composites in Aβ+ subjects for trials less than 36 months, although power remained low. To convert from normal cognition to MCI, a subject must demonstrate a clinically meaningful level of functional decline. Steep decline is observed on both the CDR-SB (Fig. 1) and on the FAQ immediately after baseline in cCN subjects. Thus, it follows that including measures with functional assessments in a composite results in a more sensitive instrument, in a population of converters. However, because conversion status in not known at baseline, the inclusion of functional assessments in a prospective study will only improve sensitivity if information available at baseline can successfully identify subjects who are on the verge of functional decline. When functional measures are excluded, the weights for both the ADAS13 and the MMSE increase. This may reflect that poor scores on a global cognitive test are likely more correlated with functional decline compared to single domain measures. Aβ+ subjects do not show substantial decline on either CDR-SB or FAQ before month 48.

4.3. Functional and cognitive outcome selection

Delayed memory recall, the MMSE, and the CDR-SB were selected via optimization for both the cCN and the pcCN composites. However, even the top-weighted measures had relatively low median weights, with 10 of the 12 measures having positive weights for cCN subjects, and 6 of 12 having positive weights for pcCN subjects (Supplementary Fig. 5). The spread of the weights suggests that many domains are declining early in the conversion process. Thus, it follows that the equal weight composites performed well. The failure of the optimization to beat the equal weight composites suggests that using either continuous weights or best subset component selection results in overfitting the training sets and a subsequent reduction of test set power. Including a large number of components in a composite may smooth over aberrations in scores in a particular assessment from visit to visit within a subject, thus lowering the within-subject variance and improving signal to noise. Similarly, the equal weight composite provided the most power in Aβ+ participants, although power did not approach levels suitable for a phase III trial (Fig. 2).

4.4. Equal weight composite: effect size, magnitude of change, and variance

Reasons for slight increases in power with the equal weight composite become clear from inspection of Fig. 3. As the number of components included in the composite increases, the magnitude of change decreases. This would result in a decrease in effect size (if the variance is held constant) and subsequently, a decrease in power. If we start with the best single component and continue by adding additional components, the magnitude of change may become diluted as less-sensitive components are included in the composite. We might expect the effect size to drop with each additional component; however, both the within-subject and between-subject errors are decreasing at a rate that overcomes the decrease in the magnitude of change, resulting in an increasing effect size, as seen at the bottom of Fig. 3. The increase in effect size plateaus in the 6–10 component range for both the converter and Aβ+ groups. The decrease in within-subject variance is clear in both groups; however, the drop in between-subject variance is steeper for converters, likely due to more consistent decline across all components. Or alternatively, the converters' scores are more variable, with more room for a reduction in within-subject variance when the number of composite components increases.

4.5. Outcome selection in other studies

The outcomes selected via optimization are consistent with the measures found to capture decline in other cohorts. The API composite in Presenilin 1 E280A mutation carriers includes the Word List Recall (CERAD), MMSE (orientation to time), and also Constructional Praxis and Raven's Progressive Matrices (Ayutyanont et al., 2014). The API composite developed from normal to MCI or AD converters includes Category Fluency, Logical Memory II (dMemory), MMSE (orientation to time), and also Ravens Progressive Matrices Subset, and Symbol Digit Modalities (Langbaum et al., 2014). The A4 composite for Aβ-positivity includes the Total Recall score from the Free and Cued Selective Reminding Test, Logical Memory II (dMemory), MMSE, and the Digit Symbol Substitution Test (Donohue et al., 2014). Delayed memory recall, orientation, and processing speed are consistently selected domains.

A variety of approaches can be used to develop composites that are sensitive to change over time (Ard et al., 2015). The development of composite measures may require the comparison of a large number of combinations of items, especially if weights are considered, leading to an increased risk of overfitting and an inflated estimate of the sensitivity and statistical power of the constructed composite. A validation procedure in a sample outside that used to identify the items and weights will be critical to accurately assess the composite's performance (Hendrix, 2012). As seen in our analysis, both types of optimization resulted in reduced power compared with the equal weight composites, likely due at least in part by overfitting the training sets. Importantly, the composites developed in this study and for the A4 study were evaluated out of sample. Neither study was able to reliably improve on equal weights.

4.6. Limitations

This study has several limitations. We evaluated assessments available in the ADNI neuropsychological battery, although it is possible or likely that there are other measures more sensitive to

decline in preclinical AD. We also did not consider item-level data from already formed composites, such as the orientation to time component of the MMSE (Langbaum et al., 2014), which may have affected the results due to carrying insensitive items along with more sensitive ones. We also make the assumption that a treatment will slow the progression of components selected for their fast decline. In reality, it is unknown which cognitive or functional components a treatment may affect and it is possible that an endpoint comprising slower progressing domains will yield more power. Additionally, we used restricted cubic splines to model the observed data and subsequently simulated clinical trials assuming an MMRM model. Maximizing the group trajectory differences assuming a spline model averages change over all time points to estimate the group curves, whereas the MMRM model allows change at each time point to be estimated more independently. Assuming an MMRM model for both steps of the analysis and allowing the weights to be differentially optimized according to trial length may yield different results. The ADNI cohort, with high levels of education possibly contributing to increased cognitive reserve, and also limited cognitive decline observed in the Aβ+ subjects compared with other cohorts, is potentially different from a population recruited for a clinical trial. The pcCN cohort is also considerably smaller with only 32 participants, reducing the stability of the estimates compared with the other cohorts. We used some of the same measures to predict conversion at screening and also track decline in the reference (equal weight) composite. It's possible that a regression to the mean effect could result in a reduction of power. However, the equal weight composite remained the most reliably performing composite with considerable power.

5. Conclusion

Our results suggest preclinical AD subjects with lower cognitive scores at baseline decline more reliably across both cognitive and functional measures compared to Aβ+ subjects without signs of subtle cognitive dysfunction. This provides a challenge to designers of preclinical AD trials to identify the level of cognitive dysfunction to be required at screening that will result in further decline, allowing a treatment effect to be demonstrated. Later stage preclinical AD may represent a more feasible target population for clinical trials designed to slow cognitive decline. In this population, suitable power for a phase III trial can be achieved with considerably lower sample sizes while capturing both cognitive and functional change to demonstrate a clinically meaningful drug effect—both while initiating treatment in subjects who are still cognitively normal. Multiple measures of delayed memory recall, orientation, processing speed, as well as multiple functional measures should be considered when forming a composite. Finally, when selecting measures, erring on the side of too many components may be preferable to too few.

Disclosure statement

Mr. Insel, Dr. Mattsson, Dr. Hansson, and Dr. Mackin report no disclosures. Dr. Donohue was a consultant for Bristol-Myers Squibb. Dr. Aisen serves on a scientific advisory board for NeuroPhage; has served as a consultant to Elan, Wyeth, Eisai, Schering-Plough, Bristol-Myers Squibb, Eli Lilly and Company, NeuroPhage, Merck, Roche, Amgen, Genentech, Abbott, Pfizer, Novartis, Bayer, Astellas, Dainippon, Biomarin, Solvay, Otsuka, Daiichi, AstraZeneca, Janssen, Medivation, Ichor, Toyama, Lundbeck, Biogen Idec, iPerian, Probiodrug, Somaxon, Biotie, Cardeus, Anavex, Kyowa Hakko Kirin Pharma, and Medtronic; and receives research support from Eli Lilly and Baxter and the NIH (NIA U01-AG10483 [PI], NIA U01-AG024904 [Coordinating Center Director], NIA R01-AG030048 [PI], and R01-AG16381 [Co-I]). Dr. Weiner has been on scientific advisory boards for Pfizer and BOLT International; has been consultant for Pfizer Inc, Janssen, KLJ Associates, Easton Associates, Harvard University, inThought, INC Research, Inc, University of California, Los Angeles, Alzheimer's Drug Discovery Foundation, and Sanofi-Aventis Groupe; has received funding for travel from Pfizer, ADPD meeting, Paul Sabatier University, Novartis, Tohoku University, MCI Group, France, Travel eDreams, Inc, Neuroscience School of Advanced Studies (NSAS), Danone Trading, BV, CTAD ANT Congres; serves as an associated editor of Alzheimer's & Dementia; has received honoraria from Pfizer, Tohoku University, and Danone Trading BV; has research support from Merck, Avid, DOD, and VA; and has stock options in Synarc and Elan.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering (P30 AG010129 and K01 AG030514), and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc; Biogen Idec Inc; Bristol-Myers Squibb Company; Eisai Inc; Elan Pharmaceuticals, Inc; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc; GE Healthcare; Innogenetics, N.V.; IXICO Ltd; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Medpace, Inc.; Merck & Co, Inc; Meso Scale Diagnostics, LLC; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc; Piramal Imaging; Servier; Synarc Inc; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). This research was also supported by The Strategic Research Area MultiPark at Lund University. 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 Southern California. This research was also supported by National Institutes of Health (NIH) grants P30 AG010129 and K01 AG030514. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.neurobiolaging.2016.08.017.

References

Ard, M.C., Raghavan, N., Edland, S.D., 2015. Optimal composite scores for longitudinal clinical trials under the linear mixed effects model. Pharm. Stat. 14, 418–426.

Ayutyanont, N., Langbaum, J.B., Hendrix, S.B., Chen, K., Fleisher, A.S., Friesenhahn, M., 2014. The Alzheimer's prevention initiative composite cognitive test score: sample size estimates for the evaluation of preclinical

Alzheimer's disease treatments in presenilin 1 E280A mutation carriers. J. Clin. Psychiatry 75, 652–660.
Breiman, L., 1996. Bagging predictors. Machine Learn. 24, 123–140.
Breiman, L., 2001. Random forests. Machine Learn. 45, 5–32.
Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. SIAM J. Scientific Comput. 16, 1190–1208.
Chambers, J.M., Hastie, T.J., 1992. Statistical Models in S. Wadsworth & Brooks/Cole.
Donohue, M.C., Sperling, R.A., Salmon, D.P., Rentz, D.M., Raman, R., Thomas, R.G., Weiner, M., Aisen, P.S., 2014. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. JAMA Neurol. 71, 961–970.
Ellis, K.A., Bush, A.I., Darby, D., De Fazio, D., Foster, J., Hudson, P., Lautenschlager, N.T., Lenzo, N., Martins, R.N., Maruff, P., Masters, C., 2009. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. Int. Psychogeriatrics 21, 672–687.
Hendrix, S.B., 2012. Measuring clinical progression in MCI and pre-MCI populations: enrichment and optimizing clinical outcomes over time. Alzheimers Res. Ther. 4, 24.
Hochberg, Y., 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75, 800–802.
Insel, P., Mattsson, N., Donohue, M.C., Mackin, S., Aisen, P., Jack, C., Shaw, L.M., Trojanowski, J.Q., Weiner, M.W., 2014. The transitional association between beta-amyloid pathology and regional brain atrophy. Alzheimers Demen. 10, P837–P838.
Insel, P.S., Mattsson, N., Mackin, R.S., Kornak, J., Nosheny, R., Tosun-Turgut, D., Donohue, M.C., Aisen, P.S., Weiner, M.W., 2015. Biomarkers and cognitive endpoints to optimize trials in Alzheimer's disease. Ann. Clin. Transl. Neurol. 2, 534–547.
Kaplan, E.F., Goodglass, H., Weintraub, S., 1982. The Boston Naming Test, second ed. Lea & Febiger, Philadelphia.
Kozauer, N., Katz, R., 2013. Regulatory innovation and drug development for early-stage Alzheimer's disease. N. Engl. J. Med. 368, 1169–1171.
Landau, S.M., Mintun, M.A., Joshi, A.D., Koeppe, R.A., Petersen, R.C., Aisen, P.S., Weiner, M.W., Jagust, W.J., 2012. Amyloid deposition, hypometabolism, and longitudinal cognitive decline. Ann. Neurol. 72, 578–586.
Langbaum, J.B., Hendrix, S.B., Ayutyanont, N., Chen, K., Fleisher, A.S., Shah, R.C., Barnes, L.L., Bennett, D.A., Tariot, P.N., Reiman, E.M., 2014. An empirically derived composite cognitive test score with improved power to track and evaluate treatments for preclinical Alzheimer's disease. Alzheimers Demen. 10, 666–674.
Mattsson, N., Insel, P.S., Donohue, M., Landau, S., Jagust, W.J., Shaw, L.M., Trojanowski, J.Q., Zetterberg, H., Blennow, K., Weiner, M.W., 2015. Independent information from cerebrospinal fluid amyloid-β and florbetapir imaging in Alzheimer's disease. Brain 138, 772–783.
Mattsson, N., Insel, P.S., Nosheny, R., Tosun, D., Trojanowski, J.Q., Shaw, L.M., Jack, C.R., Donohue, M.C., Weiner, M.W., 2014. Emerging β-amyloid pathology and accelerated cortical atrophy. JAMA Neurol. 71, 725–734.
McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., 2011. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Demen. 7, 263–269.
Morris, J.C., 1993. The clinical dementia rating (CDR): current version and scoring rules. Neurology 43, 2412–2414.
Morris, J.C., 2005. Early-stage and preclinical Alzheimer disease. Alzheimer Dis. Assoc. Disord. 19, 163–165.
Olsson, A., Vanderstichele, H., Andreasen, N., De Meyer, G., Wallin, A., Holmberg, B., Rosengren, L., VanMechelen, E., Blennow, K., 2005. Simultaneous measurement of beta-amyloid(1-42), total tau, and phosphorylated tau (Thr181) in cerebrospinal fluid by the xMAP technology. Clin. Chem. 51, 336–345.
Pfeffer, R.I., Kurosaki, T.T., Harrah Jr., C.H., Chance, J.M., Filos, S., 1982. Measurement of functional activities in older adults in the community. J. Gerontol. 37, 323–329.
Raghavan, N., Samtani, M.N., Farnum, M., Yang, E., Novak, G., Grundman, M., Narayan, V., DiBernardo, A., 2013. The ADAS-Cog revisited: novel composite scales based on ADAS-Cog to improve efficiency in MCI and early AD trials. Alzheimers Demen. 9, S21–S31.
Reitan, R.M., 1958. Validity of the trail making test as an indicator of organic brain damage. Perceptual Mot. Skills 8, 271–276.
Rey, A., 1964. L'examen clinique en psychologie. Presses Universitaires De France, Paris.
Rosen, W.G., Mohs, R.C., Davis, K.L., 1984. A new rating scale for Alzheimer's disease. Am. J. Psychiatry 141, 1356–1364.
Shaw, L.M., Vanderstichele, H., Knapik-Czajka, M., Clark, C.M., Aisen, P.S., Petersen, R.C., Blennow, K., Soares, H., Simon, A., Lewczuk, P., Dean, R., 2009. Cerebrospinal fluid biomarker signature in Alzheimer's Disease Neuroimaging Initiative subjects. Ann. Neurol. 65, 403–413.
Siddiqui, O., Hung, H.J., O'Neill, R., 2009. MMRM vs. LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets. J. Biopharm. Stat. 19, 227–246.
Snyder, P.J., Kahle-Wrobleski, K., Brannan, S., Miller, D.S., Schindler, R.J., DeSanti, S., Ryan, J.M., Morrison, G., Grundman, M., Chandler, J., Caselli, R.J., 2014. Assessing cognition and function in Alzheimer's disease clinical trials: do we have the right tools? Alzheimers Demen. 10, 853–860.
Sperling, R.A., Aisen, P.S., Beckett, L.A., Bennett, D.A., Craft, S., Fagan, A.M., Iwatsubo, T., Jack, C.R., Kaye, J., Montine, T.J., Park, D.C., 2011a. 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 Demen. 7, 280–292.
Sperling, R.A., Jack Jr., C.R., Aisen, P.S., 2011b. Testing the right target and right drug at the right stage. Sci. Transl. Med. 3, 111cm33.
Sperling, R.A., Rentz, D.M., Johnson, K.A., Karlawish, J., Donohue, M., Salmon, D.P., Aisen, P., 2014. The A4 study: stopping AD before symptoms begin? Sci. Transl. Med. 6, 228fs13.
US Dept of Health and Human Services; US Food and Drug Administration; Center for Drug Evaluation and Research. Guidance for industry: Alzheimer's disease: developing drugs for the treatment of early stage disease (draft guidance). Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM338287.pdf. Accessed August 8, 2015.
Vos, S.J., Xiong, C., Visser, P.J., Jasielec, M.S., Hassenstab, J., Grant, E.A., Cairns, N.J., Morris, J.C., Holtzman, D.M., Fagan, A.M., 2013. Preclinical Alzheimer's disease and its outcome: a longitudinal cohort study. Lancet Neurol. 12, 957–965.
Wang, J., Logovinsky, V., Hendrix, S.B., Stanworth, S.H., Perdomo, C., Xu, L., Dhadda, S., Do, I., Rabe, M., Luthman, J., Cummings, J., 2016. ADCOMS: a composite clinical outcome for prodromal Alzheimer's disease trials. J. Neurol. Neurosurg. Psychiatry 87, 993–999.
Wechsler, D.A., 1987. Wechsler Adult Intelligence Scale–Revised. Psychological Corporation, New York.