

Determining clinically meaningful decline in preclinical Alzheimer disease

Philip S. Insel, MS, Michael Weiner, MD, R. Scott Mackin, PhD, Elizabeth Mormino, PhD, Yen Ying Lim, PhD, Erik Stomrud, MD, PhD, Sebastian Palmqvist, MD, PhD, Colin L. Masters, MD, Paul T. Maruff, PhD, Oskar Hansson, MD, PhD, and Niklas Mattsson, MD, PhD

Neurology® 2019;93:e322-e333. doi:10.1212/WNL.00000000000007831

Abstract

Objective

To determine the time required for a preclinical Alzheimer disease population to decline in a meaningful way, use estimates of decline to update previous clinical trial design assumptions, and identify factors that modify β -amyloid ($A\beta$)–related decline.

Methods

In 1,120 cognitively unimpaired individuals from 3 international cohorts, we estimated the relationship between $A\beta$ status and longitudinal changes across multiple cognitive domains and assessed interactions between $A\beta$ and baseline factors. Power analyses were performed to explore sample size as a function of treatment effect.

Results

Cognitively unimpaired $A\beta+$ participants approach mild cognitive impairment (MCI) levels of performance 6 years after baseline, on average. Achieving 80% power in a simulated 4-year treatment trial, assuming a 25% treatment effect, required 2,000 participants/group. Multiple factors interacted with $A\beta$ to predict cognitive decline; however, these findings were all cohort-specific. Despite design differences across the cohorts, with large sample sizes and sufficient follow-up time, the $A\beta+$ groups declined consistently on cognitive composite measures.

Conclusions

A preclinical AD population declines to the cognitive performance of an early MCI population in 6 years. Slowing this rate of decline by 40%–50% delays clinically relevant impairment by 3 years—a potentially meaningful treatment effect. However, assuming a 40%–50% drug effect highlights the difficulties in preclinical AD trial design, as a more commonly assumed treatment effect of 25% results in a required sample size of 2,000/group. Designers of preclinical AD treatment trials need to prepare for larger and longer trials than are currently being considered. Interactions with $A\beta$ status were inconsistent and not readily generalizable.

Correspondence

Mr. Insel
philipinsel@gmail.com

RELATED ARTICLE

Editorial

The search for meaning in preclinical Alzheimer disease clinical trials

Page 139

MORE ONLINE

Podcast

Dr. Gregory Day talks with Dr. Philip Insel about his paper on determining clinically meaningful decline in preclinical Alzheimer disease.

[NPub.org/d7tb1d](https://www.ncbi.nlm.nih.gov/podcasts/npub/d7tb1d)

From the Center for Imaging of Neurodegenerative Diseases (M.W., R.S.M.), Department of Veterans Affairs Medical Center; Departments of Radiology and Biomedical Imaging (P.S.I., M.W.) and Psychiatry (P.S.I., R.S.M.), University of California, San Francisco; Clinical Memory Research Unit, Faculty of Medicine (P.S.I., E.S., S.P., O.H., N.M.), Memory Clinic (E.S., S.P., O.H.) and Department of Neurology (N.M.), Skåne University Hospital, and Wallenberg Center for Molecular Medicine (N.M.), Lund University, Sweden; Department of Neurology and Neurological Sciences (E.M.), Stanford University, CA; The Florey Institute (Y.Y.L., C.L.M., P.T.M.), The University of Melbourne; and CogState (P.T.M.), Melbourne, Australia.

Go to [Neurology.org/N](https://www.neurology.org/N) for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article. The Article Processing Charge was funded by Swedish Research Council.

This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Glossary

A β = β -amyloid; **AD** = Alzheimer disease; **ADNI** = Alzheimer's Disease Neuroimaging Initiative; **AIBL** = Australian Imaging, Biomarkers & Lifestyle; **AIC** = Akaike information criterion; **BioFINDER** = Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably; **CDR** = Clinical Dementia Rating; **CDR-SB** = CDR sum of boxes; **dADASc** = Delayed Word Recall from the Alzheimer's Disease Assessment Scale–Cognitive Subscale; **dMemory** = Logical Memory Delayed Recall; **MCI** = mild cognitive impairment; **MMSE** = Mini-Mental State Examination; **OR** = odds ratio; **PACC** = Preclinical Alzheimer's Cognitive Composite; **PiB** = Pittsburgh compound B; **SUVr** = standardized uptake value ratio; **Trails B** = Trail-Making Test B.

To effectively alter the course of Alzheimer disease (AD), interventions may need to occur during the preclinical stage of the disease, before the onset of clinical symptoms.¹ Demonstrating that treatments are effective during the preclinical stage will require understanding the magnitude of early β -amyloid (A β)–related cognitive decline in cognitively unimpaired adults.² Defining meaningful decline will help determine the time frame for subtle cognitive changes to progress to incipient functional decline and to identify an optimal treatment window.

The association between A β status and cognition in preclinical AD varies widely.^{3–9} The design of the A4 study,¹⁰ the first clinical trial in preclinical AD, was based on early estimates of A β -related decline using the Alzheimer's Disease Neuroimaging Initiative (ADNI)¹¹ and the Australian Imaging, Biomarkers & Lifestyle (AIBL) Study.¹² The effect of A β on cognitive decline in AIBL was 4-fold the magnitude of the effect in ADNI, highlighting an inconsistent picture of early cognitive decline and uncertain implications for powering a trial in early AD. Understanding how sampling variation and study design features influence estimates of cognitive decline will optimize the design of trials in preclinical AD.

The aims of this study were to harmonize several large studies in order to (1) determine the time required for a preclinical AD population to decline in a clinically meaningful way, (2) characterize how decline differs by cognitive domain, (3) update previous study design assumptions regarding sample size, power, and the required treatment effect, and (4) identify factors that modify A β -related decline.

Methods

Standard protocol approvals, registrations, and patient consents

This study was approved by the institutional review boards of all of the participating institutions. Informed written consent was obtained from all participants at each site.

Participants

Participants from each of the cohorts ADNI, AIBL, and the Swedish Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably (BioFINDER) study¹³ were included if they were classified as cognitively normal at baseline, were tested for A β biomarkers (using either CSF or PET), and

were followed longitudinally with neuropsychological examinations.^{11–13} Participants were excluded from any of the 3 studies if they had a major neurologic or psychiatric illness or a history of substance abuse. In addition, ADNI participants were excluded if the screening MRI showed evidence of infection, infarction, or other focal lesions, including multiple lacunes or lacunes in a critical memory structure. MRI results were not part of the exclusionary criteria for AIBL or BioFINDER, but BioFINDER participants were excluded if they refused MRI or lumbar puncture. Detailed exclusionary criteria for ADNI can be found at adni.loni.usc.edu/wp-content/uploads/2008/07/adni2-procedures-manual.pdf and for BioFINDER at biofinder.se/biofinder_cohorts/cognitively-healthy-elderly/. We also included 305 participants enrolled into the early mild cognitive impairment (MCI) cohort in ADNI (defined by a subjective memory complaint and a delayed logical memory score of 9–11 for those with 16 or more years of education, 5–9 for 8–15 years of education, or 3–6 for 0–7 years of education, where possible scores range from 0 to 25)¹⁴ for a comparative analysis. The extensions of ADNI introduced the distinction of MCI into early and late MCI in the attempt to define an earlier point in time for disease detection. Late MCI refers to the original definition of MCI (performance for 1.5 SD below the normative mean), whereas in early MCI, impairment is defined as performance between 1.0 SD and 1.5 SD below the normative mean on a standard test. Because of recent evidence of an artificially low reversion rate from MCI to control in ADNI,¹⁵ we excluded 7 early MCI participants who consistently had a global Clinical Dementia Rating (CDR) score of zero after screening in a sensitivity analysis.

Data on memory complaints in the controls were available in AIBL and ADNI. In AIBL, participants with a memory complaint were identified by the response to the question, “Do you have difficulties with your memory?” In ADNI, the participant was required to have a significant memory concern as reported by the participant, study partner, or clinician and a score >16 on the first 12 items of the Cognitive Change Index.

A β biomarkers

A β status was defined by PET imaging if available (all AIBL and a majority of ADNI participants), and otherwise by CSF biomarkers (all BioFINDER and a small proportion of ADNI participants). PET imaging was done using ¹⁸F-florbetapir PET in ADNI and using ¹⁸F-florbetapir, ¹¹C–Pittsburgh

compound B (PiB), or ^{18}F -flutemetamol PET in AIBL. Methods to acquire and process imaging data were described previously.^{16–18} CSF samples were collected at baseline by lumbar puncture. CSF methods have been described previously.^{19–21} In short, ADNI CSF samples were analyzed for CSF A β 42 using the AlzBio3 assay (Fujirebio, Ghent, Belgium) on the xMAP Luminex platform. BioFINDER CSF samples were analyzed for CSF A β 42 and A β 40 using ELISA assays (ADx/EUROIMMUN AG, Lübeck, Germany). For ADNI participants, A β + was defined as ^{18}F -florbetapir PET standardized uptake *evaluation* (*SUVR*) >1.1 ($n = 381$)²² or CSF A β 42 <192 ng/L ($n = 62$).¹⁹ For AIBL, A β + was defined as ^{18}F -florbetapir PET *SUVR* >1.1 ($n = 72$), ^{11}C -PiB PET *SUVR* >1.5 ($n = 201$), or ^{18}F -flutemetamol *SUVR* >0.62 ($n = 75$).²³ In BioFINDER, A β + was defined as CSF A β 42/A β 40 <0.1.²⁴

Cognitive testing

Participants were followed for up to 6 years for neuropsychological testing. In ADNI, tests were administered annually with an additional test at month 6 for most measures. In AIBL, tests were administered every 18 months. In BioFINDER, tests were administered every 2 years. The Preclinical Alzheimer's Cognitive Composite (PACC)²⁵ and its individual components were the primary outcomes compared in the 3 cohorts. This composite was developed specifically to be sensitive to early cognitive changes in AD and is being incorporated in clinical trials of disease-modifying treatments.¹⁰ Substitutions representing the same cognitive domain were made in the case where the original PACC components were not available or had limited follow-up in a cohort's neuropsychological battery, following previous procedures.^{10,25} Visits where all components or substitutions were available were included. For ADNI, the modified PACC comprised the Mini-Mental State Examination (MMSE), Logical Memory Delayed Recall (dMemory), Trail-Making Test B (Trails B), and the Delayed Word Recall from the Alzheimer's Disease Assessment Scale–Cognitive Subscale (dADASc). For AIBL, the PACC was constructed using the MMSE, dMemory, Digit Symbol Substitution Test, and the Delayed Recall from the California Verbal Learning Test (dCVLT). For BioFINDER, the PACC consisted of the MMSE, dADASc, and Trails B. To calculate the composite, z scores of the individual components were taken over all time points and then summed. This sum was then standardized to the mean and SD of the baseline score of the sum.

The PACC includes 2 measures of delayed memory recall; however, because only one delayed memory measure was available in BioFINDER, dADASc was given twice the weight in BioFINDER to reflect the contribution of delayed memory recall in the composite. Immediate recall (logical memory for ADNI and AIBL, Alzheimer's Disease Assessment Scale–Cognitive Subscale word recall for BioFINDER) was evaluated as a measure of baseline memory ability to predict changes in the PACC. The CDR sum of

boxes (CDR-SB) was also evaluated as an outcome measure.

Statistical analysis

Longitudinal measures were modeled using generalized least squares regression assuming a compound symmetric covariance structure.²⁶ To capture departures from linearity in the trajectory of the neuropsychological measures, continuous time from baseline test was parameterized using restricted cubic splines.²⁷ Cubic splines are functions of polynomials allowing flexibility in the estimation of trajectories over time. Time was modeled with 3 spline knots, 2 at the boundaries and 1 at median follow-up. Differences in trajectories between A β + and A β – groups were tested using interactions between the 2 measures for time and the group factor using likelihood ratio tests and change in the Akaike information criterion (AIC), a model selection tool.²⁸ A lower value of AIC indicates a better fitting model. Baseline age was also modeled using restricted cubic splines to capture its nonlinear effect on cognition. Models included the 2 spline-estimated measures for baseline age; sex; years of education, where education was categorized as 0–12 years, 13–15 years, and 16 or more years; the interaction between A β status and the 2 measures for time; and the main effects for A β status and time.

We also evaluated interactions between A β status and demographics (baseline age, sex, education), APOE (presence of at least one $\epsilon 4$ allele), memory complaint, and baseline memory, and their effect on changes in the PACC. These models included all the terms described above as well as the 3-way interaction between time, A β status, and the demographic term. The interaction with age was evaluated using the 2 spline-estimated measures.

To estimate power for hypothetical clinical trials, mixed models of repeated measures²⁹ were used to estimate the variance components of the change from baseline in the PACC for the A β + subjects in each cohort. To mirror current preclinical trial design,¹⁰ A β + subjects with very high cognitive scores (dMemory >15 for ADNI [$n = 32$] and AIBL [$n = 12$] and dADASc >8 in BioFINDER [$n = 29$]) were excluded in order to remove “supernormals.” This was done to mitigate the inclusion of participants with little or no sign of near-term decline in order to increase the likelihood of decline in the placebo group and improve power. Model estimates were then used to calculate the power for 4- and 6-year clinical trials, assuming a range of sample sizes and drug effects, a 6-month visit interval, and a 30% dropout rate. Individual cohort estimates of change from baseline and variance were then meta-analyzed to get combined estimates of change over time.³⁰

In order to provide a context for meaningful clinical decline in the cognitively normal participants, we compared the baseline PACC scores in the normal participants to the PACC scores in the ADNI early MCI participants (stratified by A β status). We then evaluated the mean time for the average preclinical

AD participant to reach the mean baseline PACC score in the early MCI groups.

Baseline associations between demographics and A β positivity were assessed using the Wilcoxon rank-sum test for continuous variables and a χ^2 test for categorical variables. Reductions of AIC >2 and *p* values <0.05 were considered significant. All analyses were done in R v3.4.3 (r-project.org). GLS models were fit using the *gls* function from the *nlme* package.

Data availability

Data from the ADNI and AIBL cohorts are publicly available. Data from BioFINDER may be requested.

Results

Cohort characteristics

A total of 443 cognitively healthy controls from ADNI, 348 from AIBL, and 329 from BioFINDER were included in the study. A β + groups were older, had a higher frequency of *APOE* ϵ 4 positivity, and performed significantly worse on several cognitive tests at baseline, compared to A β - groups, in all cohorts (table 1). The proportion of *APOE* ϵ 4 positivity in the A β + group was similar in BioFINDER (55%) and AIBL (53%) and lower in ADNI (44%). Education and sex were not associated with A β positivity in AIBL or BioFINDER; however, A β + ADNI participants were more likely to be female and have less education compared to A β - ADNI participants. The majority of ADNI participants had 16 or more years of education, whereas the majority of both AIBL and BioFINDER participants had fewer than 16 years of education. There was no association between subjective memory complaint and A β status in either ADNI or AIBL (subjective memory complaint data were not available in BioFINDER).

There was considerable variability in attrition rates across the 3 cohorts. At 4 years of follow-up, ADNI retained 46% of its participants; however, dropout was not associated with age, sex, education, A β status, or baseline memory performance (*p* > 0.13). At 4 years, BioFINDER retained 69% of its participants. Women were less likely to drop out (odds ratio [OR] = 0.78, *p* = 0.01), participants with more education were more likely to drop out (OR = 1.35, *p* = 0.04), and older age was associated with increased drop out (OR = 1.28 for 1 SD increase in age, *p* < 0.001). AIBL retained 90% of its participants, but older age was associated with increased drop out (OR = 1.26 for 1 SD increase in age, *p* = 0.01).

Cognitive changes

A β + participants declined significantly more on the PACC and all individual components of the PACC compared to A β - participants, in all 3 cohorts, with the exception of Trails B in BioFINDER (*p* = 0.08). Estimates and longitudinal plots of cognition are shown in figure 1. Estimates of the change from baseline, confidence intervals, and the residual SD for each visit and group are shown in table 2.

At year 4, the A β + groups declined by -0.45 points on the PACC (ADNI), -0.48 points (BioFINDER), and -0.53 points (at 4½ years, AIBL) (table 2). At year 4, the A β - group improved 0.09 points on the PACC in ADNI and declined by -0.14 points in BioFINDER and -0.02 points in AIBL.

Clinical significance

To evaluate decline and to characterize what might be considered a clinically significant change, we compared the scores of the cognitively normal participants to the baseline scores of the early MCI participants in ADNI. The mean PACC score in A β - and A β + early MCI participants at baseline was -1.01 and -1.30, respectively (figure 2). Six years after baseline, the estimated PACC score combined across cohorts of the pre-clinical AD groups was midway between the A β - and A β + early MCI performance. Similarly, the early MCI A β - and A β + scores at baseline on the CDRSB were 1.22 and 1.38, respectively, whereas the preclinical AD groups averaged about 1.0 at 6 years.

On each of the MMSE, delayed list learning, and executive function, the cognitively normal A β + groups averaged worse scores than both MCI groups by 6 years after baseline. The cognitively normal A β + groups did not approach the MCI groups' delayed logical memory scores by 6 years after baseline. Note that delayed logical memory was not available in BioFINDER.

In a sensitivity analysis, 7 early MCI participants who consistently had a global CDR of zero after screening were excluded. The reduced sample scores were slightly worse than the full MCI sample with A β - and A β + PACC scores of -1.02 and -1.33, respectively, and CDR-SB scores of 1.23 and 1.39.

Power

Using estimates of change and variance, we calculated the power for hypothetical 4- and 6-year clinical trials for each cohort, assuming a 30% dropout rate, and various sample sizes and drug effects (figure 3). In 4-year trials, assuming a 25% drug effect, i.e., a 25% slowing of cognitive decline in the treatment group, the required sample size to reach 80% power was 2,000 per group for the estimate combining all cohorts. Assuming a larger effect size of 35%, the required sample size to reach 80% power was 1,000 per group on average.

In 6-year trials, assuming a 25% drug effect, the required sample size to reach 80% power was about 600 per group for the estimate combining all cohorts. Assuming a 35% effect size, the required sample size to reach 80% power was 300 per group on average.

A β interactions

The interactions between A β status and baseline factors to predict cognitive decline on the PACC were also assessed. Plots of the amyloid groups at different levels of the significant interacting factors, *p* values, and the change in AIC are shown in figure 4. In AIBL, there were significant interactions between

Table 1 Baseline characteristics

Characteristic	A β +	A β -	<i>p</i> Value
ADNI	N = 165	N = 278	
Years of follow-up	4.1 (2.8)	4.3 (2.9)	0.28
Age	75.1 (5.5)	73.3 (5.9)	0.001
Female, n (%)	99 (60)	132 (47.5)	0.01
Education, y			0.002
0-12	18 (10.9)	25 (9.0)	
13-15	46 (27.9)	42 (15.1)	
16+	101 (61.2)	211 (75.9)	
Memory complaint, n (%)	42 (25.5)	64 (23)	0.64
APOE ϵ 4+, n (%)	73 (44.2)	53 (19.1)	<0.001
MMSE	29.1 (1.1)	29.0 (1.2)	0.68
dMemory	12.8 (3.4)	13.4 (3.2)	0.05
dADASc	7.0 (1.8)	7.2 (1.8)	0.33
Trails B	93.8 (44.4)	79.7 (39.4)	<0.001
BioFINDER	N = 85	N = 244	
Years of follow-up	3.6 (1.8)	3.6 (1.7)	0.55
Age	74.7 (5.0)	73.3 (5.0)	0.02
Female, n (%)	56 (65.9)	142 (58.2)	0.26
Education, y			0.99
0-12	51 (60.0)	146 (59.8)	
13-15	20 (23.5)	58 (23.8)	
16+	14 (16.5)	40 (16.4)	
Memory complaint, n (%)	—	—	—
APOE ϵ 4+, n (%)	46 (54.8)	46 (19)	<0.001
MMSE	29.0 (0.9)	29.1 (1.0)	0.24
dMemory	—	—	—
dADASc	7.4 (2.2)	8.2 (1.8)	0.001
Trails B	111.8 (48.7)	102.4 (50.8)	0.04
AIBL	N = 100	N = 248	
Years of follow-up	4.9 (1.9)	5.9 (2.9)	<0.001
Age	73.5 (7.3)	69.1 (6.0)	<0.001
Female, n (%)	49 (49)	136 (54.8)	0.39
Education, y			0.77
0-12	40 (40.4)	104 (41.9)	
13-15	23 (23.2)	42 (16.9)	
16+	36 (36.4)	102 (41.1)	
Memory complaint, n (%)	58 (58.6)	132 (53.4)	0.45
APOE ϵ 4+, n (%)	53 (53)	58 (23.4)	<0.001

Table 1 Baseline characteristics (*continued*)

Characteristic	A β +	A β -	<i>p</i> Value
MMSE	28.7 (1.2)	29.0 (1.2)	0.04
dMemory	11.1 (4.11)	12.1 (4.0)	0.04
dCVLT	11.5 (3.4)	12.2 (2.9)	0.10
Digit symbol	57.9 (12.9)	61.3 (13.7)	0.05

Abbreviations: A β = β -amyloid; ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging, Biomarkers & Lifestyle; BioFINDER = Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably; dADASc = Delayed Word Recall from the Alzheimer's Disease Assessment Scale-Cognitive Subscale; dCVLT = Delayed Recall from the California Verbal Learning Test; dMemory = Logical Memory Delayed Recall; MMSE = Mini-Mental State Examination; Trails B = Trail-Making Test B.

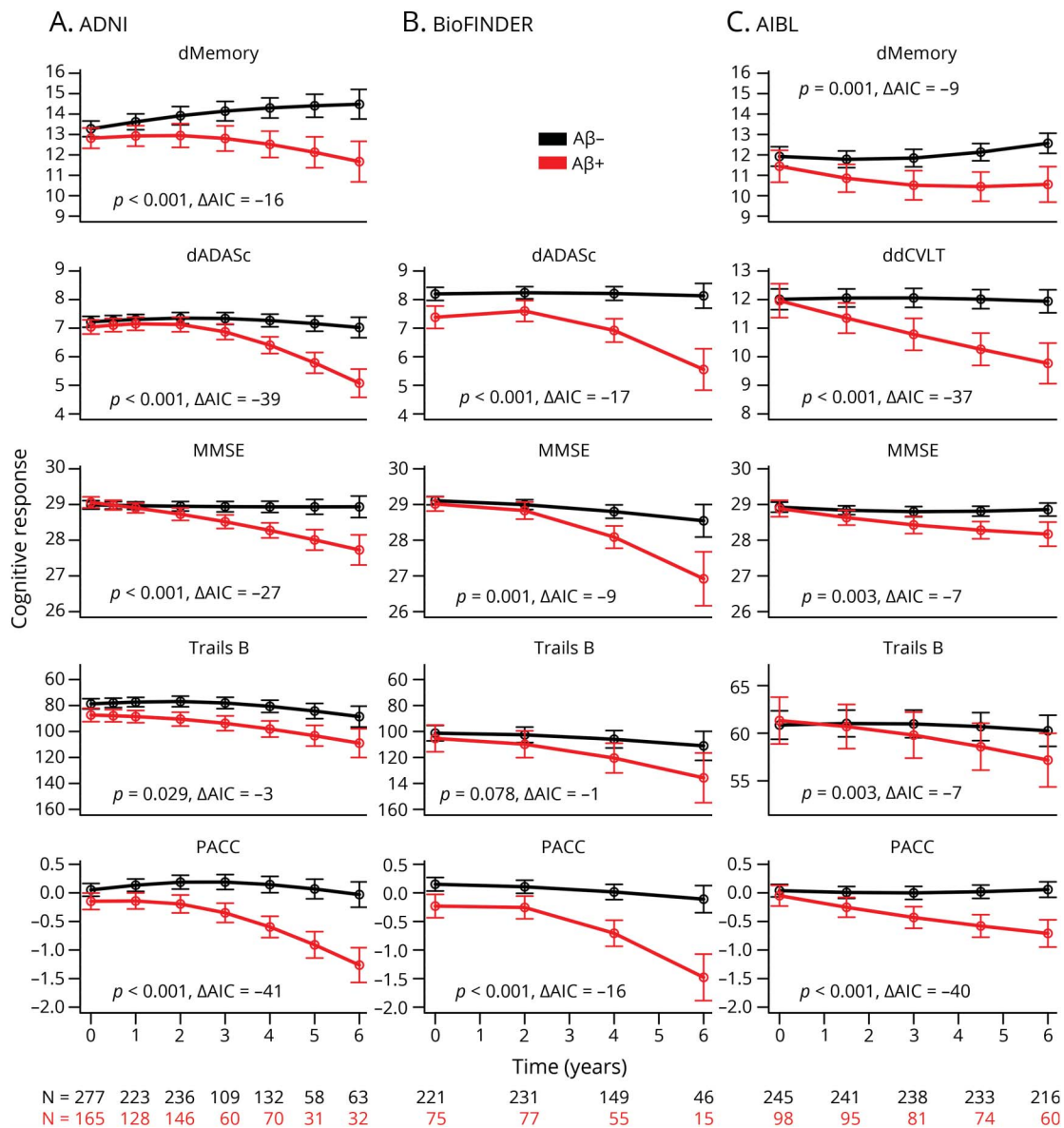
A β and education, APOE ϵ 4 positivity, and baseline memory. The only significant interaction in ADNI was between A β and sex, and the only significant interaction in BioFINDER was between A β and age. There were no significant interactions between A β and subjective memory complaint (ADNI: *p* = 0.56, AIBL: *p* = 0.87, not available for BioFINDER).

Discussion

The main findings of this study are (1) cognitively unimpaired A β + participants approach early MCI cognitive performance levels on general cognition and global outcomes, delayed list recall, and executive function by 6 years after baseline; (2) to achieve 80% power in a simulated treatment trial assuming a 25% treatment effect, 2,000 participants/group for a 4-year trial and 600 participants/group for a 6-year trial are required; (3) several baseline factors interacted with A β status to predict decline on the PACC including APOE ϵ 4 positivity, memory, and education in AIBL; age in BioFINDER; and sex in ADNI, although these findings were all cohort-specific; (4) despite considerable design differences across the cohorts, with large sample sizes and sufficient follow-up time, the cognitively unimpaired A β + groups declined consistently on cognitive composites; (5) A β + groups declined significantly faster on all cognitive tests in all cohorts, with the exception of Trails B in BioFINDER, where the A β + group declined marginally faster (*p* = 0.08), compared to the A β - group.

A key question for preclinical AD trials is how to define meaningful outcomes that will support use of therapeutic interventions in people who may remain asymptomatic for many years even without treatment. Traditional AD dementia trials are frequently powered to detect a several-point difference on a global cognitive score (e.g., Alzheimer's Disease Assessment Scale-Cognitive Subscale), as well as a global/functional co-primary outcome to establish clinical meaningfulness.³¹ Post hoc analyses of the first large trials of solanezumab in patients with mild AD showed a 34% reduction of cognitive decline and a 17% reduction of functional

Figure 1 Cognitive change over time



Cognitive responses are plotted over time for each β -amyloid ($A\beta$) group, in each cohort separately: (A) Alzheimer's Disease Neuroimaging Initiative (ADNI), (B) Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably (BioFINDER), and (C) Australian Imaging, Biomarkers & Lifestyle (AIBL). Individual Preclinical Alzheimer's Cognitive Composite (PACC) components are shown as well as the PACC in the bottom row. Akaike information criterion and p values are shown in each plot, testing for differences between $A\beta$ groups over time. dMemory = Logical Memory Delayed Recall; MMSE = Mini-Mental State Examination; Trails B = Trail-Making Test B.

decline.³² However, these effects were not replicated in a subsequent randomized trial, which failed to show a significant treatment effect, with only an 11% reduction of cognitive decline and 15% reduction of functional decline.³³ In preclinical AD, the cognitive decline observed over 3–4 years is subtle, and is typically accompanied by little or no functional decline.³⁴ However, it has not been clarified what degree of decline would warrant classification as meaningful decline. To benchmark the magnitude of cognitive decline to a measure of clinical meaningfulness, we compared the scores of the cognitively unimpaired participants to those classified as early MCI—a group with incipient functional decline. The separation between these groups was just over 1 SD on the PACC, suggesting that 1 point

of additional decline in $A\beta^+$ participants compared to $A\beta^-$ participants could be taken as an approximate benchmark for clinically meaningful decline. Combining results across cohorts shows the average $A\beta^+$ participant to have the same PACC score at 6 years post baseline as the average patient with early MCI had at baseline (figure 2). $A\beta^+$ participants also reached MCI level performance at 6 years on the other cognitive outcomes, with the exception of delayed logical memory. Possible explanations for this exception include that this measure was used as inclusion criterion for enrollment. This measure was also not available in BioFINDER, the cohort demonstrating the poorest scores on all measures by the end of follow-up. Finally, delayed logical memory demonstrated a clear practice effect in

Table 2 PACC: Change from baseline, 95% CI, and residual SD estimates

Study	Month	A β +		A β -		Difference (Δ_{diff})	95% CI	Residual SD (σ)	Δ_{diff}/σ
		N	Estimate	N	Estimate				
ADNI	12	128	0.01	223	0.08	-0.08	-0.16 to 0.01	0.89	-0.09
	24	146	-0.05	236	0.14	-0.18	-0.32 to -0.04	0.89	-0.21
	36	62	-0.20	109	0.14	-0.34	-0.50 to -0.18	0.91	-0.37
	48	70	-0.45	132	0.09	-0.54	-0.72 to -0.37	1.03	-0.53
	60	31	-0.76	58	0.02	-0.78	-1.00 to -0.56	1.15	-0.68
	72	32	-1.12	63	-0.08	-1.03	-1.35 to -0.72	1.32	-0.78
BioFINDER	24	75	-0.02	221	-0.04	0.02	-0.16 to 0.20	0.70	0.03
	48	55	-0.48	149	-0.14	-0.34	-0.56 to -0.12	0.83	-0.41
	72	15	-1.25	46	-0.26	-0.99	-1.40 to -0.57	1.29	-0.77
AIBL	18	95	-0.20	241	-0.03	-0.17	-0.29 to -0.05	0.78	-0.22
	36	81	-0.38	238	-0.04	-0.34	-0.52 to -0.17	0.83	-0.41
	54	74	-0.53	233	-0.02	-0.51	-0.69 to -0.34	1.04	-0.49
	72	60	-0.66	216	0.02	-0.68	-0.88 to -0.48	0.98	-0.70

Abbreviations: A β = β -amyloid; ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging, Biomarkers & Lifestyle; BioFINDER = Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably; CI = confidence interval; PACC = Preclinical Alzheimer's Cognitive Composite.

the A β - group (figure 2), with the cognitively unimpaired participants taking this test 6 times over follow-up, compared to one time for the MCI participants.

Based on the PACC estimates, a treatment effect of 40%–50% would be required to delay the cognitive decline of a group of A β + participants from reaching the 1 SD milestone by 3 years. Delaying the cognitive decline equivalent to the level of the average early MCI patient by 3 years may be a clinically meaningful treatment effect. But 40%–50% is a large treatment effect and highlights the difficulties in preclinical AD trial design. However, the observation that clinically meaningful decline is reached within 6 years offers strong support for the use of a cognitive composite in trials that are shorter than 6 years, since short term cognitive decline can be conceptualized as a proxy for downstream functional changes. With meaningful continuous cognitive changes occurring prior to an MCI diagnosis, these results, as well as recent reports,³⁵ argue against the use of a time-to-MCI endpoint in preclinical AD trials.

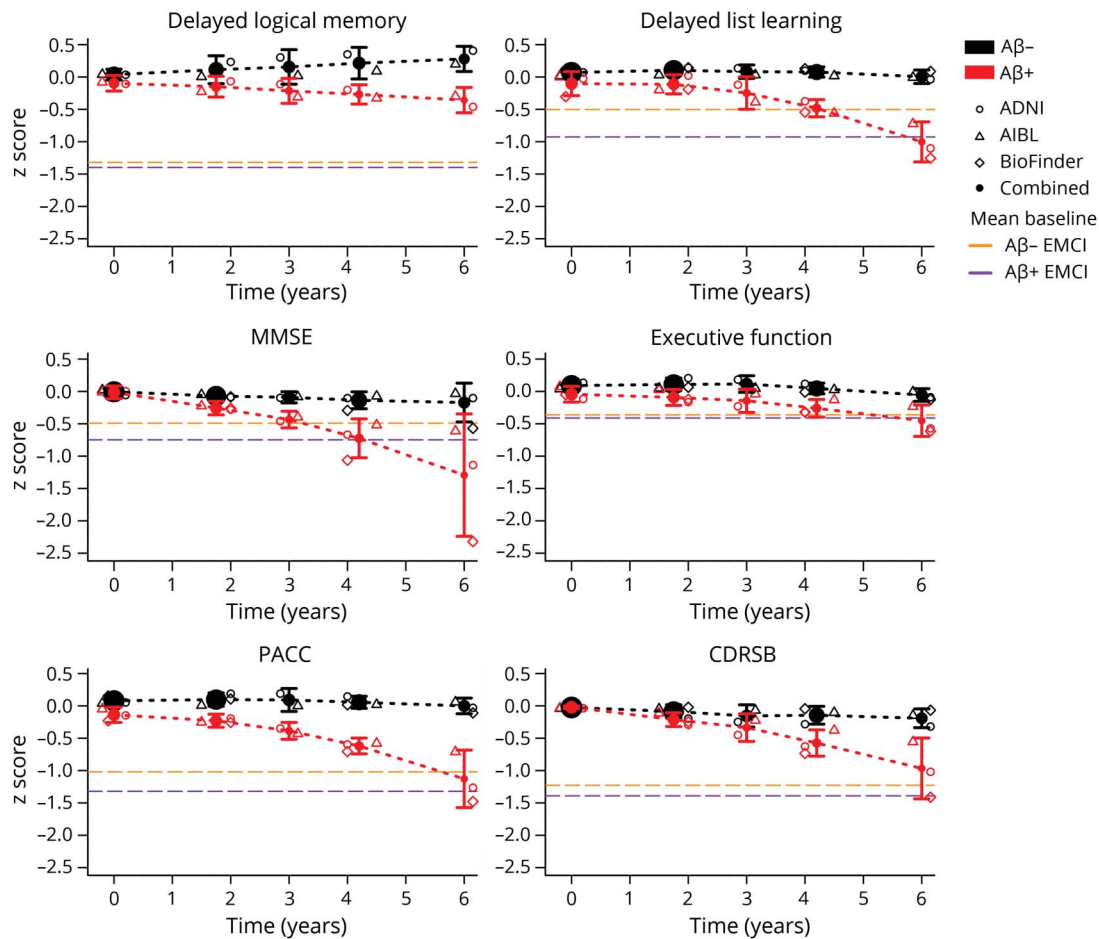
The estimated sample size or trial length requirements are sobering. Previously reported sample size and drug effect requirements of 500/group with a 30%–50% effect size in a 3-year trial were optimistic and based on approximately 20% of the data available in this study.¹⁰ In order to reliably achieve 80% power for a modest, real-world effect size of 20%–30%, investors in AD research for therapeutics development will have to prepare to support larger and longer trials than are currently envisaged.

There were several significant interactions between A β status and baseline factors. However, no interaction was observed in

more than one cohort. In AIBL, the combination of A β status and low education, *APOE* ϵ 4 positivity, or low baseline memory all led to increased rates of decline on the PACC. Decline in the A β + groups did not depend on *APOE* ϵ 4 status in ADNI or BioFINDER; however, in AIBL, little decline was observed in A β + participants who were not also *APOE* ϵ 4+ (figure 4), as was reported previously.^{36,37} Evidence for additional risk of cognitive decline for individuals who are both A β + and *APOE* ϵ 4+ had been incorporated into the design of a phase 2b/3 trial in preclinical AD (clinicaltrials.gov/ct2/show/NCT02569398); however, this pattern was observed in only one of the 3 cohorts studied here. The additional decline observed in the A β + participants who also had low baseline memory in AIBL is consistent with previous reports.³⁸ Still, despite wide separation at baseline, high and low baseline memory (and also high and low education) groups declined in parallel over time in both ADNI and BioFINDER. The lack of replicability of these interactions across cohorts suggests that if there are true underlying effects of these baseline factors that modify the A β /cognition relationship, they are mild, or they depend on other/complex interactions. Another possibility is that their identification was the consequence of type I error, although the strength of the associations in AIBL (but not ADNI, reported previously³⁹ or BioFINDER) would survive a Bonferroni correction. Our findings caution against relying on interactions between A β and demographic/clinical factors when selecting participants for preclinical AD trials.

There were considerable design differences among the 3 study cohorts including differences in geographic region, cognitive measures, visit frequency, and sampling characteristics.

Figure 2 Meta-estimates of change



Meta-estimates of change over time are shown by β -amyloid ($A\beta$) group. Individual cohort estimates are also shown. The mean baseline early mild cognitive impairment scores are shown in dashed purple for $A\beta+$ and dashed orange for $A\beta-$. ADNI = Alzheimer's Disease Neuroimaging Initiative; AIBL = Australian Imaging, Biomarkers & Lifestyle; BioFINDER = Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably; CDRSB = CDR sum of boxes; EMCI = early mild cognitive impairment; MMSE = Mini-Mental State Examination; PACC = Preclinical Alzheimer's Cognitive Composite.

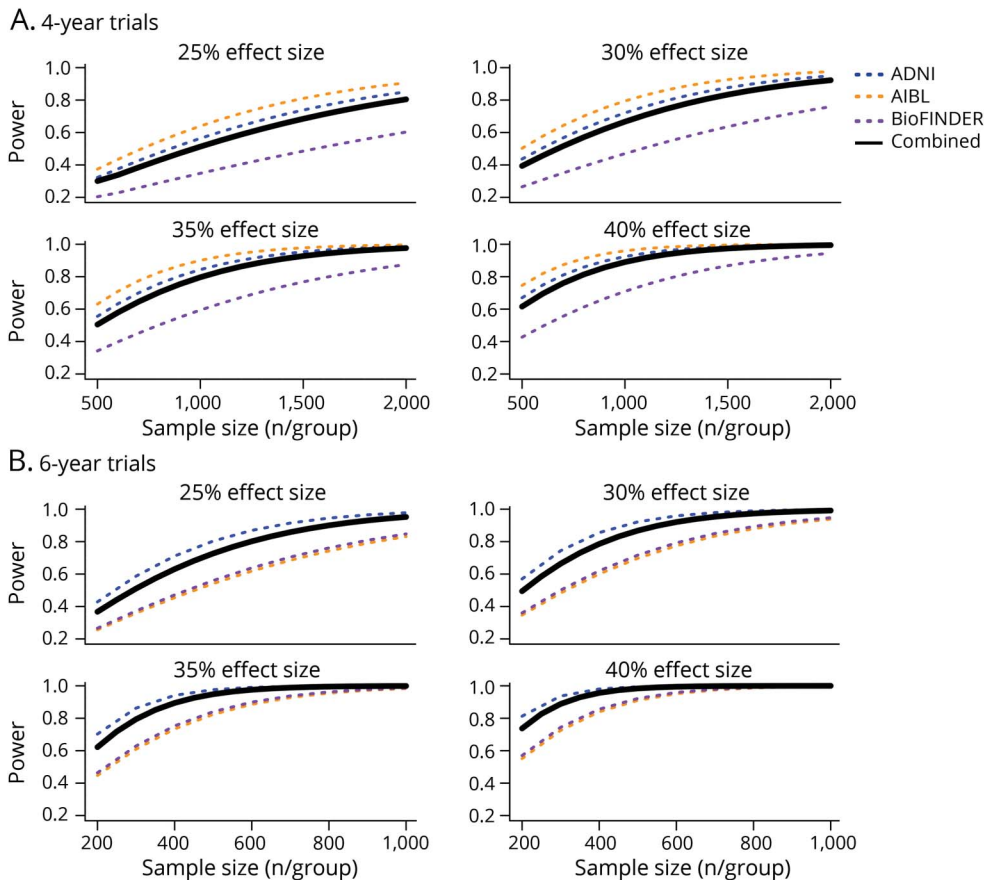
Despite these differences, the estimates of decline observed on the PACC in the $A\beta+$ groups at 4 years were remarkably similar: -0.45 points in ADNI, -0.48 in BioFINDER, and -0.53 (at 4½ years) in AIBL (table 2). Where the cohorts differed was in the change in the $A\beta-$ group: 0.09 in ADNI, -0.14 in BioFINDER, and -0.02 in AIBL. The lower power estimate for BioFINDER for a clinical trial can be traced back to the additional decline observed in the $A\beta-$ group, which may be due in part to including participants with presence of cerebrovascular pathology such as white matter lesions (not excluded from BioFINDER, but may have been excluded from ADNI).^{40,41} Cognitive reserve may also play a role, given the lower levels of education in BioFINDER compared to both ADNI and AIBL.

The $A\beta$ group trajectories on the PACC were similar, though there was variation in the shape of the trajectories for some of the individual components. One design feature that may influence trajectory differences is test frequency. ADNI participants were tested every 6 months over the first year and every

year thereafter, whereas AIBL participants were tested every 18 months and BioFINDER, every 24 months. The increased test frequency and higher levels of education in ADNI may have contributed to a tendency to improve over time as seen in dMemory (figure 1). Despite this variation in dMemory slope, $A\beta$ group separation over time was preserved in ADNI and AIBL. For delayed list learning, all $A\beta-$ groups remained stable, and all $A\beta+$ groups showed similar decline over the total follow-up time. Combining individual components into the composite seemed to mitigate individual domain trajectory differences (figure 2). Overall, the $A\beta$ groups across all 3 cohorts started to diverge reliably around 3 years after baseline.

One of the main limitations of this study is the variation of available measures used to construct the composite cognitive scores (i.e., the PACC) in each of the cohorts. While we included the domains represented in the original PACC, it remains unclear how these substitutions may affect the estimates of $A\beta$ -related cognitive decline. Another limitation is

Figure 3 Power



Hypothetical clinical trial power is plotted against sample size per treatment group, for each of 4 assumed treatment effect sizes and 2 trial lengths. Individual cohort power curves and a combined estimate are shown. Sample sizes range from (A) 500 to 2,000 per group for 4-year trials and (B) 200 to 1,000 per group for 6-year trials. ADNI = Alzheimer's Disease Neuroimaging Initiative; BioFINDER = Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably.

that with strict exclusionary criteria, the participants in these studies have few comorbidities, lack diversity, and do not mirror the general population. Clinical trials frequently use similar exclusionary criteria and may also lack generalizability. An additional limitation to all studies trying to inform disease-modifying AD trials is that without any information regarding potential effects of treatments, the power to detect a hypothetical effect is speculative.

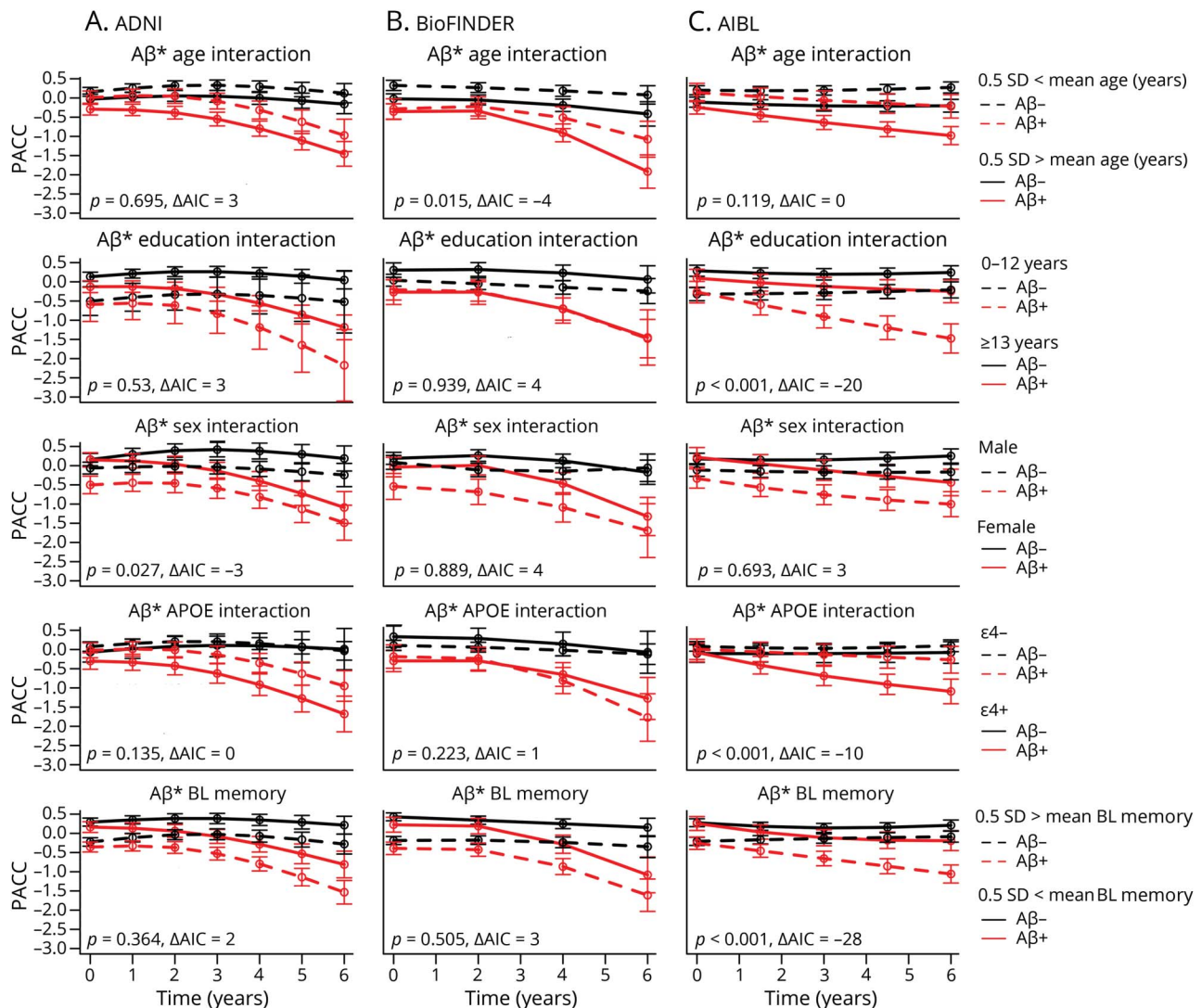
Average cognitively normal A β + participants approach early MCI cognitive performance levels 6 years after baseline. Comparing these 3 cohorts side by side demonstrates that large sample sizes and sufficiently long follow-up times result in consistent estimates of decline in preclinical AD. Despite substantial design and sampling differences, these results support the potential for internationally conducted clinical trials in preclinical AD. However, it is likely that designers of preclinical AD treatment trials will have to prepare for larger and longer trials than are currently considered.

Author contributions

P.S. Insel: drafting/ revising the manuscript, study concept or design, analysis or interpretation of data, accepts responsibility for conduct of research and final approval, statistical analysis, study supervision. M. Weiner: drafting/

revising the manuscript, analysis or interpretation of data, accepts responsibility for conduct of research and final approval, study supervision. R.S. Mackin: drafting/ revising the manuscript, analysis or interpretation of data, accepts responsibility for conduct of research and final approval. E. Mormino: drafting/ revising the manuscript, accepts responsibility for conduct of research and final approval. Y.Y. Lim: drafting/ revising the manuscript, analysis or interpretation of data, accepts responsibility for conduct of research and final approval, acquisition of data. E. Stomrud: data acquisition, accepts responsibility for conduct of research and final approval, acquisition of data, study supervision. S. Palmqvist: data acquisition, accepts responsibility for conduct of research and final approval, study supervision, obtaining funding. C.L. Masters: study concept or design, accepts responsibility for conduct of research and final approval, study supervision. P. Maruff: drafting/ revising the manuscript, data acquisition, study concept or design, accepts responsibility for conduct of research and final approval, study supervision, obtaining funding. O. Hansson: drafting/ revising the manuscript, data acquisition, study concept or design, accepts responsibility for conduct of research and final approval, acquisition of data, study supervision, obtaining funding. N. Mattsson: drafting/ revising the manuscript, data acquisition, study concept or design, analysis or interpretation of data,

Figure 4 β -amyloid ($A\beta$) interactions



Interactions between $A\beta$ group and age, education, sex, *APOE*, and baseline memory are shown for each cohort: (A) Alzheimer's Disease Neuroimaging Initiative (ADNI), (B) Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably (BioFINDER), (C) Australian Imaging, Biomarkers & Lifestyle (AIBL). Akaike information criterion (AIC) and p values are shown in each plot, testing for the significance of interactions in predicting Preclinical Alzheimer's Cognitive Composite (PACC) change over time.

accepts responsibility for conduct of research and final approval, acquisition of data, study supervision, study funding.

Study funding

Different funding agencies supported work at Lund University (for overall study coordination, and for the BioFINDER study) and the ADNI and AIBL studies. Work at Lund University in the authors' laboratory is generously supported by The Knut and Alice Wallenberg foundation, the Medical Faculty at Lund University, Region Skåne, the European Research Council, the Swedish Research Council, the Marianne and Marcus Wallenberg foundation, the Strategic Research Area MultiPark (Multidisciplinary Research in Parkinson's disease) at Lund University, the Swedish

Alzheimer Association, the Swedish Brain Foundation, the Skåne University Hospital Foundation, the Bundy Academy, and the Swedish federal government under the ALF agreement. Data collection and sharing for the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, 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 provided funds to support ADNI clinical sites in Canada. Private sector contributions were 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 Southern California. A complete listing of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. Partial financial support of AIBL was provided by the Alzheimer's Association (US), the Alzheimer's Drug Discovery Foundation, an Anonymous foundation, the Science and Industry Endowment Fund, the Dementia Collaborative Research Centres, the McCusker Alzheimer's Research Foundation, the National Health and Medical Research Council (AUS), and the Yulgilbar Foundation, plus numerous commercial interactions supporting data collection. Details of the AIBL consortium can be found at www.AIBL.csiro.au and a list of the researchers of AIBL is provided at <http://aibl.csiro.au/>.

Disclosure

P. Insel reports no disclosures relevant to the manuscript. M. Weiner has served on the Scientific Advisory Boards for Pfizer, BOLT International, Neurotrope Bioscience, Eli Lilly, University of Pennsylvania's Neuroscience of Behavior Initiative, National Brain Research Centre, India, LEARN Program at University of North Carolina, and ADNI. He has provided consulting to Synarc, Pfizer, Janssen, KLJ Associates, Easton Associates, Harvard University, University of California, Los Angeles, Alzheimer's Drug Discovery Foundation, Neurotrope Bioscience, Avid Radiopharmaceuticals, Clearview Healthcare Partners, Perceptive Informatics, Smartfish AS, Decision Resources, Inc., Araclon, Merck, Defined Health, and Genentech. The following entities have provided funding for travel: Pfizer, Paul Sabatier University, MCI Group France, Travel eDreams, Inc., Neuroscience School of Advanced Studies, Danone Trading, BV, CTAD Ant Congres, Kenes, Intl., ADRC, UCLA, UCSD, Sanofi-Aventis Groupe, University Center Hospital, Toulouse, Araclon, AC Immune, Eli Lilly, New York Academy of Sciences, National Brain Research Center, India, for Johns Hopkins Medicine, Consortium for Multiple Sclerosis Centers, Northwestern University, and University of Pennsylvania. He served on the editorial boards for *Alzheimer's & Dementia* and *Magnetic Resonance Imaging*. He received honoraria from Pfizer, Tohoku University, and Danone Trading, BV. He received research support from Merck, Avid, the Veterans Administration,

and Department of Defense. R. Mackin has received research support from The National Institute of Mental Health, Johnson & Johnson, and Avid Radiopharmaceuticals, Inc. E. Mormino has served as a paid consultant for Eli Lilly and Biogen and received funding from K01AG051718. Y. Lim reports serving as a scientific consultant to CogState, Ltd., Biogen, and Lundbeck. E. Stomrud and S. Palmqvist report no disclosures relevant to the manuscript. C. Masters reports serving as an advisor to Prana Biotechnology, Ltd., and a consultant to Eli Lilly and Company. P. Maruff is an employee at Cogstate P/L. O. Hansson has served on advisory boards for Eli Lilly and received research support from GE Healthcare and Hoffmann La-Roche. N. Mattsson has been a consultant for ADNI. Go to Neurology.org/N for full disclosures.

Publication history

Received by *Neurology* November 30, 2018. Accepted in final form March 17, 2019.

References

1. van Dyck CH. Anti-amyloid- β monoclonal antibodies for Alzheimer's disease. *Biol Psychiatry* 2017;83:311–319.
2. Sperling RA, Rentz DM, Johnson KA, et al. The A4 study: stopping AD before symptoms begin? *Sci Transl Med* 2014;6:228fs13.
3. Vos SJB, Xiong C, Visser PJ, et al. Preclinical Alzheimer's disease and its outcome: a longitudinal cohort study. *Lancet Neurol* 2014;12:957–965.
4. Vemuri P, Lesnick TG, Przybelski SA, et al. Vascular and amyloid pathologies are independent predictors of cognitive decline in normal elderly. *Brain* 2015;138:761–771.
5. Mormino EC, Papp KV, Rentz DM, et al. Early and late change on the preclinical Alzheimer's cognitive composite in clinically normal older individuals with elevated amyloid β . *Alzheimers Dement* 2017;13:1004–1012.
6. Wirth M, Oh H, Mormino EC, et al. The effect of amyloid β on cognitive decline is modulated by neural integrity in cognitively normal elderly. *Alzheimers Dement* 2013;9:687–698.
7. Baker JE, Lim YY, Pietrzak RH, et al. Cognitive impairment and decline in cognitively normal older adults with high amyloid- β : A meta-analysis. *Alzheimers Dement* 2017;6:108–121.
8. Hedden T, Oh H, Younger AP, Patel TA. Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology* 2013;80:1341–1348.
9. Insel PS, Hansson O, Mackin RS, et al. Amyloid pathology in the progression to mild cognitive impairment. *Neurobiol Aging* 2018;64:76–84.
10. Donohue MC, Sperling RA, Salmon DP, et al. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol* 2014;71:961–970.
11. Mueller SG, Weiner MW, Thal LJ, et al. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement* 2005;1:55–66.
12. Ellis KA, De Fazio D, Foster J, et al. 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 Psychogeriatr* 2009;21:672–687.
13. Palmqvist S, Zetterberg H, Blennow K, et al. Accuracy of brain amyloid detection in clinical practice using cerebrospinal fluid β -amyloid 42: a cross-validation study against amyloid positron emission tomography. *JAMA Neurol* 2014;71:1282–1289.
14. Aisen PS, Petersen RC, Donohue MC, et al. Clinical core of the Alzheimer's Disease Neuroimaging Initiative: progress and plans. *Alzheimers Dement* 2010;6:239–246.
15. Thomas K, Eppig J, Weigand A, et al. Artificially low mild cognitive impairment to normal reversion rate in Alzheimer's Disease Neuroimaging Initiative. *Alzheimers Dement* 2019;15:1–9.
16. Rowe CC, Ellis KA, Rimajova M, et al. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol Aging* 2010;31:1275–1283.
17. Landau SM, Thomas BA, Thurfjell L, et al. Amyloid PET imaging in Alzheimer's disease: a comparison of three radiotracers. *Eur J Nucl Med Mol Imaging* 2014;41:1398–1407.
18. Jagust WJ, Bandy D, Chen K, et al. The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core. *Alzheimers Dement* 2010;6:221–229.
19. Shaw LM, Vanderstichele H, Knopik-Czajka M, et al. Cerebrospinal fluid biomarker signature in Alzheimer's Disease Neuroimaging Initiative subjects. *Ann Neurol* 2009;65:403–413.
20. Olsson A, Vanderstichele H, Andreassen N, et al. Simultaneous measurement of β -amyloid(1–42), total tau, and phosphorylated tau (Thr181) in cerebrospinal fluid by the xMAP technology. *Clin Chem* 2005;51:336–345.

21. Palmqvist S, Zetterberg H, Mattsson N, et al. Detailed comparison of amyloid PET and CSF biomarkers for identifying early Alzheimer disease. *Neurology* 2015;85:1240–1249.
22. Landau SM, Mintun MA, Joshi AD, et al. Amyloid deposition, hypometabolism, and longitudinal cognitive decline. *Ann Neurol* 2012;72:578–586.
23. Villemagne VL, Doré V, Yates P, et al. En attendant centiloid. *Adv Res* 2014;2:723–729.
24. Janelidze S, Zetterberg H, Mattsson N, et al. CSF A β 42/A β 40 and A β 42/A β 38 ratios: better diagnostic markers of Alzheimer disease. *Ann Clin Transl Neurol* 2016;3:154–165.
25. Donohue MC, Sperling RA, Petersen R, et al. Association between elevated brain amyloid and subsequent cognitive decline among cognitively normal persons. *JAMA* 2017;317:2305–2316.
26. Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-PLUS*. Berlin: Springer; 2000.
27. Hastie TJ, Tibshirani RJ. *Generalized additive models*. *Monogr Stat Appl Probab* 1990;1:352.
28. Akaike H. A new look at the statistical model identification. *Autom Control IEEE Trans* 1974;19:716–723.
29. Mallinckrodt CH, Clark WS, Stacy RD. Type I error rates from mixed-effects model repeated measures versus fixed effects analysis of variance with missing values imputed via last observation carried forward. *Drug Inform J* 2001;35:1215–1225.
30. Sidik K, Jonkman JN. A comparison of heterogeneity variance estimators in combining results of studies. *Stat Med* 2007;26:1964–1981.
31. Schneider LS, Sano M. Current Alzheimer's disease clinical trials: methods and placebo outcomes. *Alzheimers Dement* 2009;5:388–397.
32. Siemers ER, Sundell KL, Carlson C, et al. Phase 3 solanezumab trials: secondary outcomes in mild Alzheimer's disease patients. *Alzheimers Dement* 2016;12:110–120.
33. Honig LS, Vellas B, Woodward M, et al. Trial of solanezumab for mild dementia due to Alzheimer's disease. *N Engl J Med* 2018;378:321–330.
34. Kozauer N, Katz R. Regulatory innovation and drug development for early-stage Alzheimer's disease. *N Engl J Med* 2013;368:1169–1171.
35. Li D, Iddi S, Aisen P, et al. The relative efficiency of time-to-progression and continuous measures of cognition in pre-symptomatic Alzheimer's. 2019;arXiv:1902.02026 [stat.AP].
36. Mormino E, Betensky RA, Hedden T, et al. Amyloid and APOE4 interact to influence short-term decline in preclinical Alzheimer disease. *Neurology* 2014;82:1760–1767.
37. Lim YY, Kalinowski P, Pietrzak RH, et al. Association of β -Amyloid and apolipoprotein E ϵ 4 with memory decline in preclinical Alzheimer disease. *JAMA Neurol* 2018;75:488–494.
38. Insel PS, Donohue MC, Mackin RS, et al. Cognitive and functional changes associated with A β pathology and the progression to mild cognitive impairment. *Neurobiol Aging* 2016;48:172–181.
39. Buckley RF, Mormino EC, Amariglio RE, et al. Sex, amyloid, and APOE ϵ 4 and risk of cognitive decline in preclinical Alzheimer's disease: findings from three well-characterized cohorts. *Alzheimers Dement* 2018;14:1193–1203.
40. Ramirez J, McNeely AA, Scott CJM, et al. White matter hyperintensity burden in elderly cohort studies: the Sunnybrook Dementia Study, Alzheimer's Disease Neuroimaging Initiative, and Three-City Study. *Alzheimers Dement* 2016;12:203–210.
41. Van Westen D, Lindqvist D, Blennow K, et al. Cerebral white matter lesions: associations with A β isoforms and amyloid PET. *Sci Rep* 2016;6:1–9.