ARTICLES
https://doi.org/10.1038/s43587-020-00003-5

nature aging

Individualized prognosis of cognitive decline and dementia in mild cognitive impairment based on plasma biomarker combinations

Nicholas C. Cullen[1,11], Antoine Leuzy[1,11], Sebastian Palmqvist[1,2], Shorena Janelidze[1], Erik Stomrud[1,2], Pedro Pesini[3], Leticia Sarasa[3], José Antonio Allué[3], Nicholas K. Proctor[4], Henrik Zetterberg[5,6,7,8], Jeffrey L. Dage[4], Kaj Blennow[5,6], Niklas Mattsson-Carlgren[1,9,10,12] ✉ and Oskar Hansson[1,2,12] ✉

We developed models for individualized risk prediction of cognitive decline in mild cognitive impairment (MCI) using plasma biomarkers of β-amyloid (Aβ), tau and neurodegeneration. A total of 573 patients with MCI from the Swedish BioFINDER study and the Alzheimer's Disease Neuroimaging Initiative (ADNI) were included in the study. The primary outcomes were longitudinal cognition and conversion to Alzheimer's disease (AD) dementia. A model combining tau phosphorylated at threonine 181 (P-tau181) and neurofilament light (NfL), but not $Aβ_{42}/Aβ_{40}$, had the best prognosis performance of all models (area under the curve = 0.88 for 4-year conversion to AD in BioFINDER, validated in ADNI), was stronger than a basic model of age, sex, education and baseline cognition, and performed similarly to cerebrospinal fluid biomarkers. A publicly available online tool for individualized prognosis in MCI based on our combined plasma biomarker models is introduced. Combination of plasma biomarkers may be of high value to identify individuals with MCI who will progress to AD dementia in clinical trials and in clinical practice.

About 50 million people live with dementia globally, with the prevalence expected to double by 2030 (ref. [1]). Fifty to seventy percent of all dementia cases are caused by Alzheimer's disease (AD)[2]. The dementia stage is preceded by mild cognitive impairment (MCI). Accurate prognosis is important in MCI, since it may either lead to cognitive decline and dementia (due to AD or other diseases) or be benign and stable[3]. If disease-modifying treatments became available for AD[4], accurate prognostics may be important to guide treatment in patients with MCI.

Even at the MCI stage, key pathological hallmarks of AD can be detected in vivo using cerebrospinal fluid (CSF) biomarkers (for example, the ratio of $Aβ_{42}$ to $Aβ_{40}$, and tau phosphorylated at threonine-181 (P-tau181)[5,6]) or positron emission tomography (PET) of Aβ and tau[7,8]. However, the use of these technologies is limited due the perceived invasiveness of lumbar punctures and the high cost and low availability of PET imaging. Blood-based biomarkers could overcome these hurdles.

Blood-based biomarkers of Aβ (A), tau (T) and neurodegeneration (N) in AD (that is, the ATN biomarkers)[9] include the $Aβ_{42}/Aβ_{40}$ ratio[10], P-tau181 (refs. [11–13]) and neurofilament light (NfL)[14,15], respectively. $Aβ_{42}/Aβ_{40}$ and P-tau181 in plasma correlate with Aβ-PET and tau-PET findings, respectively, and can distinguish AD dementia from controls and non-AD neurodegenerative disorders[10–13,16]. Blood-based NfL is associated with cortical atrophy and cognitive decline in AD[17]. P-tau217 in plasma has also recently been described[18]. However, most studies on blood-based AD biomarkers report findings at the group level. There is a gap in our understanding of how well these biomarkers predict clinical outcomes at the individual patient level and how they compare with more basic prediction models. Individualized assessment has recently been applied using CSF and related imaging biomarkers in MCI[19,20]. A similar study is lacking for blood-based biomarkers. It could be of great value for clinical practice and trials to investigate whether plasma ATN biomarkers perform as well as CSF biomarkers and better than more basic prediction models. We have previously performed a study with a multivariate approach to examine plasma biomarkers and the risk of progression from MCI to AD dementia[11], but most other studies focused on evaluating the biomarkers individually[10,12,13]. None of these studies, however, applied the ATN classification system[9], nor did they systematically aim to find the best subset of ATN biomarkers for individualized predictions.

We measured plasma $Aβ_{42}/Aβ_{40}$, P-tau181 and NfL in patients with MCI from two large cohorts and tested which subset of plasma biomarkers best predicted individual risk for cognitive decline and progression to AD dementia. We compared the prognostic ability of plasma biomarkers with the same biomarkers measured in CSF, as well as with a more basic prediction model, and cross-validated our individual-based risk assessment models both within and across cohorts.

Results

Study population characteristics (model selection sample). A total of 148 patients with MCI from the Biomarkers for Identifying Neurodegenerative Disorders Early and Reliably (BioFINDER)

[1]Clinical Memory Research Unit, Department of Clinical Sciences, Lund University, Malmö, Sweden. [2]Memory Clinic, Skåne University Hospital, Lund, Sweden. [3]Araclon Biotech, Zaragoza, Spain. [4]Eli Lilly and Company, Indianapolis, IN, USA. [5]Department of Psychiatry and Neurochemistry, Sahlgrenska Academy, University of Gothenburg, Mölndal, Sweden. [6]Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden. [7]Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK. [8]UK Dementia Research Institute, University College London, London, UK. [9]Department of Neurology, Skåne University Hospital, Lund, Sweden. [10]Wallenberg Centre for Molecular Medicine, Lund University, Lund, Sweden. [11]These authors contributed equally: Nicholas C. Cullen, Antoine Leuzy. [12]These authors jointly supervised this work: Niklas Mattsson-Carlgren, Oskar Hansson. ✉e-mail: niklas.mattsson-carlgren@med.lu.se; oskar.hansson@med.lu.se

Table 1 | Study participant characteristics (model selection)

Characteristic	BioFINDER	ADNI	*P*
n	148	86	
Age (years) (mean (s.d.))	71.36 (5.47)	71.51 (7.59)	0.865
Education (years) (mean (s.d.))	11.18 (3.49)	16.43 (2.67)	<0.001
Sex (*n* (%))			
Male	94 (63.5)	42 (48.8)	0.04
Female	54 (36.5)	44 (51.2)	
MMSE score			
Baseline (mean (s.d.))	27.21 (1.72)	28.26 (1.74)	<0.001
2 years (mean (s.d.))	24.76 (3.66)	28.16 (2.00)	<0.001
4 years (mean (s.d.))	21.78 (5.25)	27.62 (2.92)	<0.001
Conversion to AD dementia			
2 years (*n* (%))			
No	105 (74.5)	80 (93.0)	0.001
Yes	36 (25.5)	6 (7.0)	
4 years (*n* (%))			
No	43 (40.2)	66 (89.2)	<0.001
Yes	64 (59.8)	8 (10.8)	
Conversion to all-cause dementia			
2 years (*n* (%))			
No	96 (64.9)	80 (93.0)	<0.001
Yes	52 (35.1)	6 (7.0)	
4 years (*n* (%))			
No	42 (31.8)	65 (87.8)	<0.001
Yes	90 (68.2)	9 (12.2)	
Plasma biomarkers (pg ml^{-1})			
$A\beta_{42}/A\beta_{40}$ as measured by Elecsys assay (mean (s.d.))	4.15 (0.12)	NA	
$A\beta_{42}/A\beta_{40}$ as measured by mass spectrometry (mean (s.d.))	NA	0.12 (0.01)	
$A\beta_{42}/A\beta_{40}$ status (*n* (%))			
–	64 (43.2)	39 (45.3)	0.86
+	84 (56.8)	47 (54.7)	
P-tau181 as measured by MSD platform (mean (s.d.))	0.91 (0.78)	NA	
P-tau181 as measured by Simoa-based assay (mean (s.d.))	NA	2.64 (0.61)	
P-tau181 status (*n* (%))			
–	50 (33.8)	55 (64.0)	<0.001
+	98 (66.2)	31 (36.0)	
NfL (mean (s.d.))	3.14 (0.45)	3.48 (0.44)	<0.001
NfL status (*n* (%))			
–	86 (58.1)	55 (64.0)	0.458
+	62 (41.9)	31 (36.0)	

Minus and plus signs indicate negative (normal) and positive (abnormal) biomarker values, respectively. Biomarker concentrations were natural log transformed. For the BioFINDER cohort, there were $n=27$ cases of all-cause dementia at 4 years that were not AD dementia ($n=11$ vascular dementia; $n=8$ dementia with Lewy bodies/Parkinson's disease dementia; $n=2$ frontotemporal dementia; $n=6$ non-specified dementia), whereas for the ADNI cohort, there were $n=6$ ($n=1$ delirium due to West Nile encephalitis; $n=1$ dementia with Lewy bodies; $n=1$ dementia due to human immunodeficiency virus; $n=1$ normal pressure hydrocephalus; $n=1$; Down syndrome; $n=1$ non-specified dementia). *P* values were determined by two-sided *t*-test, with significance assumed at $P<0.05$ (no adjustment for multiple comparisons). F, female; M, male; MSD, Meso Scale Discovery; NA, not applicable (due to the use of different assays or time points for the BioFINDER and ADNI datasets; as a result, continuous plasma biomarker data were not compared across groups).

cohort for whom all three plasma and CSF biomarker measurements were available, along with at least one of the primary or secondary outcomes, were used for model selection (Table 1). The mean age was 71.4 years, the mean education duration was 11.2 years, 36.5% were female and the mean mini-mental state examination (MMSE) score was 27.2 ± 1.7 at the baseline. Moreover, the mean MMSE score was 21.8 ± 5.2 at 4 years after the baseline and conversion to AD dementia was 59.8% within 4 years of the baseline. There was a significant negative correlation between plasma $A\beta_{42}/A\beta_{40}$ and plasma P-tau181 (coefficient of determination (R^2) = −0.30; $P<0.0001$; see Extended Data Fig. 1) and a significant positive correlation between plasma P-tau181 and plasma NfL ($R^2=0.33$; $P<0.0001$; Extended

Data Fig. 1), but no significant correlation was observed between plasma $A\beta_{42}/A\beta_{40}$ and plasma NfL ($R^2=-0.08$; $P=0.31$; Extended Data Fig. 1). The associations between corresponding CSF biomarkers in the BioFINDER model selection sample are shown in Extended Data Fig. 2. The study procedures are outlined in Fig. 1.

The data from 86 patients with MCI from the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort for whom all three plasma biomarker measurements were available alongside at least one of the primary or secondary outcomes were used to replicate model selection (Table 1). The mean age was 71.5 years, the mean education duration was 16.4 years, 51.2% were female and the mean MMSE score was 28.3 ± 1.7 at the baseline. Moreover, the mean MMSE score was 27.6 ± 2.9 at 4 years after the baseline and conversion to AD dementia was 10.8% within 4 years of the baseline. There was a significant negative correlation between plasma $A\beta_{42}/A\beta_{40}$ and plasma P-tau181 ($R^2=-0.31$; $P=0.0023$; Extended Data Fig. 2) but no significant correlation between plasma P-tau181 and plasma NfL ($R^2=0.15$; $P=0.16$; Extended Data Fig. 2) or between plasma $A\beta_{42}/A\beta_{40}$ and plasma NfL ($R^2=-0.16$; $P=0.11$; Extended Data Fig. 2).

To validate our prognostic models, we used data from 425 patients with MCI from the ADNI cohort for whom plasma P-tau181 and NfL measurements were available alongside at least one of the primary or secondary outcomes available. The mean age of this cohort was 71.0 years, the mean education duration was 16.1 years, 51.8% were female and the mean MMSE score was 28.2 ± 1.7 at the baseline. Moreover, the mean MMSE was 26.6 ± 4.1 4 years after the baseline and conversion to AD dementia was 33.1% within 4 years of the baseline. There was a significant positive correlation between plasma P-tau181 and plasma NfL ($R^2=0.35$; $P<0.0001$; Extended Data Fig. 3).

Model selection for longitudinal cognition. With 4-year MMSE as the outcome in the BioFINDER model selection sample ($n=118$), the model that included plasma $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL as predictors (full model; $R^2=0.36$; Akaike information criterion (AIC) = 684) fit the data significantly better than the basic model that included age, sex, education and baseline MMSE score ($R^2=0.24$; AIC = 702; $P=0.0001$ compared with the full model). However, the best-fitting model according to the AIC was the one that included plasma P-tau181 and NfL but not $A\beta_{42}/A\beta_{40}$ ($R^2=0.36$; AIC = 683). In the best-fitting model, there was a significant individual effect of P-tau181 ($\beta=-1.65$; $P<0.0001$) but not NfL ($\beta=-0.70$; $P=0.13$) (Fig. 2a,b and Supplementary Table 1).

With 4-year MMSE as the outcome in the ADNI model selection sample used for replication ($n=64$), the full plasma model ($R^2=0.25$; AIC = 310) fit the data better than the basic model ($R^2=0.15$; AIC = 316; $P=0.01$ compared with the full ATN model), and the best-fitting model according to the AIC again included plasma P-tau181 and NfL only ($R^2=0.25$; AIC = 309). In the best-fitting model, the individual effect of P-tau181 was nearly significant ($\beta=-0.64$; $P=0.06$) while the individual effect of NfL was significant ($\beta=-1.02$; $P=0.02$) (Fig. 2a,b and Supplementary Table 1).

Model selection for clinical conversion. With 4-year conversion to AD as the outcome in the BioFINDER model selection sample ($n=107$), the full plasma model that included all three biomarkers (area under the curve (AUC) = 0.88; AIC = 106) fit the data significantly better than the basic model (AUC = 0.70 (0.60–0.79); AIC = 140; $P<0.0001$ compared with the full model). The best-fitting model according to the AIC included P-tau181 and NfL but not $A\beta_{42}/A\beta_{40}$, resulting in an AUC of 0.88 (0.82–0.95). A maximum specificity of 95.3% could be achieved given at least 50% sensitivity, and a maximum sensitivity of 96.8% could be achieved given at least 50% specificity (Fig. 3a,b and Supplementary Table 2). In the best-fitting model, there was a significant individual effect of P-tau181 (odds ratio (OR) = 5.87; $P=0.0001$) but not NfL (OR = 1.73; $P=0.10$). Using Cox regression (Supplementary Table 3), the model with P-tau181 and NfL was still the best-fitting model

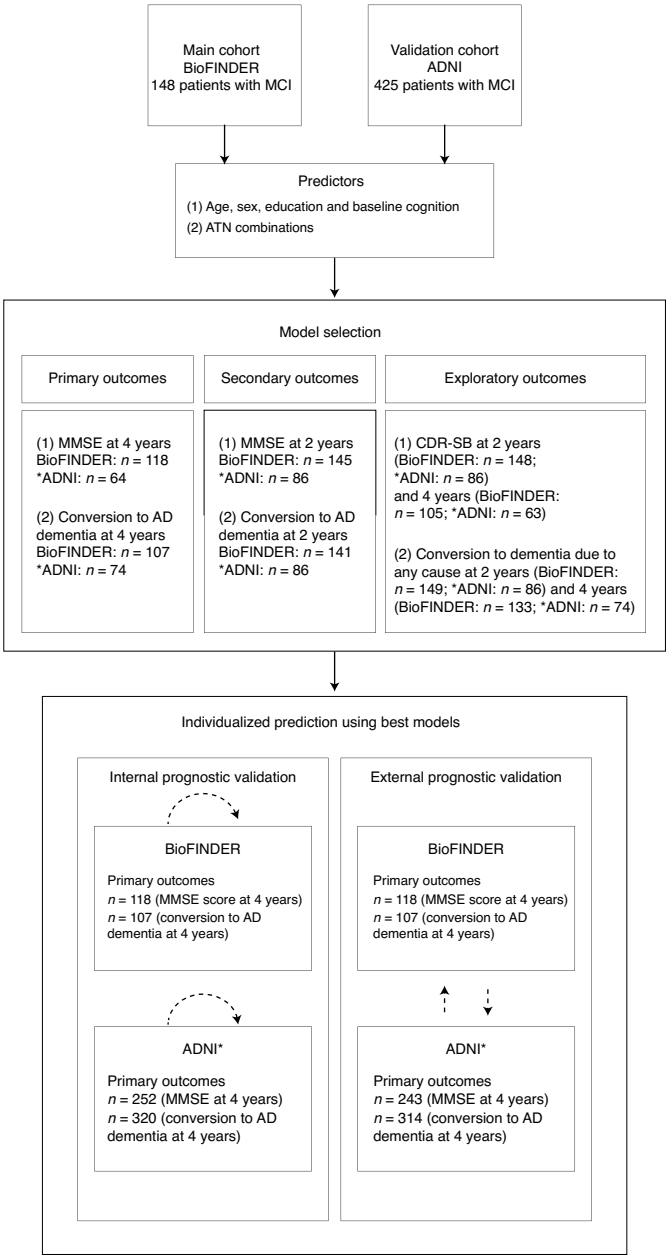


Fig. 1 | Flow chart of study procedures. In the ADNI cohort, 425 participants had data for plasma P-tau181 and NfL. Of these, 86 also had data for plasma $A\beta_{42}/A\beta_{40}$. The number of participants available for different outcomes differed for both BioFINDER and ADNI. Asterisks in the section 'Model selection' represent the ADNI model selection sample, which included individuals in the ADNI cohort for whom data were available for all plasma biomarkers (that is, $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL). In the section 'Individualized prediction using best models', 'internal prognostic validation' refers to fivefold cross-validation for the internal prognostic validation step, whereas 'external prognostic validation' refers to validating the model selection (that is, validating that the best-performing model for BioFINDER was the best-performing model for ADNI, and vice versa). Asterisks in the section 'Individualized prediction using best models' represent the ADNI prognostic validation sample, which included individuals in the ADNI cohort for whom data were available for P-tau181 and NfL.

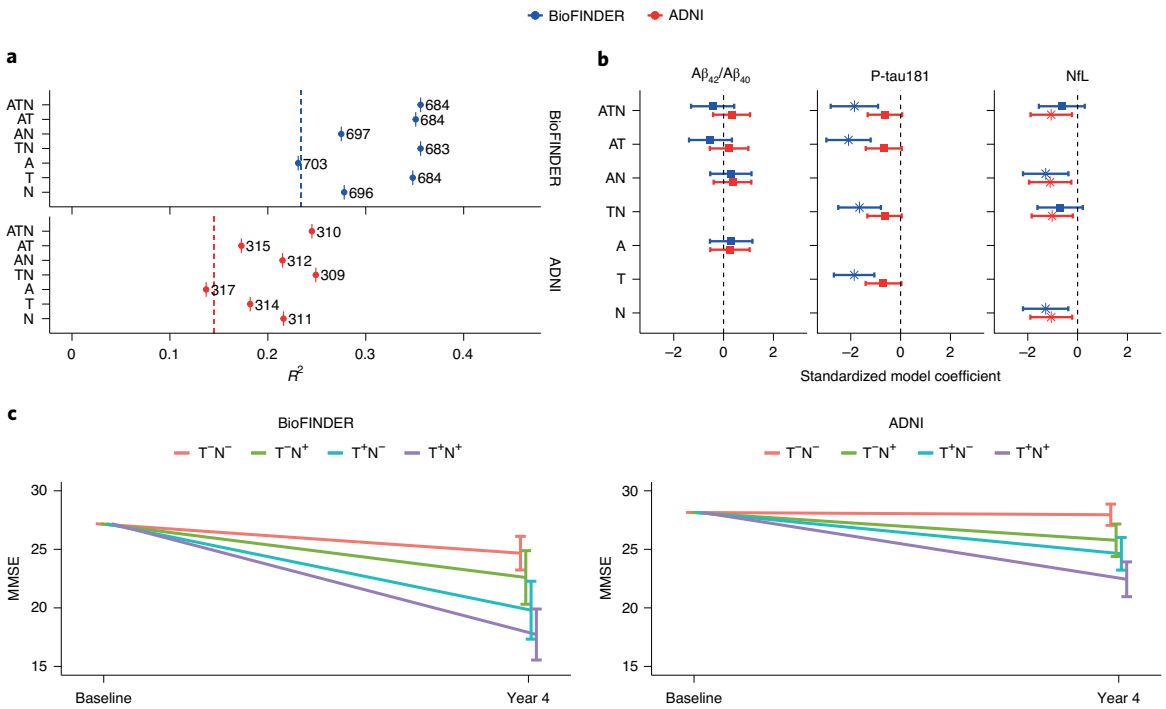


Fig. 2 | Modeling cognitive decline using plasma $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL. Results from modeling cognitive decline in patients with MCI using plasma biomarkers. **a**, R^2 (*x* axis) and AIC values (numbers in plots) for each plasma-based model, with MMSE evaluated 4 years after the baseline as the outcome, in the ADNI and BioFINDER cohorts. The dashed vertical lines show basic model performances for reference. Error bars indicate 95% confidence intervals. All models also included age, sex, education and baseline MMSE as predictors. **b**, Coefficients from each plasma-based model, with MMSE evaluated 4 years after the baseline as the outcome, in the ADNI and BioFINDER cohorts. Statistically significant variables, as determined by two-sided *t*-test on regression coefficients, are plotted with an asterisk instead of a square, and the error bars represent 95% confidence intervals. **c**, Estimated MMSE trajectories (colored lines), together with the estimated trajectories from the best-fitting model (P-tau181 and NfL; $A\beta_{42}/A\beta_{40}$ was not taken forward for assessing predictive performance), according to biomarker status, adjusted for age, sex, education and baseline MMSE. Error bars represent 95% confidence intervals on estimated trajectories. No corrections for multiple comparisons were performed.

and there was a significant individual effect of both P-tau181 (hazard ratio (HR) = 2.52; $P = 0.006$) and NfL (HR = 2.70; $P = 0.02$).

With 4-year conversion to AD as the outcome in the ADNI model selection sample used for replication ($n = 74$), the full plasma model (AUC = 0.88 (0.80–0.98); AIC = 50) fit the data better than the basic model (AUC = 0.74 (0.52–0.95); AIC = 57; $P = 0.005$ compared with the full ATN model) and the best-fitting model according to AIC again included P-tau181 and NfL only. The AUC for this model was 0.89. A maximum specificity of 90.1% could be achieved given at least 50% sensitivity, and a maximum sensitivity of 100% could be achieved given at least 50% specificity (Fig. 3a,b and Supplementary Table 2). In the best-fitting model, the individual effect of P-tau181 was significant (OR = 4.58; $P = 0.009$) while the individual effect of NfL was not significant (OR = 2.15; $P = 0.20$). Using Cox regression (Supplementary Table 3), the model with P-tau181 and NfL was again still the best-fitting model, with a significant effect in this model for P-tau181 (HR = 2.22; $P < 0.0001$) but not NfL (HR = 1.27; $P = 0.13$). Receiver operating characteristic (ROC) curves for the main models are presented in Fig. 4. ROC curves for all of the models in the model selection analysis with 4-year conversion to AD dementia as the outcome are presented in Extended Data Fig. 4.

Effect of the *APOE* ε4 genotype on model selection. Including *APOE* ε4 genotype status (that is, ε4 carrier versus non-carrier) in the basic model did not substantially affect model selection for the primary cognitive outcome (Supplementary Table 4) or for the primary clinical outcome (Supplementary Table 5). Moreover, the full plasma ATN model still outperformed the basic model for both the primary cognitive and clinical outcomes.

Sensitivity analysis using secondary and exploratory outcomes. We performed the same model selection procedure as outlined above using the secondary and exploratory outcomes. We found that the best-fitting models identified here varied across outcomes, whereas the full plasma model including $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL still always outperformed the basic model alone (Supplementary Tables 4–13).

Sensitivity analysis using an alternative plasma $A\beta_{42}/A\beta_{40}$ assay. Because plasma $A\beta_{42}/A\beta_{40}$ was not selected as a predictor in the best-fitting models for the co-primary outcomes, we tested whether this result differed when using a mass spectrometry assay (Araclon Biotech) instead of the Elecsys assay used in the BioFINDER cohort. Here, the best-fitting models according to AIC when using the mass spectrometry plasma $A\beta_{42}/A\beta_{40}$ assay still did not include $A\beta_{42}/A\beta_{40}$. Instead, the best-fitting model with 4-year MMSE as the outcome included P-tau181 and NfL only and the best-fitting model with 4-year conversion to AD as the outcome included P-tau181 only (Supplementary Tables 12 and 13).

Sensitivity analysis using plasma P-tau217 in the BioFINDER cohort. Primary analyses (that is, 4-year MMSE and 4-year conversion to AD) were repeated using plasma P-tau217 (see Palmqvist et al.[18] for an assay description) instead of P-tau181. As plasma P-tau217 data are not available for ADNI, these analyses were performed

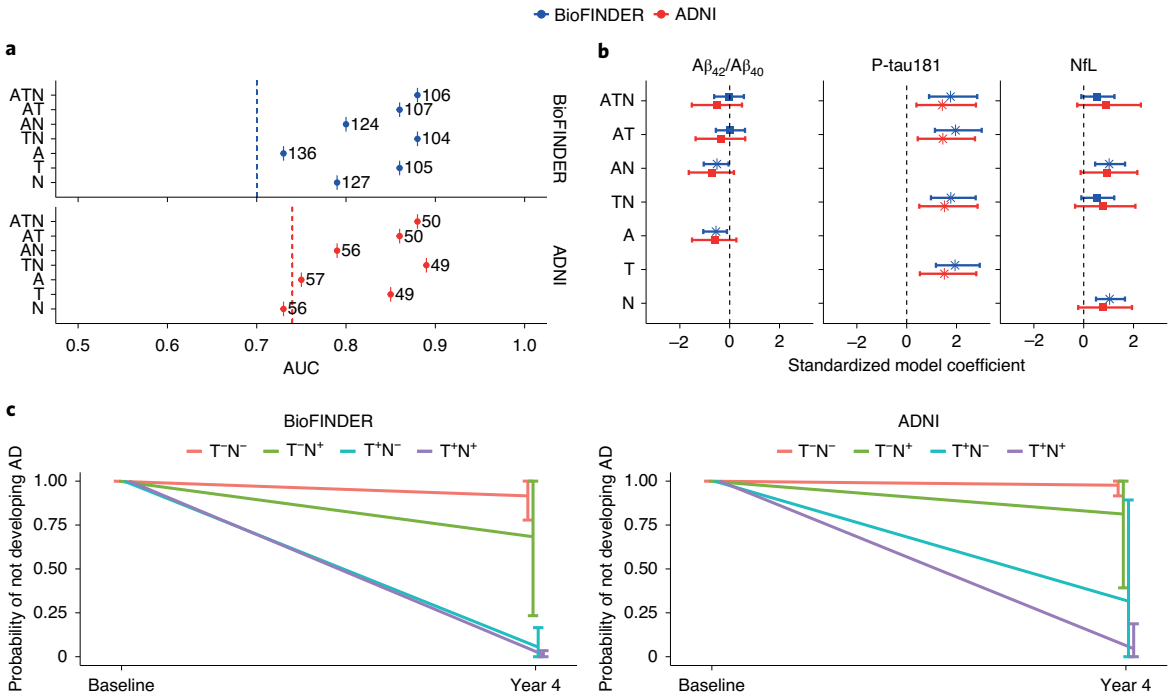


Fig. 3 | Modeling clinical conversion using plasma $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL. Results from modeling clinical conversion in patients with MCI using plasma biomarkers. **a**, AUC (*x* axis) and AIC values (numbers in plots) for each plasma-based model, with conversion to AD within 4 years after baseline as the outcome, in the BioFINDER and ADNI cohorts. The dashed vertical lines show basic model performances for reference. Error bars indicate 95% confidence intervals. All models also included age, sex, education and baseline MMSE as predictors. **b**, Coefficients from each plasma-based model, with conversion to AD within 4 years after the baseline as the outcome, in the BioFINDER and ADNI cohorts. Statistically significant variables, as determined by two-sided *t*-test on regression coefficients, are plotted with an asterisk instead of a square, and the error bars represent 95% confidence intervals. **c**, Estimated probability of not converting to AD, as predicted from the best-fitting model (P-tau181 and NfL; $A\beta_{42}/A\beta_{40}$ was not taken forward for assessing predictive performance), according to biomarker status, adjusted for age, sex, education and baseline MMSE. Error bars represent 95% confidence intervals on estimated trajectories.

using BioFINDER participants only. Findings from these analyses replicated those found when using P-tau181 (Supplementary Tables 14 and 15): P-tau in combination with NfL remained the best model for both outcomes, and R^2 and AUC values were similar when using P-tau217 or P-tau181 (for MMSE, compare Supplementary Table 14 with Supplementary Table 1; for conversion to AD, compare Supplementary Table 15 with Supplementary Table 2).

Study population characteristics (prognostic validation sample). Since the model that included both plasma P-tau181 and NfL (but not $A\beta_{42}/A\beta_{40}$) provided the best fit across co-primary outcomes in both cohorts, this model was taken forward in the prognostic validation stage. For this analysis, we therefore only required participants to have available data on plasma P-tau181 and NfL measurements and at least one of the primary or secondary outcomes. As such, 148 patients with MCI from the BioFINDER cohort (no difference from the BioFINDER model selection sample) and 425 patients with MCI from the ADNI cohort (Table 2; see Extended Data Fig. 5 for the association between plasma P-tau181 and NfL in this group) were included.

Patient-level prognostic validation within cohorts. We performed an internal validation where the patient-level predictive performance of the best-fitting plasma model (the basic model plus plasma P-tau181 and NfL) was evaluated within each cohort and compared with the basic model alone and with the full CSF model (the basic model plus CSF $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL).

With 4-year MMSE as the outcome in the BioFINDER prognostic validation sample ($n=118$), the best-fitting plasma model improved cross-validated, out-of-sample prediction compared with the basic model alone (mean absolute error (MAE) = 3.07 points versus 3.36 points; $P<0.001$; 8.5% improvement) and showed no significant difference compared with the full CSF model ($P=0.68$ over 1,000 bootstrapped trials). With 4-year MMSE as the outcome within the ADNI prognostic validation sample ($n=252$), the same best-fitting plasma model improved out-of-sample prediction compared with the basic model alone (MAE = 2.42 points versus 2.49 points; $P<0.001$; 2.9% improvement).

With 4-year conversion to AD as the outcome in the BioFINDER prognostic validation sample ($n=107$), the best-fitting plasma model improved out-of-sample prediction compared with the basic model alone (AUC = 0.63 versus 0.83 (note that these cross-validated AUCs are, as expected, lower than the AUCs from corresponding models fit on all data for model selection); $P<0.001$; 31.7% improvement) and significantly outperformed the full CSF model ($P=0.002$ over 1,000 bootstrapped trials; 5% improvement). With 4-year conversion to AD as the outcome in the ADNI prognostic validation sample ($n=320$), the plasma model significantly improved out-of-sample prediction compared with the basic model alone (AUC = 0.66 versus 0.76; $P<0.001$; 15.4% improvement). For a visual depiction of individualized predicted probabilities of conversion to AD across models, see Fig. 5.

Patient-level prognostic validation across cohorts. We performed an external validation whereby the patient-level predictive performance of the best-fitting plasma model (the basic model plus plasma P-tau181 and NfL) was evaluated across each cohort by first fitting the model on BioFINDER participants and then testing on

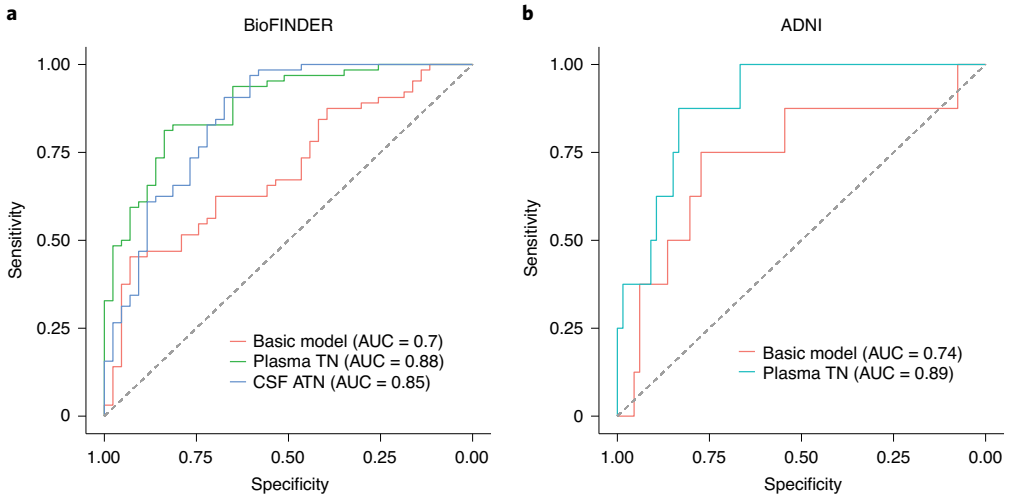


Fig. 4 | ROC curves for clinical conversion across cohorts. a,**b**, ROC curves from modeling clinical conversion in patients with MCI using plasma biomarkers in the BioFINDER (**a**) and ADNI (**b**) cohorts. The basic model (age, sex, education and baseline MMSE) is shown along with the best-performing model selected for further prognostic validation (basic model + plasma P-tau181 + plasma NfL (plasma TN)). Results from the full CSF model (basic model + CSF $A\beta_{42}/A\beta_{40}$ + CSF P-tau181 + CSF NfL (CSF ATN)) are also shown for the BioFINDER cohort. These results are not shown for the ADNI cohort as the full CSF model was not available (but note that the CSF TN model for ADNI was tested at the prognostic validation stage).

ADNI participants, and vice versa. For this analysis, biomarkers were dichotomized.

With 4-year MMSE as the outcome in the BioFINDER ($n = 118$, of whom 28 were T^-N^-, 13 were T^-N^+, 46 were T^+N^- and 31 were T^+N^+) and ADNI prognostic validation samples ($n = 243$, of whom 118 were T^-N^-, 35 were T^-N^+, 46 were T^+N^- and 44 were T^+N^+), the plasma model significantly improved prediction on the test cohort compared with the basic model, both when the model was fit on BioFINDER and tested on ADNI (MAE = 3.74 versus 4.08; $P = 0.0006$; 8.3% improvement) and when the model was fit on ADNI and tested on BioFINDER (MAE = 4.15 versus 5.19; $P < 0.0001$; 20.1% improvement).

With 4-year conversion to AD as the outcome in the BioFINDER ($n = 107$; of whom 20 were T^-N^-, five were T^-N^+, 49 were T^+N^- and 33 were T^+N^+) and ADNI prognostic validation samples ($n = 314$, of whom 139 were T^-N^-, 45 were T^-N^+, 62 were T^+N^- and 68 were T^+N^+), the plasma model improved prediction on the unseen cohort both when the model was fit on ADNI and tested on BioFINDER (AUC = 0.61 versus 0.79; $P < 0.0001$; 29.3% improvement) and when the model was fit on BioFINDER and tested on ADNI (AUC = 0.62 versus 0.73; $P < 0.0001$; 18.3% improvement).

Online individualized risk prediction tool. We provide an illustrative online tool for use with the current dataset at predictprogression.com, where individualized predictions can be made for MMSE, conversion to AD dementia and clinical dementia rating scale–sum of boxes (CDR-SB) score, at 2 and 4 years after the baseline, in patients with MCI at the baseline. The tool allows the user to enter data on age, sex, baseline cognition (MMSE and/or CDR-SB) and dichotomous biomarker status for CSF or plasma $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL. It is also possible to test predictions with sparse models including subsets of biomarkers. For example, for a 70-year-old female with MCI and a baseline MMSE score of 27, if all plasma $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL values are known and negative, the probabilities are 6% (90% prediction interval = 2–20%) at 2 years and 16% (90% prediction interval = 5–38%) at 4 years (Extended Data Fig. 6). If all plasma $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL values are positive, the probabilities change to 43% (90% prediction interval = 25–62%) at 2 years and 92% (90% prediction interval = 77–97%) at 4 years (Extended Data Fig. 6). In the absence of biomarker information, the predicted probability of conversion to AD is 33% (90% prediction interval = 23–45%) at 2 years and 69% (90% prediction interval = 56–80%) at 4 years.

Discussion

We addressed the patient-level prognostic value of plasma AD biomarkers ($A\beta_{42}/A\beta_{40}$, P-tau181 and NfL) in MCI. Plasma P-tau181 in combination with NfL best predicted primary outcomes of decline in MMSE score and clinical progression to AD dementia over 4 years. These results were robust to the time horizon (2 or 4 years of follow-up), selection of outcome (MMSE, CDR-SB, conversion to AD dementia or conversion to all-cause dementia), two different cohorts and choice of Aβ assay. In general, prognostic performance using the plasma-based models was either non-inferior or even better than when using CSF biomarkers, and better than a basic model including age, sex, education and baseline cognition. These biomarker-driven prediction models can be applied using our online tool for accurate individualized prognosis in MCI (that is, to predict, for a given patient, both what their MMSE score will be 4 years after their baseline visit and their percentage risk of progressing from MCI to AD dementia over the same time interval). This tool might improve treatment and care[21] and could increase power for clinical trials for prodromal AD by only including those with a high risk of future progression.

Our study is novel in the way we address the individualized predictive value of plasma AD biomarkers, but it can be compared with previous work examining CSF and imaging biomarker-driven prognosis at the MCI stage[20]. Using four separate prognostic models—including age, sex, CSF $A\beta_{42}$, T-tau and MMSE, as well as an ATN variant combining CSF $A\beta_{42}$ and P-tau181 with hippocampal volume—van Maurik and colleagues[20] looked at the likelihood of progression to dementia from MCI. While all models performed well, the highest performance was seen using the CSF ATN model. Similarly, we found that a combined ATN model ($A\beta_{42}/A\beta_{40}$, P-tau181 and NfL) in plasma outperformed a basic model with demographics and baseline MMSE score as predictors. Importantly, plasma models improved the prediction of longitudinal MMSE despite adjustment for MMSE score at the baseline, which in itself is a very strong predictor of future MMSE score.

Table 2 | Study participant characteristics in ADNI (prognostic validation)

Characteristic	ADNI
n	425
Age (years) (mean (s.d.))	70.98 (7.81)
Education (years) (mean (s.d.))	16.06 (2.65)
Sex (*n* (%))	
Male	205 (48.2)
Female	220 (51.8)
MMSE score	
Baseline (mean (s.d.))	28.15 (1.67)
2 years (mean (s.d.))	27.27 (2.76)
4 years (mean (s.d.))	26.57 (4.10)
Conversion to AD dementia	
2 years (*n* (%))	
No	337 (79.7)
Yes	86 (20.3)
4 years (*n* (%))	
No	230 (66.9)
Yes	114 (33.1)
Conversion to all-cause dementia	
2 years (*n* (%))	
No	336 (79.1)
Yes	89 (20.9)
4 years (*n* (%))	
No	228 (65.7)
Yes	119 (34.3)
Plasma biomarkers (pg ml^{-1})	
$A\beta_{42}/A\beta_{40}$ as measured by mass spectrometry (mean (s.d.))	0.12 (0.01)
$A\beta_{42}/A\beta_{40}$ status (*n* (%))	
–	39 (45.3)
+	47 (54.7)
P-tau181 as measured by Simoa-based assay (mean (s.d.))	2.70 (0.61)
P-tau181 status (*n* (%))	
–	258 (60.7)
+	167 (39.3)
NfL (mean (s.d.))	3.50 (0.47)
NfL status (*n* (%))	
–	283 (66.6)
+	142 (33.4)

This table includes data for a subset of the ADNI cohort with plasma P-tau181 and NfL measurements and at least one of the primary or secondary outcomes available (prognostic validation). Minus and plus signs indicate negative (normal) and positive (abnormal) biomarker values, respectively. Biomarker concentrations were natural log transformed.

Inclusion of P-tau181 in the best models may reflect that P-tau181 detects AD-type changes[11]. In contrast, plasma $A\beta_{42}/A\beta_{40}$ was not included in the best model, suggesting that plasma Aβ biomarkers do not provide additional prognostic information in MCI when an efficient plasma P-tau measure is included. This is logical, since symptoms in AD are linked to tau pathology[22], and elevations in tau biomarkers appear to be dependent on Aβ pathology[23,24]. Findings for plasma $A\beta_{42}/A\beta_{40}$ have also been more varied than for plasma P-tau181 (refs. [10,25]) and have only shown modest reductions in AD dementia (10–15% compared with cognitively unimpaired), while P-tau181 is greatly increased in AD dementia (>100% compared with cognitively unimpaired)[11–13]. However, it is possible that plasma $A\beta_{42}/A\beta_{40}$ may have added value at the preclinical stage of the disease, when it has reached pathological levels[10], while tau and neurodegeneration markers continue to increase during the symptomatic stages of the disease[11,26]. The best models also included NfL, which is a more general marker of neurodegeneration[17] and appears to give prognostic information complimentary to P-tau181.

In addition to the CSF studies by van Maurik on individualized biomarker-based risk predictions of dementia in patients with MCI[19,20], recent work from our group has examined the association between plasma-based biomarkers and the risk of AD dementia[11]. Although similar in terms of including plasma $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL, the present study differs from our previous work in a number of important ways. We now focus on identifying optimal models within the ATN framework, rather than individual biomarkers, and on patient-level (and not group-level) predictions. We also compare plasma-based models with a more basic model without biomarkers (but with baseline MMSE score) and with CSF-based models. We performed extensive internal and external validation analyses (including novel plasma P-tau181 measurements from the ADNI cohort). In terms of results, one important difference compared with the previous study[11] is the finding that both NfL and P-tau181 (rather than just P-tau181) contribute to the best-performing models. Although little difference in AIC was seen between plasma models with and without $A\beta_{42}/A\beta_{40}$, we opted not to include this biomarker as we aimed to select the most parsimonious model (that is, the one performing best or as good as all the others but including the least number of biomarkers).

Although the relative importance of biomarkers may vary across contexts and intended applications, plasma biomarkers are promising due to their high accessibility and low cost. With respect to the potential future clinical implementation of plasma biomarkers, the use of binarized data (that is, abnormal versus normal) will probably prove much easier to implement as different assays will have different cut-offs; moreover, different cut-offs may even be required across different laboratories, even when using the same assay, due to local differences in how plasma is collected, handled and analyzed. That said, in the longer term, assays will need to be standardized so that the same values are obtained even when samples are measured in different laboratories; this represents a considerable task, based on the field's experience with CSF AD biomarkers[27].

The strengths of our study include the use of CSF-based ATN models as an internal performance benchmark and the focus on risk estimates at the participant level. The results were robust across different assays used to measure plasma P-tau181 in the BioFINDER and ADNI cohorts. Validation in two independent cohorts with greatly differing demographic makeup speaks to the robustness and relevance of our findings. BioFINDER patients were recruited in a consecutive fashion at three different memory clinics, with approximately 90% of these referred by primary care physicians. ADNI patients were recruited from many different clinics and may be more representative of a highly selected clinical trial population. Interestingly, the additional effect that including plasma biomarkers had on individualized prognostic performance was higher in the BioFINDER cohort than in the ADNI cohort. This phenomenon is most likely explained by the differences in the demographic makeup of the cohorts, as previously explained. Sensitivity analyses showed that our results were robust across clinical outcomes and for the method used to measure plasma $A\beta_{42}/A\beta_{40}$.

We showed that including the *APOE* ε4 genotype in the basic model did not dampen the effect of including plasma biomarkers. This is a promising result given the ethical difficulties in disclosing

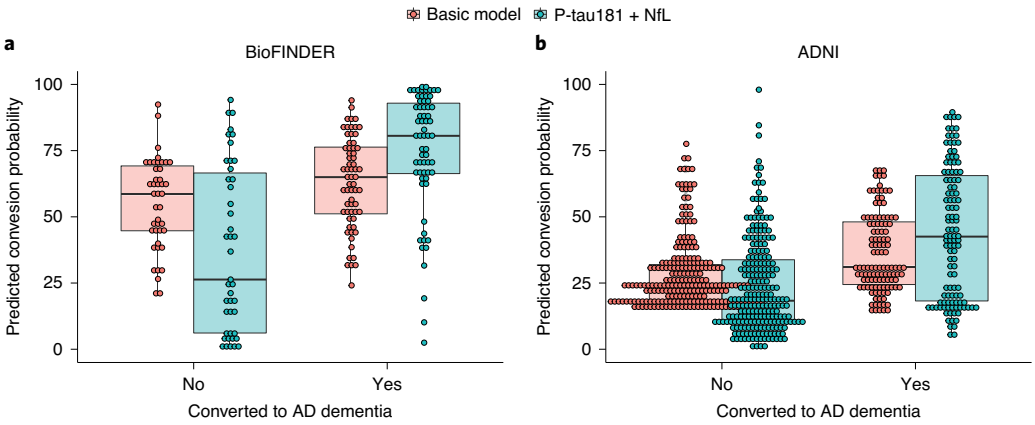


Fig. 5 | Individualized prediction of 4-year conversion from MCI to AD dementia. a,**b**, Results from internal cross-validation for clinical conversion for the best-performing models, as identified in the first stage of analysis using all available BioFINDER (**a**; $n=107$) and ADNI (**b**; $n=320$) patients. The box plots show the minimum and maximum values (whiskers), as well as the median (central line) and interquartile range (that is, the first quartile (25th percentile; bottom edge of box) and third quartile (75th percentile; top edge of box)). The values plotted here show the predicted probability of conversion from MCI to AD dementia for each individual in the BioFINDER and ADNI cohorts, showing 32.3% improvement (AUC = 0.62 for the basic model versus AUC = 0.82 for the P-tau181 and NfL model) of the plasma-based model over the basic model in BioFINDER and a 15.4% improvement (AUC = 0.66 for the basic model versus AUC = 0.76 for the P-tau181 and NfL model) in ADNI.

genetic status to patients, particularly as it relates to patients' relatives. Future work will be focused on understanding whether *APOE* ε4 allele carriers can especially benefit from plasma biomarker measurement or whether current plasma ATN biomarkers can largely replace the need for genetic testing, as indicated by the present results.

This study has limitations. First, our sample size was relatively modest; therefore, significant differences between models are difficult to establish. We relied on selecting the most parsimonious model where differences in AIC values were small (when the models perform similarly, we believe that a less complex model with fewer biomarker predictors is preferable). Further studies on larger and more diverse populations, including in primary care, may result in more precise and generalizable models. Second, as data on plasma P-tau217 were not available from the ADNI cohort, the sensitivity analysis for primary outcomes replacing P-tau181 with P-tau217 was only performed for the BioFINDER cohort. Although plasma P-tau217 has recently been shown to outperform P-tau181 for discriminating AD dementia from non-AD conditions[18], there are, as yet, no data indicating that P-tau217 is superior to P-tau181 for predicting progression to AD in MCI. As our sensitivity analysis resulted in findings that were quite similar to those using P-tau181, P-tau217 might not be a clearly better predictor of future cognitive decline than P-tau181 in individuals with MCI. Third, the finding that the plasma-based model outperformed the CSF-based model when using 4-year conversion to AD—but not when using 4-year MMSE score—as the outcome requires replication. Possibly, this finding could relate to our sample size or to changes to the blood–brain barrier, rendering plasma P-tau181 more specific for AD pathology than CSF P-tau181. It is also possible that variability in CSF production and turnover could differentially impact the performance of CSF P-tau181 versus plasma P-tau181. This too requires further study.

Lastly, as indicated above, our online prediction tool is specific to the dataset we examined and was included only to illustrate the possible future use of such a tool in clinical settings. Further studies are needed to address its performance using data from other cohorts.

In conclusion, plasma-based AD biomarkers can provide patient-level prognostic information in MCI, comparable to CSF biomarkers. Plasma P-tau181 in combination with NfL seems to best predict cognitive decline and clinical progression. Plasma biomarkers of core AD features may aid in individualized risk assessment for patients with MCI, which represents a critical step towards accessible precision medicine for cognitive diseases. Standardized assays with universal cut-offs and replication of the findings in large cohorts are needed.

Methods

This study was conducted in accordance with Standards for Reporting of Diagnostic Accuracy (STARD) guidelines. The STARD checklist can be found as part of the Supplementary Information.

Participants. In this prospective study, we studied consecutively enrolled patients with MCI from the Swedish BioFINDER cohort (clinical trial number NCT01208675; www.biofinder.se). The patients were recruited and evaluated at memory clinics in the cities of Lund, Malmö and Ängelholm between July 2008 and June 2019. They were between 60 and 80 years old and fulfilled the consensus criteria for MCI suggested by Petersen[28] (including: cognitive complaints, preferably corroborated by an informant; objective cognitive impairment, adjusted for age and education; preservation of general cognitive functioning and an MMSE score of 24–30; and no or minimal impairment of daily life activities) but did not fulfill criteria for dementia, as described previously in detail[29]. Exclusion criteria included cognitive impairment that could better be accounted for by another non-neurodegenerative condition, severe somatic disease and current alcohol or substance abuse. After their baseline visit, all patients were seen annually, to assess clinical progression.

For validation, data were obtained from patients with MCI in the ADNI database (the convenience sample; clinical trial number NCT00106899; adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership led by principal investigator M. W. Weiner. For up-to-date information, see www.adni-info.org. The data included in the present study were collected between September 2005 and December 2019.

All participants gave written informed consent. For BioFINDER, ethical approval was given by the Regional Ethical Committee of Lund University. Ethical approval in relation to the ADNI cohort was given by the local ethical committees of all of the involved sites (see Supplementary Notes for a complete list). For both cohorts, no data points were excluded from the analyses. As both BioFINDER and ADNI are observational cohort studies, allocation of participants into experimental groups was not performed (that is, there was no randomization). All analyses were adjusted for age, sex, education and baseline cognition.

Outcomes. The co-primary outcomes were the global cognitive measure MMSE and clinical conversion to AD dementia, evaluated 4 years after the baseline (4-year MMSE and 4-year conversion to AD, respectively). Clinicians who evaluated cognitive decline and conversion to dementia were blinded to biomarker data. As secondary outcomes, we used 2-year MMSE and 2-year conversion to AD dementia. As exploratory outcomes, we used 2-year and 4-year CDR-SB along with 2-year and 4-year conversion to dementia due to any cause. Outcomes were selected on the basis of their clinical relevance.

In BioFINDER, clinical status of dementia due to AD or other diseases was evaluated according to *Diagnostic and Statistical Manual of Mental Disorders* (version 5) criteria for major neurocognitive disorder (that is, dementia). Dementia in ADNI was defined using the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association criteria for probable AD[30].

Predictors. All models included age, sex, education and baseline MMSE score as predictors (the basic model). We also measured $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL in both CSF and plasma. These analyses were performed by technicians blinded to the clinical data.

For the BioFINDER cohort, plasma $A\beta_{42}/A\beta_{40}$ was measured using Elecsys immunoassays on a Cobas e601 analyzer (Roche Diagnostics)[10]. A sensitivity analysis was performed using a mass spectrometry-based plasma $A\beta_{42}/A\beta_{40}$ assay (Araclon Biotech). Plasma P-tau181 was measured on a Meso Scale Discovery platform using an assay developed by Eli Lilly[11]. Plasma NfL was analyzed using a single molecule array (Simoa)-based assay[17]. Moreover, CSF levels of $A\beta_{42}$ (used in place of $A\beta_{42}/A\beta_{40}$ due to no available $A\beta_{40}$ data for the ADNI cohort) and P-tau181 were measured using Elecsys assays (Roche Diagnostics), while CSF NfL was measured using an enzyme-linked immunosorbent assay (UmanDiagnostics).

For the ADNI cohort, plasma $A\beta_{42}/A\beta_{40}$ was analyzed using an immunoprecipitation and mass spectrometry-based method[31]. P-tau181 was analyzed on a Simoa HD-X Analyzer (Quanterix) using an assay developed in the Clinical Neurochemistry Laboratory, University of Gothenburg, Sweden[12]. Plasma NfL was analyzed using the same Simoa-based assay as was used for the BioFINDER cohort.

All biomarker values were natural log transformed. Biomarkers were binarized when validating models across cohorts, whereby cut-offs were defined using Youden's index to maximize the separation between Aβ-negative cognitively unimpaired participants and Aβ-positive patients with AD dementia from within each cohort; these participants have been described previously[11,12]. We derived cut-offs separately for plasma and CSF biomarkers in the BioFINDER cohort, and for plasma biomarkers only in the ADNI cohort (see Supplementary Table 16). For the BioFINDER cohort, there were 528 cognitively unimpaired participants (age = 72.5 years (s.d. = 5.5 years); education = 12.3 years (s.d = 3.6 years); sex = 57% female) and 81 patients with AD (age = 76.7 years (s.d = 5.1 years); education = 9.7 years (s.d = 3.0 years); sex = 59% female). For the ADNI cohort, there were 126 cognitively unimpaired participants (age = 71.5 years (s.d = 6.5 years); education = 16.1 years (s.d. = 2.9 years); sex = 41% female) and 106 patients with AD (age = 73.8 years (s.d = 8.3 years); education = 15.7 years (s.d = 2.7 years); sex = 58% female). Note that none of the participants used to define cut-offs were used in the statistical analysis.

Statistical analysis. In the first analysis stage (model selection), different linear regression models were fit with the cognitive outcomes described above as response variables: a basic model (age, sex, education and baseline MMSE) and plasma biomarker models (the basic model plus seven different biomarker combinations including: $A\beta_{42}/A\beta_{40}$ only; P-tau181 only; NfL only; $A\beta_{42}/A\beta_{40}$ and P-tau181; $A\beta_{42}/A\beta_{40}$ and NfL; P-tau181 and NfL; or all three biomarkers). Because *APOE* ε4 is the strongest genetic risk factor for AD, we tested whether the addition of *APOE* ε4 genotype status (represented as a binary variable split based on individuals with at least one ε4 allele) to the basic model reduced the effectiveness of using plasma biomarkers. Models were compared using R^2 and AIC values (lower is better). The best-fitting model was that which included the fewest predictors among the models within two points of the lowest AIC value; this procedure is well established for selecting the most parsimonious model based on AIC values[32,33]. The statistical significance of different models with the same outcome variable was assessed using the likelihood ratio test. Additionally, logistic regression models were fit with the clinical conversion outcomes described above as response variables, with the same set of predictors and the same method of comparison but with AUC instead of R^2 as the performance metric; a sensitivity analysis was performed using Cox regression models, to ensure that the timing of conversion to AD did not affect model selection.

In the second analysis stage (prognostic validation), the best-fitting model identified in the first stage was carried forward and its predictive accuracy was evaluated. Prognostic validation was first done separately within each cohort using 1,000 repetitions of fivefold cross-validation (internal validation), and then by fitting the model on BioFINDER participants and testing on ADNI participants, and vice versa (external validation; all biomarkers were dichotomized for this analysis, to compare across assays). For internal validation in the BioFINDER cohort, the best-fitting plasma model was compared with a CSF model that included CSF $A\beta_{42}/A\beta_{40}$, P-tau181 and NfL. Prognostic performance was evaluated using MAE with cognitive outcomes, which represents the absolute deviation in MMSE score predicted by the model (for example, MAE = 3 means that the model predicts an individual's cognitive value to within three points, on average) and using AUC with clinical conversion outcomes.

In the model selection stage, only participants with all three plasma biomarkers available were included. In the prognostic validation stage, participants were only required to have measurements from the plasma biomarkers included in the best-fitting model. Both Q–Q plots and normality of residuals were visually inspected for primary (basic and full ATN) regression models. No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications[11,19]. All analyses were performed using the R programming language (version 4.0.0), with significance set at $P < 0.05$ (two sided).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Plasma and CSF data from ADNI were downloaded from https://ida.loni.usc.edu. Anonymized data from the BioFINDER study will be shared by request from a qualified academic investigator for the sole purpose of replicating procedures and results presented in this Article and as long as the data transfer is in agreement with European Union legislation on general data protection regulation and decisions by the Ethical Review Board of Sweden and Region Skåne, which should be regulated by a material transfer agreement.

Code availability

The code that support the findings of this study is available from the corresponding author upon request. All models were built using publicly available packages and functions in the R programming language.

Received: 10 July 2020; Accepted: 30 October 2020;
Published online: 30 November 2020

References

1. *World Alzheimer Report 2019* (Alzheimer's Disease International, 2019).
2. Winblad, B. et al. Defeating Alzheimer's disease and other dementias: a priority for European science and society. *Lancet Neurol.* **15**, 455–532 (2016).
3. Burnham, S. C. et al. The dawn of robust individualised risk models for dementia. *Lancet Neurol.* **18**, 985–987 (2019).
4. Abbasi, J. Promising results in 18-month analysis of Alzheimer drug candidate. *J. Am. Med. Assoc.* **320**, 965 (2018).
5. Buchhave, P. et al. Cerebrospinal fluid levels of β-amyloid 1–42, but not of tau, are fully changed already 5 to 10 years before the onset of Alzheimer dementia. *Arch. Gen. Psychiatry* **69**, 98–106 (2012).
6. Hansson, O. et al. CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: a study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimers Dement.* **14**, 1470–1481 (2018).
7. Rabinovici, G. D. et al. Association of amyloid positron emission tomography with subsequent change in clinical management among Medicare beneficiaries with mild cognitive impairment or dementia. *J. Am. Med. Assoc.* **321**, 1286–1294 (2019).
8. Ossenkoppele, R. et al. Discriminative accuracy of [^{18}F]flortaucipir positron emission tomography for Alzheimer disease vs other neurodegenerative disorders. *J. Am. Med. Assoc.* **320**, 1151–1162 (2018).
9. Jack, C. R. Jr et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement.* **14**, 535–562 (2018).
10. Palmqvist, S. et al. Performance of fully automated plasma assays as screening tests for Alzheimer disease-related β-amyloid status. *JAMA Neurol.* **76**, 1060–1069 (2019).
11. Janelidze, S. et al. Plasma P-tau181 in Alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to Alzheimer's dementia. *Nat. Med.* **26**, 379–386 (2020).
12. Karikari, T. K. et al. Blood phosphorylated tau 181 as a biomarker for Alzheimer's disease: a diagnostic performance and prediction modelling study using data from four prospective cohorts. *Lancet Neurol.* **19**, 422–433 (2020).
13. Thijssen, E. H. et al. Diagnostic value of plasma phosphorylated tau181 in Alzheimer's disease and frontotemporal lobar degeneration. *Nat. Med.* **26**, 387–397 (2020).
14. Zetterberg, H. Neurofilament light: a dynamic cross-disease fluid biomarker for neurodegeneration. *Neuron* **91**, 1–3 (2016).
15. Quiroz, Y. T. et al. Plasma neurofilament light chain in the presenilin 1 E280A autosomal dominant Alzheimer's disease kindred: a cross-sectional and longitudinal cohort study. *Lancet Neurol.* **19**, 513–521 (2020).
16. Mielke, M. M. et al. Plasma phospho-tau181 increases with Alzheimer's disease clinical severity and is associated with tau- and amyloid-positron emission tomography. *Alzheimers Dement.* **14**, 989–997 (2018).
17. Mattsson, N., Cullen, N. C., Andreasson, U., Zetterberg, H. & Blennow, K. Association between longitudinal plasma neurofilament light and neurodegeneration in patients with Alzheimer disease. *JAMA Neurol.* **76**, 791–799 (2019).
18. Palmqvist, S. et al. Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders. *J. Am. Med. Assoc.* **324**, 772–781 (2020).

19. Van Maurik, I. S. et al. Interpreting biomarker results in individual patients with mild cognitive impairment in the Alzheimer's Biomarkers In Daily Practice (ABIDE) project. *JAMA Neurol.* **74**, 1481–1491 (2017).
20. Van Maurik, I. S. et al. Biomarker-based prognosis for people with mild cognitive impairment (ABIDE): a modelling study. *Lancet Neurol.* **18**, 1034–1044 (2019).
21. Gomersall, T., Smith, S. K., Blewett, C. & Astell, A. 'It's definitely not Alzheimer's': perceived benefits and drawbacks of a mild cognitive impairment diagnosis. *Br. J. Health Psychol.* **22**, 786–804 (2017).
22. Ossenkoppele, R. et al. Associations between tau, Aβ, and cortical thickness with cognition in Alzheimer disease. *Neurology* **92**, e601–e612 (2019).
23. Mattsson-Carlgren, N. et al. Aβ deposition is associated with increases in soluble and phosphorylated tau that precede a positive Tau PET in Alzheimer's disease. *Sci. Adv.* **6**, eaaz2387 (2020).
24. Leuzy, A. et al. Diagnostic performance of RO948 F 18 Tau positron emission tomography in the differentiation of Alzheimer disease from other neurodegenerative disorders. *JAMA Neurol.* **77**, 955–965 (2020).
25. Doecke, J. D. et al. Total $A\beta_{42}/A\beta_{40}$ ratio in plasma predicts amyloid-PET status, independent of clinical AD diagnosis. *Neurology* **94**, e1580–e1591 (2020).
26. Mattsson, N., Andreasson, U., Zetterberg, H. & Blennow, K. Alzheimer's Disease Neuroimaging Initiative Association of plasma neurofilament light with neurodegeneration in patients with Alzheimer disease. *JAMA Neurol.* **74**, 557–566 (2017).
27. Kuhlmann, J. et al. CSF $A\beta_{1-42}$—an excellent but complicated Alzheimer's biomarker—a route to standardisation. *Clin. Chim. Acta* **467**, 27–33 (2017).
28. Petersen, R. C. Mild cognitive impairment as a diagnostic entity. *J. Intern. Med.* **256**, 183–194 (2004).
29. Petrazzuoli, F. et al. Brief cognitive tests used in primary care cannot accurately differentiate mild cognitive impairment from subjective cognitive decline. *J. Alzheimers Dis.* **75**, 1191–1201 (2020).
30. McKhann, G. et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34**, 939–944 (1984).
31. Ovod, V. et al. Amyloid β concentrations and stable isotope labeling kinetics of human plasma specific to central nervous system amyloidosis. *Alzheimers Dement.* **13**, 841–849 (2017).
32. Akaike, H. Likelihood of a model and information criteria. *J. Econom.* **16**, 3–14 (1981).
33. Burnham K. P. & Anderson D. R. *Model Selection and Multimodel Inference* (Springer, 2002).

Acknowledgements

Work at the authors' laboratory at Lund University was supported by the Swedish Research Council, the Wallenberg Center for Molecular Medicine, the Knut and Alice Wallenberg Foundation, Lund University's Medical Faculty, Region Skåne, the Marianne and Marcus Wallenberg Foundation, the Strategic Research Area MultiPark (Multidisciplinary Research focused on Parkinson's Disease) at Lund University, the Swedish Alzheimer's Foundation, the Swedish Brain Foundation, the Swedish Medical Association, the Konung Gustaf V:s och Drottning Victorias Frimurarestiftelse, the Bundy Academy, the Skåne University Hospital Foundation and the Swedish federal government under the ALF agreement. Data collection and sharing for the ADNI part of the study was funded by the ADNI (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). The ADNI is funded by the National Institute on Aging and the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; Bioclinica; Biogen; Bristol Myers Squibb; CereSpir; Cogstate; Eisai; Elan Pharmaceuticals; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche and its affiliated company Genentech; Fujirebio; GE Healthcare; IXICO; Janssen Alzheimer Immunotherapy Research and Development; Johnson & Johnson Pharmaceutical Research and Development; Lumosity; Lundbeck; Merck & Co.; Meso Scale Diagnostics; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. 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). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Author contributions

A.L., N.M.-C. and O.H. conceived of the study. A.L. and N.C.C. performed the statistical analysis. E.S., S.P. and O.H. recruited participants and collected clinical data. H.Z., L.S., J.A.A., J.L.D., K.B., N.K.P., P.P. and S.J. were responsible for the biochemical analyses. N.M.-C. developed an online tool implementing the statistical models. A.L., N.M.-C., L.S., J.A.A., P.P., N.M.C. and O.H. drafted the initial manuscript. All authors contributed to revision and editing of the manuscript.

Competing interests

N.C.C., A.L., S.P., S.J. and N.M.-C. declare no competing interests. P.P., L.S. and J.A.A. are current employees of Araclon Biotech. N.K.P. and J.L.D. are employees and stockholders of Eli Lilly and Company. H.Z. has served on scientific advisory boards for Denali, Roche Diagnostics, Wave, Samumed, Siemens Healthineers, Pinteon Therapeutics and CogRx and has delivered lectures in symposia sponsored by Fujirebio, AlzeCure and Biogen. K.B. has served as a consultant on advisory boards or data monitoring committees for Abcam, Axon, Biogen, JOMDD/Shimadzu, Julius Clinical, Lilly, MagQu, Novartis, Roche Diagnostics and Siemens Healthineers and is a co-founder of Brain Biomarker Solutions in Gothenburg, which is a part of the GU Ventures Incubator Program. O.H. has acquired research support (for the institution) from Roche, Pfizer, GE Healthcare, Biogen, Eli Lilly and AVID Radiopharmaceuticals. In the past 2 years, he has received consultancy/speaker fees (paid to the institution) from Biogen and Roche. E.S. has no competing interests.

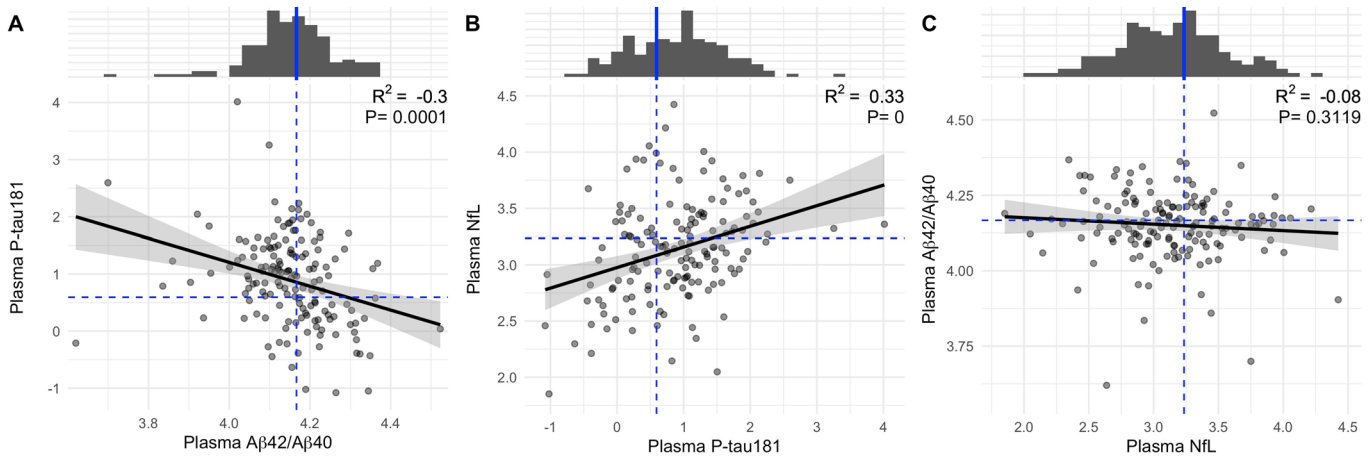
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s43587-020-00003-5.

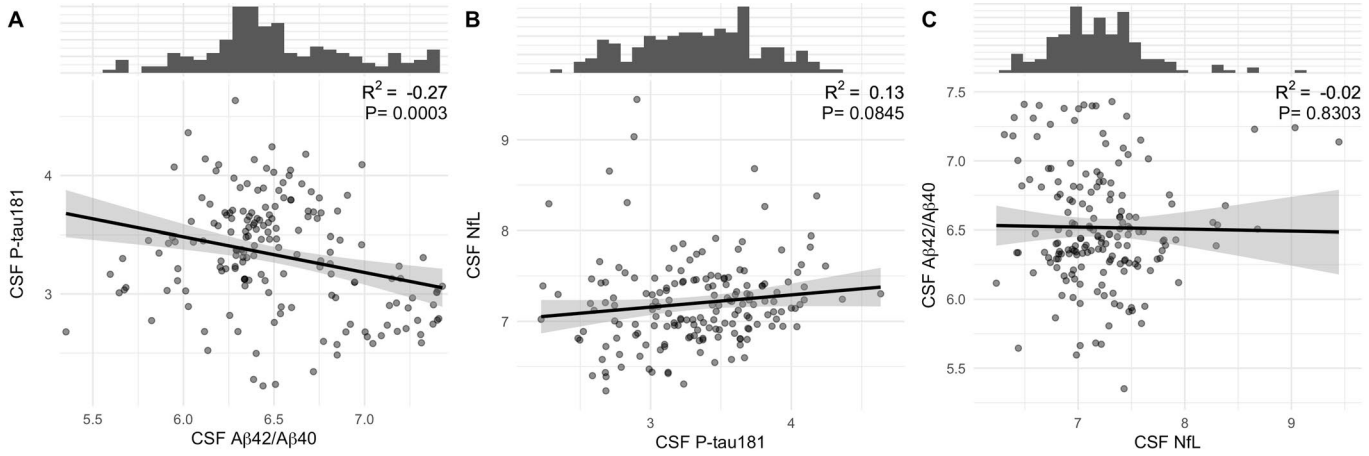
Supplementary information is available for this paper at https://doi.org/10.1038/s43587-020-00003-5.

Correspondence and requests for materials should be addressed to N.M.-C. or O.H.

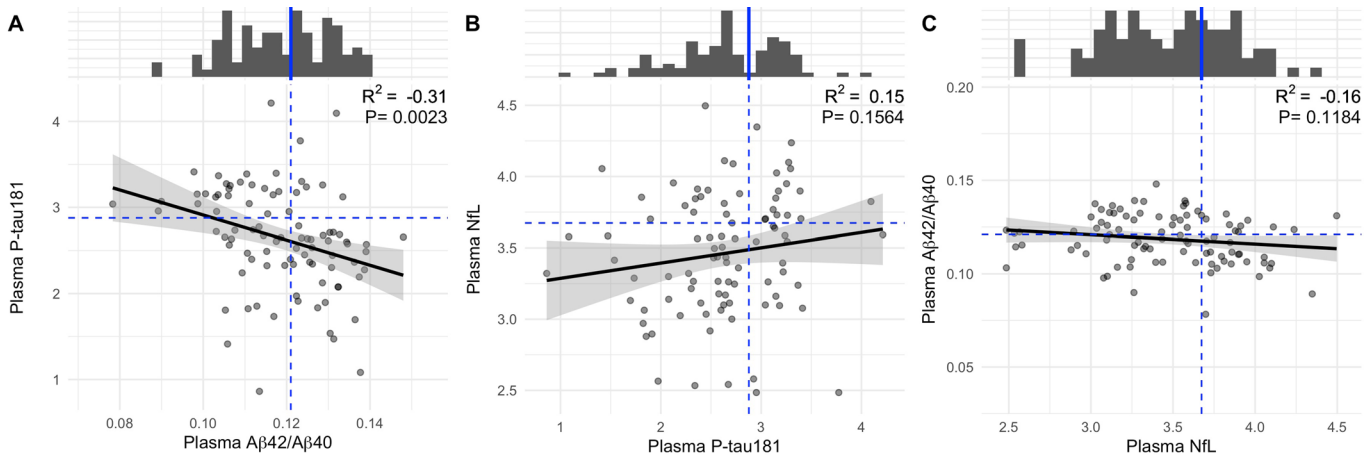
Peer review information *Nature Aging* thanks Suzanne Schindler, Charlotte Teunissen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



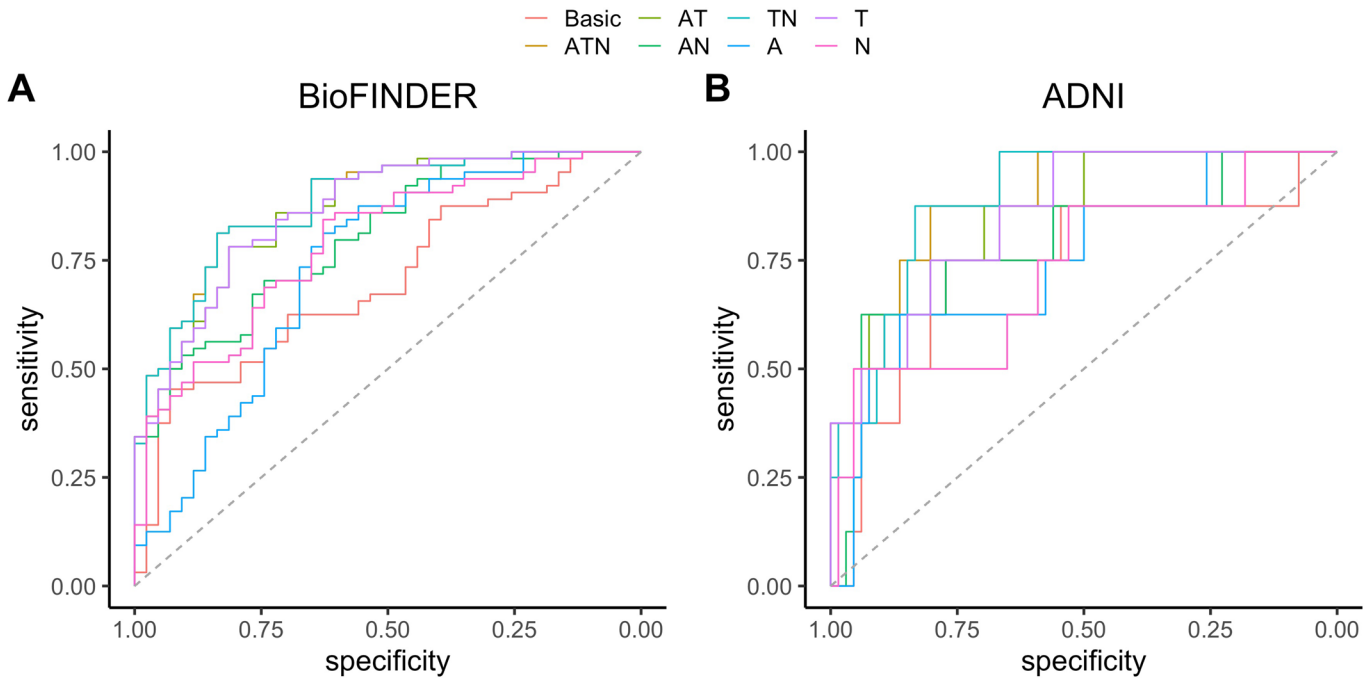
Extended Data Fig. 1 | Association between plasma biomarkers in MCI patients in the BioFINDER model selection/prognostic validation cohort. This figure shows the association between the different possible combinations of ATN plasma biomarkers (left to right: **a**, plasma P-tau181 vs plasma Aβ42/Aβ40; **b**, plasma NfL vs plasma P-tau181; and **c**, plasma Aβ42/Aβ40 vs plasma NfL). Each filled circle represents a single MCI patient from the BioFINDER cohort (n=148); the histograms above the scatterplots show the distribution of concentration values for the plasma biomarker on the x-axis (that is panel A, Aβ42/Aβ40; panel B, P-tau181; panel C, NfL). The blue lines (solid and dashed) indicate the cutoff values for each biomarker.



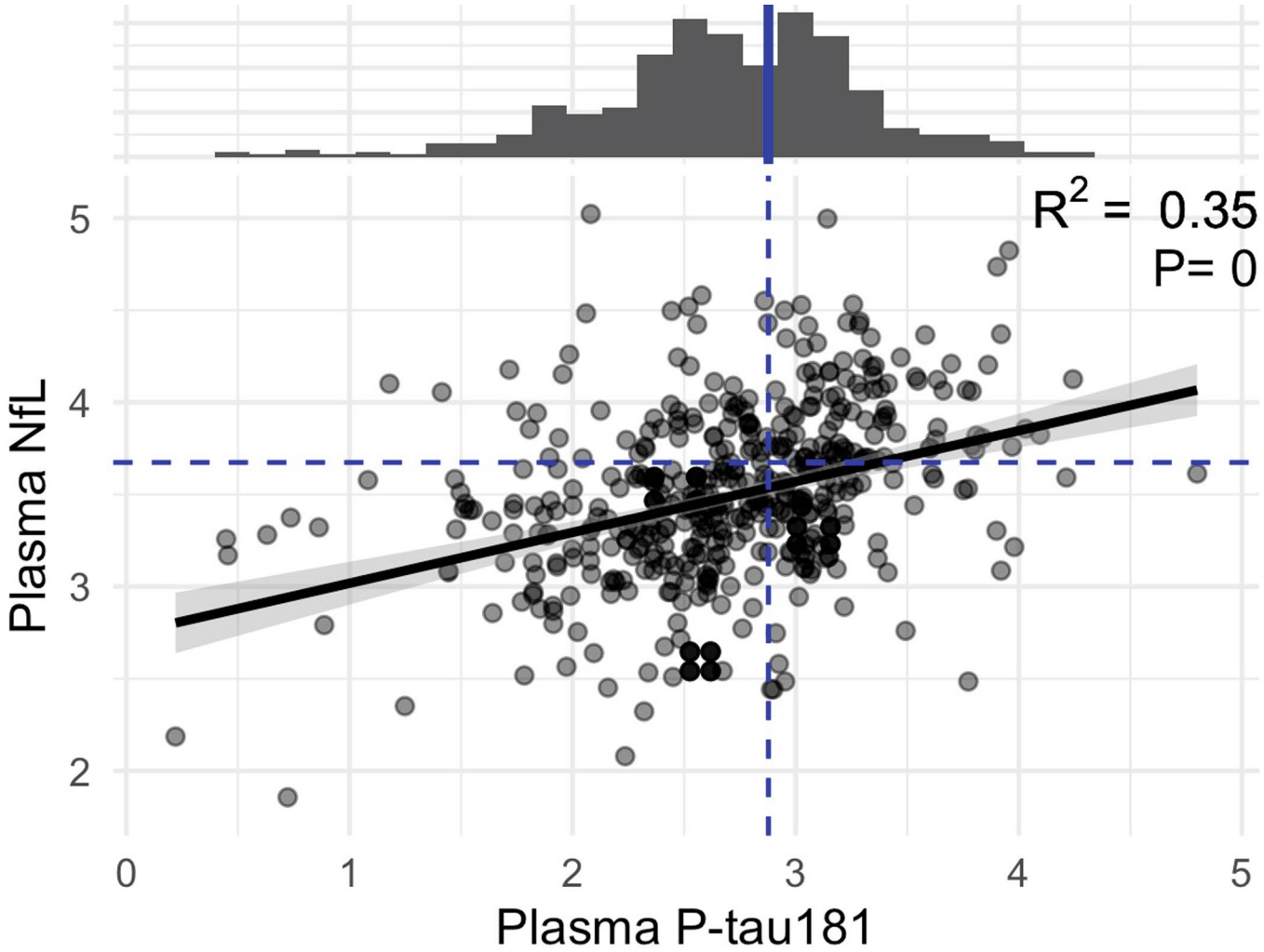
Extended Data Fig. 2 | Association between CSF biomarkers in MCI patients in the BioFINDER model selection/prognostic validation cohort. This figure shows the association between the different possible combinations of ATN CSF biomarkers (left to right: **a**, CSF P-tau181 vs CSF Aβ42/Aβ40; **b**, CSF NfL vs CSF P-tau181; and **c**, CSF Aβ42/Aβ40 vs CSF NfL). Each filled circle represents a single MCI patient from the BioFINDER cohort (n=148); the histograms above the scatterplots show the distribution of concentration values for the plasma biomarker on the x-axis (that is panel A, Aβ42/Aβ40; panel B, P-tau181; panel C, NfL).



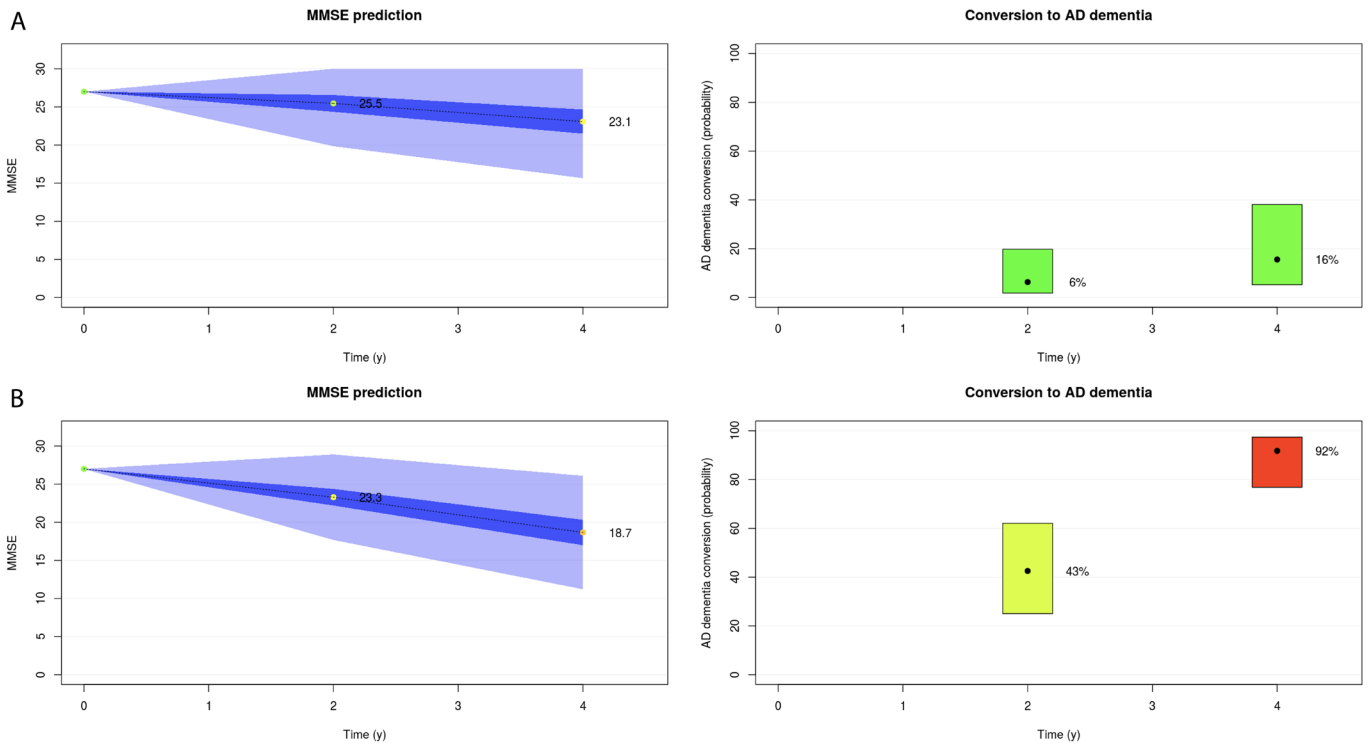
Extended Data Fig. 3 | Association between plasma biomarkers in MCI patients in the ADNI model selection cohort. This figure shows the association between the different possible combinations of ATN plasma biomarkers (left to right: **a**, plasma P-tau181 vs plasma Aβ42/Aβ40; **b**, plasma NfL vs plasma P-tau181; and **c**, plasma Aβ42/Aβ40 vs plasma NfL). Each filled circle represents a single MCI patient from the ADNI selection cohort (n=86); the histograms above the scatterplots show the distribution of concentration values for the plasma biomarker on the x-axis (that is panel A, Aβ42/Aβ40; panel B, P-tau181; panel C, NfL). The blue lines (solid and dashed) indicate the cutoff values for each biomarker.



Extended Data Fig. 4 | ROC curves for all models in the model selection analysis with four-year conversion to AD dementia as outcome. This figure shows ROC curves derived from the model selection analysis with four-year conversion to AD dementia as outcome. AUC values were the following in the BioFINDER cohort: Basic Model = 0.70, ATN=0.88, AT=0.86, AN=0.80, TN=0.88, A=0.73, T=0.86, N=0.79. AUC values were the following in the ADNI cohort: Basic Model = 0.74, ATN=0.88, AT=0.86, AN=0.79, TN=0.89, A=0.75, T=0.85, N=0.73. Full data is available in Supplementary Table 2.



Extended Data Fig. 5 | Association between plasma biomarkers in MCI patients in the ADNI prognostic validation cohort. Scatterplot showing plasma NfL versus plasma P-tau181 for the ADNI prognostic validation cohort (n=483). Each filled circle represents a single MCI patient; the histograms above the scatterplots show the distribution of concentration values for plasma P-tau181. The blue lines (solid and dashed) indicate the cutoff values for each biomarker.



Extended Data Fig. 6 | Online individualized risk prediction tool. Screenshots from our webtool predictprogression.com with predictions for a 70-year female with mild cognitive impairment and baseline MMSE 27. The upper panels (A) show predicted MMSE (left) and probability of conversion to AD dementia (right) at 2 and 4 years in a scenario where plasma Aβ42/Aβ40, P-tau181 and NfL are negative. In the lower panels (B), predicted MMSE (left) and probability of conversion to AD dementia (right) are shown at 2 and 4 years in a scenario where plasma Aβ42/Aβ40, P-tau181 and NfL are positive. In both A and B, the dark blue ribbons are 95% confidence interval of the estimates and the light blue ribbon are 90% prediction intervals. The predictions are derived from linear and logistic regression models that were established in the BioFINDER study on non-demented individuals with mild cognitive impairment.

nature research

Corresponding author(s): Oskar Hansson

Last updated by author(s): Oct 20, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	ll models were built using publicly available packages and functions in the R programming language (v4.0.0). The code that support the findings of this study are available from the corresponding author upon request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Plasma and CSF data from ADNI were downloaded online at https://ida.loni.usc.edu. Anonymized data from the BioFINDER study will be shared by request from a qualified academic investigator for the sole purpose of replicating procedures and results presented in the article and as long as data transfer is in agreement with EU legislation on the general data protection regulation and decisions by the Ethical Review Board of Sweden and Region Skåne, which should be regulated in a material transfer agreement.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We included all subjects who had data available for the variables of interest in BioFINDER and ADNI.
Data exclusions	No data was excluded from our analyses.
Replication	Prognostic validation was first done separately within each cohort using 1000 repetitions of five-fold cross validation (internal validation), and then by fitting the model on BioFINDER subjects and testing on ADNI subjects, and vice-versa (external validation). We replicated results at independent occasions before manuscript submission and during each revision step.
Randomization	As both BioFINDER and ADNI are observational cohort studies, allocation of participants into experimental groups was not performed (i.e. no randomization). All analyses were adjusted for age, sex, education and baseline cognition.
Blinding	Clinicians who evaluated cognitive decline and conversion to dementia were blinded to biomarker data. CSF and plasma analyses were performed by technicians blinded to clinical data.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Complete details for the assays used in the analysis of plasma and CSF can be found in the following publications: for BioFinder, plasma Aβ42/Aβ40 (Palmqvist et al., JAMA Neurol 2019), plasma P-tau181 (Janelidze et al. Nature Med 2020), plasma NfL (Mattsson et al., JAMA Neurol 2019), CSF Aβ42 and P-tau181 (Hansson et al. Alzheimers Dement 2018), CSF NfL (Mattsson et al., EMBO Mol Med 2016). In ADNI, plasma Aβ42/Aβ40 (Ovod et al., Alzheimers Dement 2017), plasma P-tau181 (Karikari et al., Lancet Neurol 2020).
Validation	These details are provided in the above mentioned publications.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Detailed information on the cohorts is provided in Tables 1 and 2 (and in the Supplement - Text 2 and 3). In short, we present results for analyses from two different prospective cohorts with similar study designs. For the model selection step, were included 148 MCI patients from the BioFINDER cohort (mean (IQR) age, 71.36 (5.47), 54 were women) and 86 MCI patients from ADNI (age 71.51 (7.59), 44 were women). For the prognostic validation step, we included 425 MCI patients from ADNI (70.98 (7.81), 220 were women).
Recruitment	MCI patients in the Swedish BioFINDER study were consecutively recruited memory clinics in the cities of Lund, Malmö and Ängelholm between July 2008 and June 2019. Approximately 90% of these were referred by primary care physicians. In the

ADNI study, MCI patients were recruited from many different clinics between September 2005 and December 2019, and thereby represent a more selected sample (i.e. closer to a clinical trial population). Though this difference could introduce bias, the fact that our findings held when validated in both cohorts—and were consistent even when performing sensitivity analyses—speaks to their robustness.

Ethics oversight

For BioFINDER, ethical approval was given by the Regional Ethical Committee of Lund University. Ethical approval in ADNI was given by the local ethical committees. The Ethics committees/institutional review boards that approved the ADNI study are: Albany Medical Center Committee on Research Involving Human Subjects Institutional Review Board, Boston University Medical Campus and Boston Medical Center Institutional Review Board, Butler Hospital Institutional Review Board, Cleveland Clinic Institutional Review Board, Columbia University Medical Center Institutional Review Board, Duke University Health System Institutional Review Board, Emory Institutional Review Board, Georgetown University Institutional Review Board, Health Sciences Institutional Review Board, Houston Methodist Institutional Review Board, Howard University Office of Regulatory Research Compliance, Icahn School of Medicine at Mount Sinai Program for the Protection of Human Subjects, Indiana University Institutional Review Board, Institutional Review Board of Baylor College of Medicine, Jewish General Hospital Research Ethics Board, Johns Hopkins Medicine Institutional Review Board, Lifespan - Rhode Island Hospital Institutional Review Board, Mayo Clinic Institutional Review Board, Mount Sinai Medical Center Institutional Review Board, Nathan Kline Institute for Psychiatric Research & Rockland Psychiatric Center Institutional Review Board, New York University Langone Medical Center School of Medicine Institutional Review Board, Northwestern University Institutional Review Board, Oregon Health and Science University Institutional Review Board, Partners Human Research Committee Research Ethics, Board Sunnybrook Health Sciences Centre, Roper St. Francis Healthcare Institutional Review Board, Rush University Medical Center Institutional Review Board, St. Joseph's Phoenix Institutional Review Board, Stanford Institutional Review Board, The Ohio State University Institutional Review Board, University Hospitals Cleveland Medical Center Institutional Review Board, University of Alabama Office of the IRB, University of British Columbia Research Ethics Board, University of California Davis Institutional Review Board Administration, University of California Los Angeles Office of the Human Research Protection Program, University of California San Diego Human Research Protections Program, University of California San Francisco Human Research Protection Program, University of Iowa Institutional Review Board, University of Kansas Medical Center Human Subjects Committee, University of Kentucky Medical Institutional Review Board, University of Michigan Medical School Institutional Review Board, University of Pennsylvania Institutional Review Board, University of Pittsburgh Institutional Review Board, University of Rochester Research Subjects Review Board, University of South Florida Institutional Review Board, University of Southern, California Institutional Review Board, UT Southwestern Institution Review Board, VA Long Beach Healthcare System Institutional Review Board, Vanderbilt University Medical Center Institutional Review Board, Wake Forest School of Medicine Institutional Review Board, Washington University School of Medicine Institutional Review Board, Western Institutional Review Board, Western University Health Sciences Research Ethics Board, and Yale University Institutional Review Board.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT01208675 (BioFINDER) and NCT00106899 (ADNI)
Study protocol	Clinical protocols can be obtained online: BioFINDER, www.biofinder.se; ADNI, adni.loni.usc.edu
Data collection	For BioFINDER, all data was collected between July 2008 and June 2019 at three memory clinics in the cities of Lund, Malmö and Ängelholm in the south of Sweden. For ADNI, data included in the present study were collected across 63 sites in the US and Canada between September 2005 and December 2019. These dates include clinical evaluations. Full details are provided at https://clinicaltrials.gov/ct2/show/NCT01208675?term=biofinder&rank=2 (BioFINDER) and https://clinicaltrials.gov/ct2/show/NCT00106899 (ADNI).
Outcomes	Primary outcomes were the global cognitive measure MMSE and clinical conversion to AD dementia evaluated four years after baseline ("four-year MMSE" and "four-year conversion to AD", respectively). As secondary outcomes, we used two-year MMSE and two-year conversion to AD dementia. As exploratory outcomes, we used two-year and four-year Clinical Dementia Rating Scale - Sum of Boxes (CDR-SB) along with two-year and four-year conversion to dementia due to any cause. Full details are provided at https://clinicaltrials.gov/ct2/show/NCT01208675?term=biofinder&rank=2 (BioFINDER) and https://clinicaltrials.gov/ct2/show/NCT00106899 (ADNI).