Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures

Sebastian Palmqvist (✉ sebastian.palmqvist@med.lu.se)
Lund University
Pontus Tideman
Lund University
Nicholas Cullen
Lund University
Henrik Zetterberg
University of Gothenburg https://orcid.org/0000-0003-3930-4354
Kaj Blennow
University of Gothenburg https://orcid.org/0000-0002-1890-4193
Jeffrey Dage
Eli Lilly and Company https://orcid.org/0000-0002-2192-1699
Erik Stomrud
Lunds University
Shorena Janelidze
Lund University https://orcid.org/0000-0003-2869-8378
Niklas Mattsson-Carlgren
Lund University
Oskar Hansson
Lund University https://orcid.org/0000-0001-8467-7286

Article

Keywords: Biomarkers, Subjective Cognitive Decline, Mild Cognitive Impairment, Logistic Regression Models, Memory, Executive Function, APOE Genotyping

Posted Date: January 5th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-134146/v1

Version of Record: A version of this preprint was published at Nature Medicine on May 24th, 2021. See the published version at https://doi.org/10.1038/s41591-021-01348-z.

Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures

Sebastian Palmqvist, MD, PhD[1,2*], Pontus Tideman, MSc[1,2], Nicholas Cullen, MSc[1], Henrik Zetterberg, MD, PhD[3,4,5,6], Kaj Blennow, MD, PhD[3,4], for the Alzheimer's Disease Neuroimaging Initiative, Jeffery L. Dage, PhD[7], Erik Stomrud, MD, PhD[1,2], Shorena Janelidze, PhD[1], Niklas Mattsson-Carlgren, MD, PhD[1,8,9], Oskar Hansson MD, PhD[1,2*]

[1] Clinical Memory Research Unit, Department of Clinical Sciences, Lund University, Sweden.
[2] Memory Clinic, Skåne University Hospital, Sweden.
[3] Department of Psychiatry and Neurochemistry, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden.
[4] Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden.
[5] Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London, United Kingdom.
[6] UK Dementia Research Institute at UCL, London, United Kingdom.
[7] Eli Lilly and Company, Indianapolis IN 46285 U.S.A.
[8] Department of Neurology, Skåne University Hospital, Sweden.
[9] Wallenberg Center for Molecular Medicine, Lund University, Lund, Sweden

* Corresponding authors

Oskar Hansson
e-mail: oskar.hansson@med.lu.se

Sebastian Palmqvist
e-mail: sebastian.palmqvist@med.lu.se

Both at
Memory Clinic
Skåne University Hospital
205 02 Malmö
Sweden

One sentence summary

Plasma P-tau in combination with brief and accessible measures predict future AD dementia in two independent cohorts with high accuracy that is clearly superior to specialists' clinical diagnostic prediction.

ABSTRACT

A combination of plasma phospho-tau (P-tau) and other accessible biomarkers may provide accurate prediction about the risk of developing Alzheimer's disease (AD) dementia. We examined this in patients with subjective cognitive decline (SCD) and mild cognitive impairment (MCI) who were consecutively recruited in the BioFINDER study (n=340). The results were validated in SCD/MCI participants in the ADNI study (n=543). Plasma P-tau, plasma Aβ42/Aβ40, plasma neurofilament light, *APOE* genotype, brief cognitive tests and MRI (cortical thickness in AD-specific regions) were examined as predictors of progression to AD dementia primarily within 4 years. The accuracy was determined using the area under the ROC curve (AUC) from logistic regression models. Within 4 years, plasma P-tau217 predicted AD dementia accurately (AUC 0.83) in BioFINDER. A model of plasma P-tau217, memory, executive function, and *APOE* had higher accuracy (AUC 0.91, $p<0.001$). In ADNI, this model produced a similar AUC (0.90) using plasma P-tau181 instead of P-tau217. A cross-validated version of this model was implemented online for prediction of the individual probability of progressing to AD dementia. Within 2 and 6 years, parsimonious models performed similar in both cohorts (AUCs 0.90-0.91). Using cerebrospinal fluid P-tau, Aβ42/Aβ40 and neurofilament light instead of plasma biomarkers did not improve the accuracy. The clinical prediction by memory clinic physicians had significantly lower accuracy than all models (4-year AUC 0.71). In summary, plasma P-tau in combination with brief cognitive tests and *APOE* genotyping may greatly improve the diagnostic prediction of AD dementia and facilitate recruitment for AD trials.

INTRODUCTION

Correctly determining if a patient with subtle cognitive symptoms, such as memory decline, suffers from prodromal or preclinical Alzheimer's disease (AD) and will progress to AD dementia within the near future remains a challenge for clinicians. The task is nonetheless of utmost importance for a timely referral to a memory clinic, a correct and early AD diagnosis, initiation of symptomatic treatment, planning for the future, and hopefully soon for initiating disease-modifying treatments. Although there have been impressive developments in biomarkers for AD and progression to AD dementia, such as cerebrospinal fluid analysis of β-amyloid (Aβ42[1] or the ratio of Aβ42/Aβ40[2]), phosphorylated tau (P-tau)[3,4] and neurofilament light (NfL)[5], as well as Aβ-PET[6,7] and tau-PET[8,9], the invasive nature, high cost, and limited availability restrict their use to a limited number of highly specialized centers. A possible turning point has emerged with the recent development of blood-based biomarkers, making it possible to measure NfL[10,11], Aβ42/Aβ40[2,12,13], and P-tau (either phosphorylated at threonine 181 or 217)[14-16] in plasma.

Plasma P-tau181 and P-tau217 in particular have shown especially high diagnostic performance for discriminating AD dementia from other neurodegenerative diseases.[14-16] Plasma P-tau has also recently been shown to be suitable for individualized prediction of cognitive decline in individuals with mild cognitive impairment (MCI).[17] In the clinical work-up of patients with cognitive complaints, however, it is unlikely that plasma P-tau (or any other biomarker) will achieve the highest potential predictive accuracy on its own due to the multifactorial nature of AD etiology and its heterogenous clinical presentation. There is therefore now a need to identify which other measures plasma P-tau should be combined with to produce the most accurate prediction of future AD and establish an optimal diagnostic algorithm of non-invasive, cost-effective and easily available methods for early diagnosis of AD.

Further, before establishing plasma P-tau in clinical practice, alone or as part of an algorithm, it is important to examine whether it actually performs better than the clinical prediction made by a treating physician, which has previously not been examined. The aim of the current study was therefore to examine the accuracy of plasma P-tau among patients with mild cognitive symptoms for predicting future AD dementia when combined with other accessible and non-invasive biomarkers. The prediction included not only the discrimination between progression to AD dementia and stable cognitive symptoms but also versus progression to other dementias. The accuracies were compared with the diagnostic prediction of memory clinic physicians who had performed extensive clinical assessments and evaluated cognitive testing and structural brain imaging at baseline. Selection of variables and accuracies from the models were examined in two independent multicenter cohorts. The primary cohort was the Swedish BioFINDER study and the validation cohort was the Alzheimer's Disease Neuroimaging Initiative (ADNI). The primary outcome was progression to AD dementia within 4 years and secondary outcomes were progression to AD dementia within 2 and 6 years, respectively. Finally, a cross-validated model was established and implemented as an online tool for predicting the individual risk of progressing to AD dementia (https://brainapps.shinyapps.io/PredictionADdementia/).

RESULTS

Participants in BioFINDER

The cohort included 164 patients with SCD and 176 patients with MCI of whom 91 progressed to AD dementia at follow-up, 49 to other dementias and 200 who did not progress to any dementia. The mean (standard deviation [SD]) age was 70.7 (5.6) years and 49% were women. The mean (SD) time to dementia was 2.9 (1.5) years and the mean (SD) follow-up time in participants that did not progress to dementia was 4.5 (1.6) years. Participant characteristics are described in Table 1 and the enrollment process in Supplementary Fig. 1.

Prediction of AD dementia in BioFINDER

Fig. 1 summarizes the model selection process and main results. First, a data-driven model selection was performed to select the model with the lowest Akaike information criterion (AIC), *i.e.*, the best model fit for predicting AD dementia. Variables screened for included key demographics, number of *APOE* ε4 alleles, brief tests from four cognitive domains, an MRI measure (cortical thickness from temporal regions prone to atrophy in AD[18]) and plasma biomarkers (NfL, P-tau217 and Aβ42/Aβ40). Then, a parsimonious model was created by removing as many variables as possible while maintaining a similar model fit defined as being within two AIC points of the lowest AIC model identified in the first step (ΔAIC <2). Thereafter, variables were removed further in a stepwise procedure to examine the performance of more basic models. Using the primary outcome, prediction of AD dementia within 4 years, the best model fit was found using plasma P-tau217, number of *APOE* ε4 alleles, executive function, memory function, cortical thickness, and plasma NfL (Fig. 1, Supplementary Table 1). This model resulted in an AUC of 0.92 (95% CI 0.89-0.95). Removing plasma NfL resulted in similar model fit (ΔAIC <2) and accuracy (AUC 0.91, 95% 0.88-0.95). Also removing cortical thickness (i.e. a model with plasma P-tau217, cognition,

and *APOE*) retained the same accuracy (AUC 0.91, 95% CI 0.87-0.94; p=0.09-0.82 vs the more complex models including the model with the best fit) but with slightly poorer model fit (ΔAIC +6.9) indicating that the model had slightly less confidence for predictions at the individual level. The AUC of the latter model was significantly higher than using only plasma P-tau217 (AUC 0.83, 95% CI 0.78-0.89; p<0.001; Fig. 1 and Supplementary Table 1).

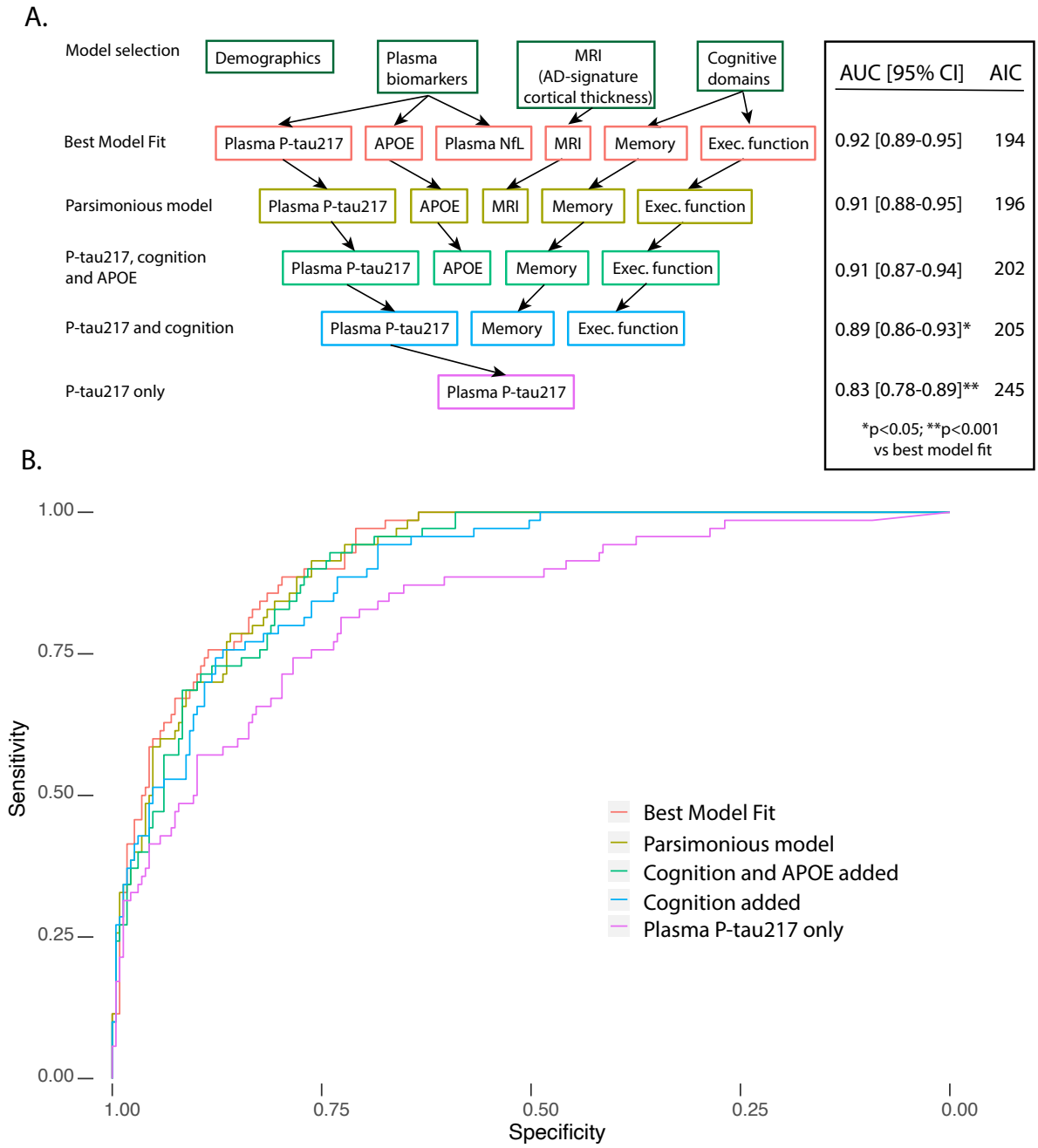

Figure 1. Model selection process and performance for predicting AD dementia within 4 years. A, The model selection process. Best Model Fit shows the data-driven model selection with the lowest AIC (*i.e.* the best model fit). The parsimonious model shows the model that had a similar model fit (*Δ*AIC <2) with as few significant predictors as possible. In subsequent models, modalities were removed in a step-wise procedure. Model specifications including comparisons between all models are shown in Supplementary Table 1. **B,** ROC curves of the different models.
Abbreviations: AD, Alzheimer's disease; AIC, Akaike Information Criterion; *APOE*, Apolipoprotein E genotype (number of

ε4 alleles); AUC, Area under the ROC curve; Exec. function, Executive Function; MRI, Cortical thickness of a temporal AD-specific region; ROC, Receiver Operating Characteristic

The best model fit for predicting AD dementia within 2 years resulted in similarly included variables and accuracies (Supplementary Fig. 3 and Supplementary Table 2). Model selection for prediction within 6 years, selected the same variables as for the 4-year outcome, but with the addition of plasma Aβ42/Aβ40 (AUC 0.94, 95% CI 0.91-0.97; Supplementary Fig. 4 and Supplementary Table 3).

Separate AUCs of univariable models for predicting progression to AD dementia for each time point from 2-6 years are shown in Supplementary Fig. 2. For cognition, the AUC for executive function was highest for short-term prediction but decreased over time, while memory continued to retained the accuracy through the prediction span. For the biomarkers of neurodegeneration (plasma NfL and cortical thickness), a similar trend was seen with higher AUCs for short- to mid-term prediction and a decrease after 4 years. In contrast, the AUCs for plasma P-tau217, plasma Aβ42/Aβ40 and presence of the *APOE* ε4 allele (associated with Aβ accumulation) increased from 2 to 6 years (Supplementary Fig. 2).

Comparison with the clinical prediction of AD dementia in BioFINDER

The subsample where the memory clinic physicians at baseline determined the most probable underlying cause of the cognitive impairment (*i.e.*, "clinical prediction") comprised 285 patients of whom 72 converted to AD dementia during follow-up. Using the primary outcome (progression to AD dementia within 4 years) the AUC for the clinical prediction was 0.72 (95% CI 0.65-0.78) and for plasma P-tau217 alone 0.81 (95% CI 0.75-0.87; p=0.03 vs the clinical prediction). Adding memory, executive function and *APOE* to plasma P-tau217 provided a further significantly higher accuracy than the clinical prediction (AUC 0.90, 95% CI 0.86-0.94; p<0.001). Additional improvements compared with the clinical prediction were

seen in the more complex models (Fig. 2, Supplementary Table 4), and they had similar accuracies to the performance in the whole population.

Significantly better accuracies for the models vs the clinical prediction were also seen at 2- and 6-years (Supplementary Tables 5-6).

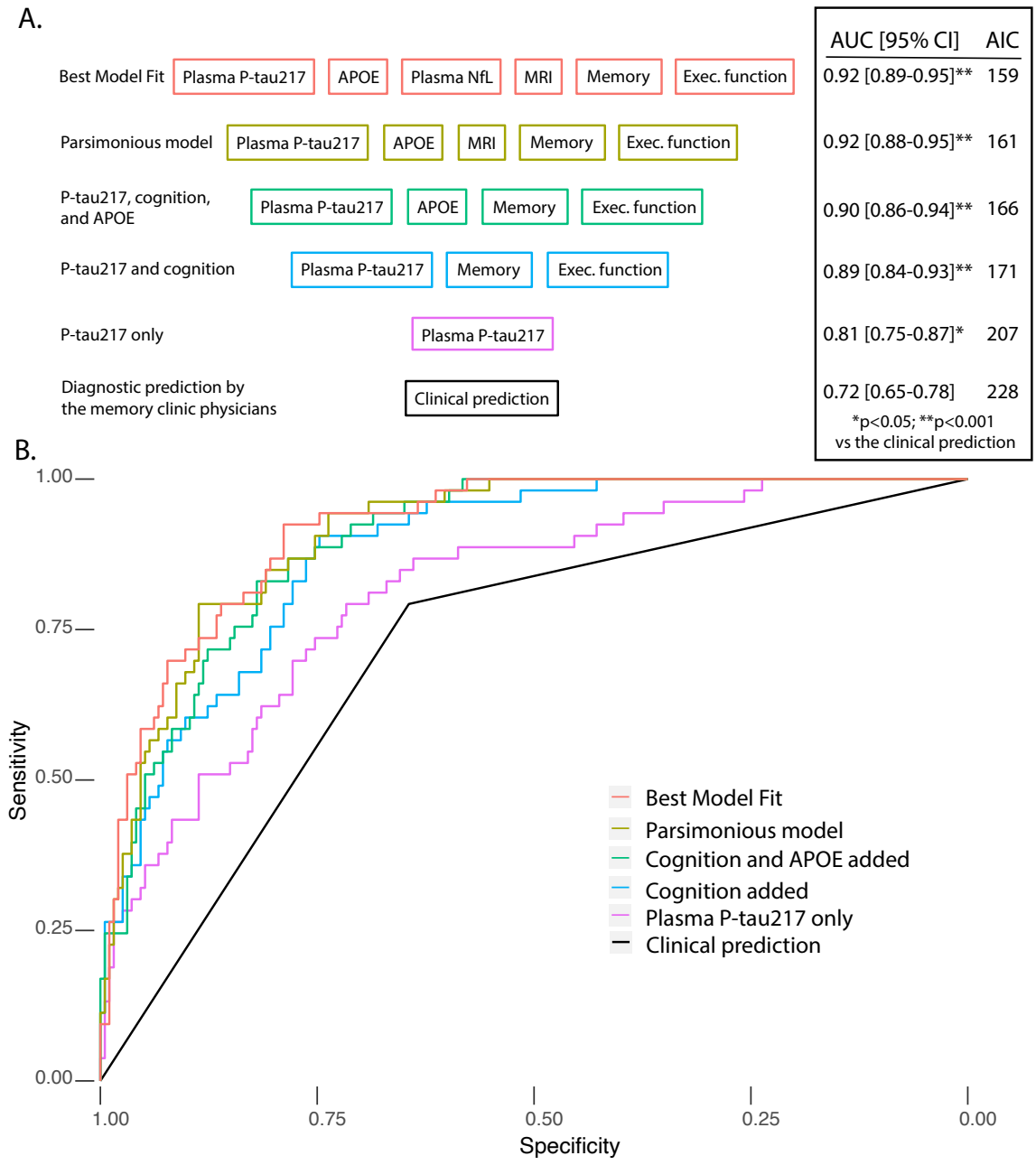

Figure 2. Comparison with the clinical prediction for predicting AD dementia within 4 years. A, Models from Fig. 1 compared to the clinical diagnostic prediction of memory clinic physicians at the baseline visit (blinded to fluid biomarker data). **B,** ROC curves of the different models. Note that comparison with the clinical prediction was performed on a subsample where the clinical prediction was available (n=285), hence the slightly different AUCs (and 95% CIs) compared with those shown in Fig. 1.
Abbreviations: AD, Alzheimer's disease; AIC, Akaike Information Criterion; *APOE*, Apolipoprotein E genotype (number of *ε4 alleles);* AUC, Area under the ROC curve; Exec. function, Executive Function; MRI, Cortical thickness of a temporal AD-specific region; ROC, Receiver Operating Characteristic

Comparison with CSF biomarkers

To examine the effect of using CSF biomarkers instead of plasma biomarkers, we tested different models with CSF P-tau, Aβ42/Aβ40, and NfL for the main outcome – prediction of AD dementia within 4 years (Supplementary Table 7). In the best model fit (Fig. 1, top model), the use of CSF P-tau and CSF NfL instead of plasma P-tau and plasma NfL produced a non-significantly different AUC (0.93, 95% CI 0.90-0.96; p=0.44 vs using the corresponding plasma biomarkers). In a new data-driven model selection using CSF instead of plasma biomarkers, CSF P-tau, CSF Aβ42/Aβ40, memory, executive function, and cortical thickness were selected for the best model fit. This model also produced a non-significantly different AUC when compared to the best model fit using plasma biomarkers (0.94, 95% CI 0.92-0.97; p=0.085). The more basic model with P-tau, *APOE*, memory and executive function had very similar accuracies (AUC 0.91 using either CSF or plasma P-tau, p=0.55). Finally, in univariate analyses, both CSF and plasma P-tau had an AUC of 0.83 (p=0.95).

Validation in the ADNI cohort

The cohort included 106 SCD and 437 MCI participants of whom 102 progressed to AD dementia at follow-up and 28 to other dementias (Table 1). The validation from BioFINDER was carried out in two steps. First, the same type of model selection was performed in ADNI (ranking models based on the AIC) to examine if similar variables were selected (with the exception that plasma P-tau181 was available instead of P-tau217 and that plasma Aβ42/Aβ40 measures only were available in a small subsample and therefore not included in the analysis;[19,20] see Supplementary Methods for details). Second, the key models identified from the BioFINDER cohort were tested in ADNI. Third, a cross-validated model was constructed and implemented online (see next section).

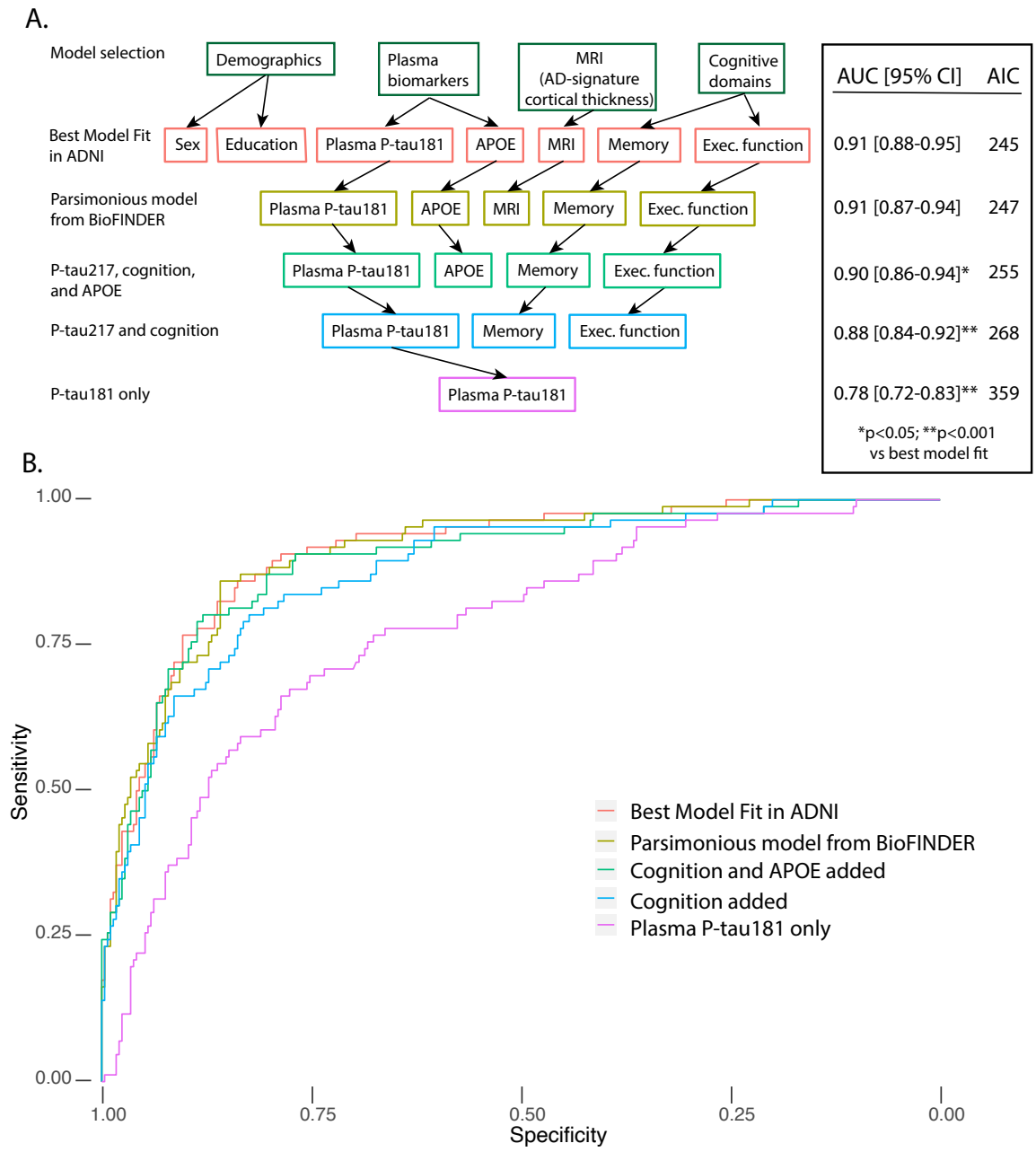

Figure 3. Model selection and performance in ADNI for predicting AD dementia within 4 years with comparisons using models selected in BioFINDER. A, Model selection process. Best Model Fit shows the data-driven model selection with the lowest AIC in the ADNI cohorts. This was compared to the models selected in BioFINDER and applied in ADNI. Model specifications are shown in Supplementary Table 8. **B,** ROC curves of the different models.
Abbreviations: AD, Alzheimer's disease; AIC, Akaike Information Criterion; *APOE*, Apolipoprotein E genotype (number of *ε4 alleles);* AUC, Area under the ROC curve; Exec. function, Executive Function; MRI, Cortical thickness of a temporal AD-specific region; ROC, Receiver Operating Characteristic

For the primary outcome (predicting AD dementia within 4 years), the same biomarkers as in BioFINDER were selected for the best model fit in ADNI, with the exception that plasma NfL was not chosen in ADNI (Fig. 3, Supplementary Table 8). Note that even though plasma NfL

was selected in BioFINDER it was not a significant predictor (Supplementary Table 1). When testing the variables selected in the parsimonious model from BioFINDER in ADNI (plasma P-tau, MRI, *APOE*, memory, and executive function), the accuracy (AUC 0.91, 95% CI 0.87-0.94) was not different from the best model fit established in ADNI (AUC 0.91, 95% CI 0.87-0.94; p=0.41, ΔAIC <2; Fig. 3, Supplementary Table 8). The more basic model with just P-tau, *APOE*, memory and executive function performed very similar in both cohorts (AUC 0.90, 95% CI 0.86-0.94 in ADNI vs 0.91, 95% CI 0.87-0.94 in BioFINDER).

For prediction of AD dementia within 2 years, the parsimonious model from BioFINDER again had a similar AUC in ADNI (0.91, 95% CI 0.87-0.94 vs 0.90, 95% CI 0.85-0.94 in BioFINDER; Supplementary Tables 2 and 9) and did not differ from the best model fit in ADNI (AUC 0.92, 95% CI 0.89-0.95; p=0.08, Supplementary Table 9). For prediction of AD dementia within 6 years, the best BioFINDER model not including plasma Aβ42/Aβ40 (*i.e.*, plasma P-tau, *APOE* and memory) performed similarly in both cohorts (AUC 0.90, 95% CI 0.86-0.94 in ADNI vs 0.91, 95% CI 0.87-0.95 in BioFINDER; Supplementary Tables 3 and 10).

A cross-validated model for calculating the individual probability of progressing to AD dementia within 4 years

As seen in the comparison between the BioFINDER and ADNI cohorts, a model consisting of plasma P-tau, memory, executive function and number of *APOE* ε4 alleles provided a good balance between simplicity, accuracy and generalizability for prediction of AD dementia within 4 years (AUCs 0.90-0.91; Fig. 1 and 4). We therefore created a new model in BioFINDER, where the estimates of the model could be directly tested in ADNI (and other cohorts). Because of the different plasma P-tau variants used in the cohorts, P-tau217 and P-tau181 were converted to binary variables (abnormal/normal). Unbiased plasma P-tau cutoffs

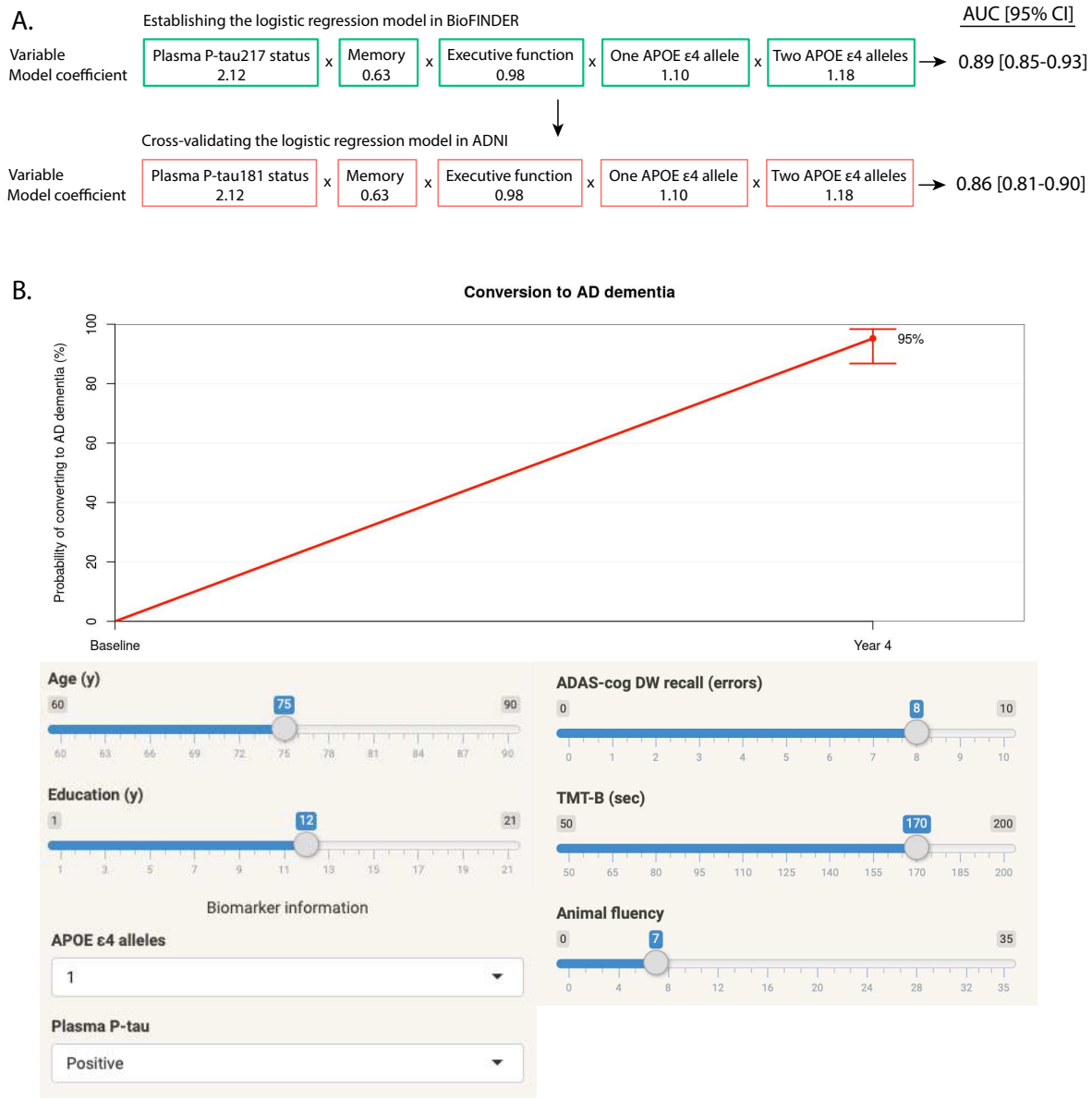

Figure 4. **A**. Cross-validation of the logistic regression model using plasma P-tau status instead of continuous plasma P-tau levels. Model coefficient were established in BioFINDER (AUC 0.89) and tested in ADNI (AUC 0.86). Z-scores have been inverted so that higher scores equal poorer results. **B**. Implementation of the logistic regression model at https://brainapps.shinyapps.io/PredictionADdementia/ where one can enter the raw cognitive test scores that constitute the z-scores, number of *APOE* ε4 alleles and plasma P-tau status (either from P-tau217 or P-tau181). Age and education are not part of the logistic regression model, but used to calculate cognitive z-scores. The example shows that the risk of progressing to AD dementia within 4 years is 95% in a 75-year old individual with cognitive complaints that has 12 years of education, one *APOE* ε4 allele, abnormal plasma P-tau status and scores 8 errors on 10-word delayed recall (*i.e.*, remembers 2 words), names 7 animals in one minute and completes the Trail-Making Test B in 170 seconds. Error bars represent the 95% prediction interval.

were established in independent populations in BioFINDER and ADNI at the mean level + 2 SD in Aβ-negative controls (see Supplementary Methods). Cognitive domain scores were established based on the distribution in an independent control sample in BioFINDER without tau or Aβ pathology, adjusted for age and education (see Supplementary Methods).[21,22] The

new model had an AUC of 0.89 (95% CI 0.85-0.93) in BioFINDER and when the estimates of the model were validated in ADNI, the AUC was 0.86 (95% CI 0.81-0.90), see Fig. 4A. This model was implemented online at https://brainapps.shinyapps.io/PredictionADdementia/ where the individual probability of progressing to AD dementia within 4 years can be calculated by entering the person's plasma P-tau status (abnormal/normal), number of *APOE* ε4 alleles, age and education level (for calculating cognitive domain z-scores), and the test scores from the 10-word delayed recall test, Trail-Making test B, and animal fluency (Fig. 4B).

DISCUSSION

We examined how plasma P-tau best could be combined with other easily accessible and cost-effective measures to predict progression to AD dementia, primarily within 4 years, in a heterogenous and consecutively included memory clinic cohort. Although plasma P-tau alone could predict AD dementia accurately (AUC 0.83) within 4 years, the most marked increase in accuracy was seen when it was combined with brief cognitive tests of memory and executive function and *APOE* genotype (AUC 0.91; Fig. 1, Supplementary Table 1). Minor further improvements were seen when also including cortical thickness and plasma NfL (Fig. 1). Similar accuracies were found when using CSF biomarkers instead of plasma biomarkers (Supplementary Table 7). Plasma P-tau217 alone and in combination with other variables had significantly higher accuracy than the clinical diagnostic prediction of memory clinic physicians after a comprehensive baseline clinical assessment including medical history, cognitive testing and CT or MRI of the brain (Fig. 2, Supplementary Table 4). The high generalizability of these predictors was demonstrated by a similar variable selection and performance in the independent ADNI cohort (Fig. 3). Especially the combination of plasma P-tau, memory, executive function, and *APOE* genotype had a robust performance and high accuracy in both cohorts (AUC 0.91 in both BioFINDER and ADNI; Fig. 1 and 4). This selection of variables was therefore used to create a cross-validated model using binary plasma P-tau results (abnormal or normal), that can be used to predict the individual risk of progressing to AD dementia (Fig. 4; https://brainapps.shinyapps.io/PredictionADdementia/).

Although there has been great progress recently in validating plasma P-tau as a biomarker for AD[14-16,23-27] and for individualized prediction of cognitive decline in individuals classified as having MCI[17], the present paper is the first to present how plasma P-tau can be combined with other easily available and cost-effective measures for predicting development AD dementia in patients seeking medical care based on diverse mild cognitive

symptoms. Another novelty is the comparison with the clinical prediction (Fig. 2), which shows the true value of implementing plasma P-tau alone or in combination with other measures to improve the diagnostic prediction in clinical practice. The other measures that were examined in this study were based on the literature of cognitive tests sensitive to the cognitive decline in AD[28-30], plasma and MRI biomarkers that have been shown to measure the underlying disease processes in AD at different stages[11,12,16,18] and known demographic and genetic risk factors of AD and cognitive decline.[31,32]

The variables that were selected for the 2-, 4- and 6-year models (Supplementary Tables 1-3) as well as the AUCs from the univariate models (Supplementary Fig. 2) showed the different measures were important depending on the timeframe. Regarding the cognitive measures, memory had high accuracy from short- to long-term predictions of AD dementia, while executive function had lower accuracies for long-term prediction. This suggest that memory changes earlier than executive function during the development of AD. Regarding the biomarkers, cortical thickness representing AD-specific neurodegeneration were best for short- to mid-term prediction, plasma P-tau217 for mid- to long-term prediction and plasma Aβ42/Aβ40 better for long-term prediction (Supplementary Fig. 2, Supplementary Tables 1-3). This is congruent with the model for the development of AD that begins with the accumulation of Aβ, then phosphorylation of tau and the deposition of tau tangles, and finally neurodegeneration.[4,33]

When comparing the effect of the biomarkers between cohorts, the accuracy of P-tau differed slightly. Plasma P-tau217 had higher predictive accuracy in BioFINDER than P-tau181 in ADNI. Although this might be caused by population-specific differences or pre-analytical differences, the somewhat higher accuracy of P-tau217 is in agreement with previous comparisons between P-tau217 and P-tau181 both in plasma[16,23] and in CSF[3,34].

Nonetheless, when combing plasma P-tau with other measures, the accuracy for the models was similar across both cohorts and P-tau variants (Fig. 1 and 4).

Using BioFINDER as the primary cohort of interest had some valuable strengths for determining optimal combinations of tests for use in clinical practice. The population consisted of *consecutively* recruited patients that had been referred to participating memory clinics, making the cohort heterogenous and representative of a future target population (Table 1). Nonetheless, similar results were obtained in ADNI that consists of a selected population focused on AD. The cognitive span in both cohorts ranged from subjective to objective cognitive symptoms (*i.e.*, both SCD and MCI participants), which best mimics the clinical scenario where physicians would use the combination of tests. Although the commonly utilized division into individuals with MCI and cognitively unimpaired individuals (SCD and controls) can make sense from a research standpoint for studying e.g. disease mechanisms in AD[33], this cognitive classification system is difficult to replicate between cohorts and translate into clinical practice. Depending on MCI definitions, use of cognitive tests, and cutoffs, those classified as MCI will greatly vary between populations.[35-37] And even if a unified definition and set of cognitive tests were determined, the comprehensive cognitive test battery needed would not fit the testing routines in *e.g.* primary care.[38]

This study has limitations. First, plasma Aβ42/Aβ40 was not available in an adequately large sample size in ADNI and could therefore not be included in that cohort. However, in BioFINDER where plasma Aβ42/Aβ40 was available, it was only selected for the prediction within 6 years and removing it from the model reduced the AUC by less than 0.02 suggesting that including plasma Aβ42/Aβ40 in ADNI may not have affected the overall results. Second, the cognitive tests that were available in both cohorts were limited. This resulted in suboptimal tests for the verbal and visuospatial domains, which could have contributed to their lower accuracy (Supplementary Fig. 2). On the other hand, those

representing the memory and executive domains have been extensively validated as sensitive measures of early cognitive decline in AD.[28-30] Third, the cross-validated model including plasma P-tau, memory, executive function and number of *APOE* ε4 alleles (Fig. 4) used binary P-tau data (abnormal/normal) since plasma P-tau217 was available in BioFINDER and plasma P-tau181 in ADNI; that is, estimates from a continuous, unified P-tau measure could not be cross-validated. Such a cross-validation would nonetheless be premature at this stage of the validation process since it is yet to be determined which variant of P-tau (181 or 217), assay and platform will be used in clinical practice. There are currently several different assays/platforms being set-up for clinical use and it is probable that not one single method will be used. This favors the use of the present cross-validated model that only required a binary result from a plasma P-tau measure. Foremost, this study provides robust evidence for which measures to combine with P-tau regardless of platform or type of P-tau used, since the P-tau assays differed between BioFINDER and ADNI but still rendered similar model results.

As for the potential diagnostic improvements in clinical practice, our comparison with the clinical-based prediction shows a clear advantage of using plasma P-tau in combination with the other measures (increases in AUC from 0.72 to 0.89-0.92; Fig. 2). The clinical prediction was here based on the baseline assessment of memory clinic physicians, showing the potential improved value at a specialist center. In addition, the presented models showed similarly high accuracy as when using CSF biomarkers, instead of plasma biomarkers (Supplementary Table 7). This suggests that the plasma P-tau models may provide a comparable substitute for CSF analyses in settings where these are not accessible or too expensive. A venous puncture (for plasma analyses) is also easier for the patient to undergo compared to lumbar puncture (for CSF analysis). In primary care, the implementation of these models is even more important because of the restricted availability of accurate diagnostic tools and the fact that only 20%-50% of cases with dementia are routinely recognized and

documented.[39,40] Presuming that primary care physicians make less accurate predictions of future AD dementia than memory clinic physicians, the advantage of using brief diagnostic algorithms based on plasma P-tau in primary care would be even greater. This implementation, however, would require further validation in large, unselected, and ethnically diverse primary care populations. A third area of use would be for recruitment of subjects with early AD to clinical trials. The presented models there may provide substantial cost-benefits compared to using only CSF analysis and PET to screen for eligible participants, which could speed up recruitments and hence facilitate the drug development process of future disease-modifying treatments for AD.[29]

MATERIAL AND METHODS

Participants

Participants from the Swedish BioFINDER study (http://biofinder.se; NCT01208675) consisted of consecutively included non-demented patients with mild cognitive symptoms referred to participating memory clinics, mostly directly from primary care units.[38] The patients were then categorized as having either subjective cognitive decline (SCD) or objective mild cognitive impairment (MCI) based on an extensive neuropsychological battery as previously described.[38] Participants with at least one follow-up visit and a complete baseline dataset of all variables included in the logistic regression models were selected for the present study. See Supplementary Fig. 1 for an enrollment flowchart. The participants were followed longitudinally with yearly follow-ups including cognitive testing, informant-based activities of daily living (ADL) questionnaires and detailed assessments by physicians experienced in neurocognitive disorders. All patients gave their written informed consent to participate and the study was approved by the regional ethics committee in Lund, Sweden.

The clinical diagnostic prediction

In a subgroup of patients (those included from the memory clinics in Malmö and Lund), the treating physician at the memory clinic was prospectively registering the most likely underlying cause of the cognitive impairment (here called the "clinical prediction") in the clinical research form at baseline. The clinical prediction was based on the first visit to the clinic (1.5 hour long visit with the patient and informant), informant-based cognitive symptom (CIMP-QUEST[41]) and ADL questionnaires (FAQ[42]), a broad cognitive test battery (https://biofinder.se/data-biomarkers/clinical-evaluation/), and a CT or MRI scan. The physicians were however blinded to all CSF, plasma and PET biomarker data, because these were performed after this initial visit.

Cognitive tests

Brief cognitive test that were available in both BioFINDER and ADNI were selected to approximately represent different cognitive domains. Trail Making Test B (TMT-B) and verbal fluency (animals) were selected as a measure of executive/attention performance based on their validated use in the preclinical Alzheimer's cognitive composite, PACC, which is sensitive in tracking cognitive decline in AD[30]. The 10-word delayed recall test from ADAS-cog has also been validated for detecting early cognitive decline in AD[28,29] and was chosen for the memory domain. The naming objects and fingers task from ADAS-cog was used for verbal performance[43], and the clock drawing test for visuospatial performance.[44] Each domain was converted to a z-score based on the test score distribution in the present population. See Supplementary Methods for the z-scores used in the cross-validated model.

Plasma biomarkers

Blood samples were collected at baseline and analyzed according to a standardized protocol.[12] Plasma P- tau217 concentrations were measured on a Meso-Scale Discovery platform (MSD, Rockville, MD), using an assay developed by Eli Lilly.[34] Plasma Aβ42 and Aβ40 concentrations were analyzed using the Elecsys immunoassays on a cobas e601 analyzer (Roche Diagnostics GmbH, Penzberg, Germany) and plasma NfL was measured by Simoa as previously described.[10]

CSF biomarkers

CSF was collected and handled according to a structured protocol as previously described.[1] P-tau and the Aβ42/Aβ42 were analyzed using the Elecsys immunoassays on a cobas e601 analyzer (Roche Diagnostics GmbH, Penzberg, Germany).[45] NfL was analyzed as previously described.[5]

MRI

The MRI protocol for BioFINDER has previously been described.[46] As MRI measure, cortical thickness in temporal brain regions susceptible to atrophy in AD was used (referred to as the "AD signature region"), as previously described.[18] This was chosen instead of hippocampal volume, which performed poorer for predicting progression to AD dementia (data not shown).

Outcomes

The primary outcome was prediction of progression to AD dementia vs progression to any other dementia or not progressing to any dementia within 4 years. Four years was chosen to reflect a clinically relevant timeframe in which it seems reasonable for a physician to give prognostic advice to an elderly patient and also a suitable timeframe for clinical trials to detect differences in conversion to dementia. Those who converted to AD dementia within

that timeframe were coded as "1" and stable SCD/MCI and conversion to any other dementia within the timeframe were coded as "0". Non-dementia converters with follow-ups <4 years were excluded from this analysis. Conversion to AD dementia within 2 and 6 years, respectively, were secondary outcomes with corresponding selection criteria for the examined population. The follow-up diagnosis was based on the treating physician's follow-up assessments and reviewed by a consensus group including memory clinic physicians and a senior neuropsychologist. The diagnosis was based on the DSM-5 criteria for major neurocognitive disorder due to probable AD. In addition, the patient was required to show signs of abnormal amyloid accumulation either according to CSF analysis[45] or Aβ PET[1].

Validation cohort – the Alzheimer's Disease Neuroimaging Initiative (ADNI)

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org.

According to our aim, we selected only non-demented subjects with cognitive symptoms at baseline. This included participants with MCI from the MCI cohort and participants from the healthy control cohort who had significant memory concerns (SMC; here referred to as "SCD"). Inclusion/exclusion criteria are described in detail at www.adni-info.org. Briefly, all subjects in the present study were between the ages of 55 and 91 years, had completed at least 6 years of education, were fluent in Spanish or English, and were free

of any significant neurologic disease other than AD. Subjects with SCD had Mini Mental State Examination score (MMSE) ≥24 and Clinical Dementia Rating (CDR) score 0, but expressed concerns of memory impairment. Subjects classified as MCI had MMSE ≥24, objective memory loss as shown on scores on delayed recall on the Wechsler Memory Scale Logical Memory II, CDR 0.5, preserved activities of daily living, and absence of dementia. All SCD and MCI participants with at least one follow-up visit and a complete dataset of variables included in the logistic regression models, were included in the present study. The variables used in the model selection were the same as in BioFINDER except for the plasma P-tau biomarker (in ADNI, P-tau181 was available instead of P-tau217). In addition, ADNI had two different plasma Aβ42/Aβ40 assays and each one was only available in very small datasets. Therefore, plasma Aβ42/Aβ40 was not included in the ADNI models. To create a similar outcome variable as in BioFINDER, participants were deemed converters to AD dementia if they had a follow-up diagnosis of AD dementia *and* were Aβ positive according to the Aβ PET scan.[47] Cognitively stable participants and converters to other dementias or Aβ-negative AD were thus coded as non-AD dementia converters.

Statistical analysis

Conversion to AD dementia was used as the dependent variable in logistic regression models. All continuous independent variables were transformed to z-scores based on the distribution in the present population. *APOE* ε4 genotype was coded into two different variables; presence of just one ε4 allele and presence of two ε4 alleles, as per previously described differences in their risk of AD.[29,31] The initial model selection was performed using the R package *MuMIn*, which tests all different combinations of variables and then ranks the models according to the Akaike information criterion (AIC). The model with the lowest AIC (*i.e.*, had the best model fit) were selected as the model with the best model fit. The next step was then to find models

with as few variables as possible that produced a similar model fit (defined as ΔAIC<2 from the optimal model[27]). Therefore, a stepwise removal of variables was performed as long as the ΔAIC was <2 from the optimal model to end up with the "parsimonious model". In addition, only variables with p<0.10 were kept in the parsimonious model. Then, variables were removed using a stepwise procedure in subsequent models to illustrate the added value of the different variables to plasma P-tau. This process was repeated for the time points 2, 4 and 6 years. Comparisons of area under the receiver operating characteristics curve (AUC) were performed using DeLong statistics. A two-sided P-value <0.05 was considered statistically significant. R version 4.0 was used for all statistical analyses.

DISCLOSURES

HZ has served at scientific advisory boards for Denali, Roche Diagnostics, Wave, Samumed, Siemens Healthineers, Pinteon Therapeutics and CogRx, has given lectures in symposia sponsored by Fujirebio, Alzecure and Biogen, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program. KB has served as a consultant, at advisory boards, or at 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 AB (BBS), which is a part of the GU Ventures Incubator Program. JLD is an employee of Eli Lilly and Company. OH has acquired research support (for the institution) from AVID Radiopharmaceuticals, Biogen, Eli Lilly, Eisai, GE Healthcare, Pfizer, and Roche. In the past 2 years, he has received consultancy/speaker fees from AC Immune, Alzpath, Biogen, Cerveau and Roche. SP, SJ, PT, NC, ES, and NMC report no disclosures.

FUNDING

Work at the authors' research center was supported by the Swedish Research Council (2016-00906 and 2018-02052), the Knut and Alice Wallenberg foundation (2017-0383), the Marianne and Marcus Wallenberg foundation (2015.0125), the Strategic Research Area MultiPark (Multidisciplinary Research in Parkinson's disease) at Lund University, the Swedish Alzheimer Foundation (AF-745911, AF-930929, and AF-940046), the Swedish Brain Foundation (FO2019-0326, FO2019-0300, and FO2020-0271), The Parkinson foundation of Sweden (1280/20), the Skåne University Hospital Foundation (2020-O000028), Regionalt Forskningsstöd (2020-0314 and Projekt-2020-0383) and the Swedish federal government under the ALF agreement (2018-Projekt0279 and 2018-Projekt0226). HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council (2018-02532), the European Research Council (681712), Swedish State Support for Clinical Research (ALFGBG-720931), the Alzheimer Drug Discovery Foundation (ADDF), USA (201809-2016862), AD Strategic Fund and the Alzheimer's Association (ADSF-21-831376-C, ADSF-21-831381-C and ADSF-21-831377-C), the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860197 (MIRIADE), and the UK Dementia Research Institute at UCL. KB is supported by the Swedish Research Council (2017-00915), the Alzheimer Drug Discovery Foundation (ADDF), USA (RDAPB-201809-2016615), the Swedish Alzheimer Foundation (AF-742881), Hjärnfonden, Sweden (FO2017-0243), the Swedish state under the agreement between the Swedish government and the County Councils, the ALF-agreement (ALFGBG-715986), and European Union Joint Program for Neurodegenerative Disorders (JPND2019-466-236).The precursor of ^{18}F-flutemetamol was sponsored by GE Healthcare. The precursor of ^{18}F-RO948 was provided by Roche.

DATA AVAILABILITY

For BioFINDER data: Anonymized data 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 Swedish Ethical Review Authority and Region Skåne, which should be regulated in a material transfer agreement. For ADNI data: Data is stored (publicly available) at the loni database (https://ida.loni.usc.edu/).

CODE AVAILABILITY

No custom code or mathematical algorithm that was central to the conclusions were used in the study.

TABLES

Table 1. Baseline characteristics for the BioFINDER and ADNI cohorts

	BioFINDER			ADNI		
	Converting to AD dementia (N=91)	**Not converting to AD dementia (N=249)**	**P-value**	**Converting to AD dementia (N=102)**	**Not converting to AD dementia (N=441)**	**P-value**
Age (yrs)	72.1 (4.91)	70.2 (5.73)	**0.007**	73.2 (6.98)	71.2 (7.14)	**0.003**
Gender (female)	44 (48%)	124 (50%)	0.81	54 (53%)	229 (52%)	0.85
Education (yrs)	11.8 (3.44)	11.9 (3.56)	0.75	16.2 (2.64)	16.3 (2.63)	0.64
Baseline diagnosis						
MCI	71 (78.0%)	105 (42.2%)	**<0.001**	102 (100%)	335 (76.0%)	**<0.001**
SCD	20 (22.0%)	144 (57.8%)	**<0.001**	0 (0%)	106 (24.0%)	**<0.001**
Time to any dementia (yrs)	2.87 (1.59)	2.95 (1.45)		2.54 (1.90)	3.20 (2.38)	
Follow-up time, stable participants (yrs)	NA	4.55 (1.62)		NA	4.30 (2.38)	
Follow-up diagnosis						
AD	91 (100%)	0		118 (110%)	0	
DLB/PDD/PSP	0	10 (4.0%)		0	1 (0.2%)	
FTD	0	7 (2.8%)		0	1 (0.2%)	
VaD	0	23 (9.2%)		0	2 (0.2%)	
Other	0	8 (3.2%)		0	24 (5.5%)**	
Non-dementia	0	201 (80.7%)		0	416 (94.3%)	
MMSE (points)	27.1 (1.87)	28.2 (1.65)	**<0.001**	27.4 (1.81)	28.5 (1.56)	**<0.001**
Memory function (z-score)	-0.70 (0.82)	0.25 (0.94)	**<0.001**	-0.82 (1.01)	0.20 (0.89)	**<0.001**

	BioFINDER			ADNI		
	Converting to AD dementia (N=91)	**Not converting to AD dementia (N=249)**	**P-value**	**Converting to AD dementia (N=102)**	**Not converting to AD dementia (N=441)**	**P-value**
Verbal function (z-score)	-0.096 (0.89)	0.035 (0.75)	0.11	-0.39 (1.17)	0.090 (0.93)	**<0.001**
Executive function (z-score)	-0.46 (0.74)	0.17 (0.87)	**<0.001**	-0.493 (0.927)	0.114 (0.777)	**<0.001**
Visuospatial function (z-score)	-0.17 (0.96)	0.061 (1.01)	0.02	-0.33 (1.26)	0.075 (0.91)	**0.002**
Plasma P-tau217 (pg/mL)	0.40 (0.25)	0.17 (0.14)	**<0.001**	NA	NA	
Plasma P-tau181 (pg/mL)	NA	NA		24.4 (10.8)	15.8 (11.4)	**<0.001**
Plasma Aβ42/Aβ40 (pg/mL)	0.062 (0.006)	0.066 (0.008)	**<0.001**	NA	NA	
Plasma NfL (pg/mL)*	25.2 (10.9)	23.6 (20.8)	**0.002**	44.9 (18.1)	35.1 (19.0)	**<0.001**
Cortical thickness of the AD signature region (mm)*	2.30 (0.24)	2.48 (0.27)	**<0.001**	2.65 (0.18)	2.76 (0.15)	**<0.001**

* Calibration-related (NfL) or camera-related (MRI) differences make it difficult to directly compare the results between the cohorts for these biomarkers.
** There were 20 participants with a follow-up diagnosis of AD dementia who were Aβ-negative. These were coded as having other dementias.
Note that before analyzed in logistic regression models, biomarker concentrations were transformed so that higher values corresponded to more abnormal results.

REFERENCES

1 Palmqvist, S. *et al.* Accuracy of brain amyloid detection in clinical practice using cerebrospinal fluid beta-amyloid 42: a cross-validation study against amyloid positron emission tomography. *JAMA Neurol* **71**, 1282-1289, doi:10.1001/jamaneurol.2014.1358 (2014).

2 Janelidze, S. *et al.* CSF Abeta42/Abeta40 and Abeta42/Abeta38 ratios: better diagnostic markers of Alzheimer disease. *Ann Clin Transl Neurol* **3**, 154-165, doi:10.1002/acn3.274 (2016).

3 Barthelemy, N. R. *et al.* Cerebrospinal fluid phospho-tau T217 outperforms T181 as a biomarker for the differential diagnosis of Alzheimer's disease and PET amyloid-positive patient identification. *Alzheimers Res Ther* **12**, 26, doi:10.1186/s13195-020-00596-4 (2020).

4 Mattsson, N. *et al.* Aβ deposition is associated with increases in soluble and phosphorylated tau that precede a positive Tau PET in Alzheimer's disease. *Science Advances* **6**, eaaz2387 (2020).

5 Mattsson-Carlgren, N. *et al.* The implications of different approaches to define AT(N) in Alzheimer disease. *Neurology* **94**, e2233-e2244, doi:10.1212/WNL.0000000000009485 (2020).

6 Mattsson, N., Palmqvist, S., Stomrud, E., Vogel, J. & Hansson, O. Staging beta-Amyloid Pathology With Amyloid Positron Emission Tomography. *JAMA Neurol*, doi:10.1001/jamaneurol.2019.2214 (2019).

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. *JAMA* **321**, 1286-1294, doi:10.1001/jama.2019.2000 (2019).

8 Leuzy, A. *et al.* Diagnostic Performance of RO948 F18 Tau Positron Emission Tomography in the Differentiation of Alzheimer Disease From Other Neurodegenerative Disorders. *JAMA Neurol* (2020).

9 Ossenkoppele, R. *et al.* Discriminative Accuracy of [18F]flortaucipir Positron Emission Tomography for Alzheimer Disease vs Other Neurodegenerative Disorders. *JAMA* **320**, 1151-1162, doi:10.1001/jama.2018.12917 (2018).

10 Gisslen, M. *et al.* Plasma Concentration of the Neurofilament Light Protein (NFL) is a Biomarker of CNS Injury in HIV Infection: A Cross-Sectional Study. *EBioMedicine* **3**, 135-140, doi:10.1016/j.ebiom.2015.11.036 (2016).

11 Preische, O. *et al.* Serum neurofilament dynamics predicts neurodegeneration and clinical progression in presymptomatic Alzheimer's disease. *Nature medicine* **25**, 277-283, doi:10.1038/s41591-018-0304-3 (2019).

12 Palmqvist, S. *et al.* Performance of Fully Automated Plasma Assays as Screening Tests for Alzheimer Disease-Related beta-Amyloid Status. *Jama Neurology* **76**, 1060-1069, doi:10.1001/jamaneurol.2019.1632 (2019).

13 Schindler, S. E. *et al.* High-precision plasma beta-amyloid 42/40 predicts current and future brain amyloidosis. *Neurology* **93**, e1647-e1659, doi:10.1212/WNL.0000000000008081 (2019).

14 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. *Nature medicine* **26**, 379-386, doi:10.1038/s41591-020-0755-1 (2020).

15 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, doi:10.1016/S1474-4422(20)30071-5 (2020).
16 Palmqvist, S. *et al.* Discriminative Accuracy of Plasma Phospho-tau217 for Alzheimer Disease vs Other Neurodegenerative Disorders. *JAMA* **324**, 772-781, doi:10.1001/jama.2020.12134 (2020).
17 Cullen, N. C. *et al.* Individualized prognosis of cognitive decline and dementia in mild cognitive impairment based on plasma biomarker combinations. *Nature Aging*, doi:10.1038/s43587-020-00003-5 (2020).
18 Jack, C. R., Jr. *et al.* Defining imaging biomarker cut points for brain aging and Alzheimer's disease. *Alzheimers Dement* **13**, 205-216, doi:10.1016/j.jalz.2016.08.005 (2017).
19 Toledo, J. B. *et al.* Factors affecting Abeta plasma levels and their utility as biomarkers in ADNI. *Acta Neuropathol* **122**, 401-413, doi:10.1007/s00401-011-0861-8 (2011).
20 Ovod, V. *et al.* Amyloid beta concentrations and stable isotope labeling kinetics of human plasma specific to central nervous system amyloidosis. *Alzheimers Dement* **13**, 841-849, doi:10.1016/j.jalz.2017.06.2266 (2017).
21 Borland, E. *et al.* The Montreal Cognitive Assessment: Normative Data from a Large Swedish Population-Based Cohort. *J Alzheimers Dis* **59**, 893-901, doi:10.3233/JAD-170203 (2017).
22 Borland, E., Stomrud, E., van Westen, D., Hansson, O. & Palmqvist, S. The age-related effect on cognitive performance in cognitively healthy elderly is mainly caused by underlying AD pathology or cerebrovascular lesions: implications for cutoffs regarding cognitive impairment. *Alzheimers Res Ther* **12**, 30, doi:10.1186/s13195-020-00592-8 (2020).
23 Barthelemy, N. R., Horie, K., Sato, C. & Bateman, R. J. Blood plasma phosphorylated-tau isoforms track CNS change in Alzheimer's disease. *J Exp Med* **217**, doi:10.1084/jem.20200861 (2020).
24 Karikari, T. K. *et al.* Diagnostic performance and prediction of clinical progression of plasma phospho-tau181 in the Alzheimer's Disease Neuroimaging Initiative. *Mol Psychiatry*, doi:10.1038/s41380-020-00923-z (2020).
25 Mattsson-Carlgren, N. *et al.* Longitudinal plasma p-tau217 is increased in early stages of Alzheimer's disease. *Brain*, doi:10.1093/brain/awaa286 (2020).
26 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, doi:10.1016/j.jalz.2018.02.013 (2018).
27 Palmqvist, S. *et al.* Cerebrospinal fluid and plasma biomarker trajectories with increasing amyloid deposition in Alzheimer's disease. *EMBO Mol Med*, e11170, doi:10.15252/emmm.201911170 (2019).
28 Insel, P. S. *et al.* Determining clinically meaningful decline in preclinical Alzheimer disease. *Neurology* **93**, e322-e333, doi:10.1212/WNL.0000000000007831 (2019).
29 Palmqvist, S. *et al.* Accurate risk estimation of beta-amyloid positivity to identify prodromal Alzheimer's disease: Cross-validation study of practical algorithms. *Alzheimers Dement* **15**, 194-204, doi:10.1016/j.jalz.2018.08.014 (2019).
30 Papp, K. V., Rentz, D. M., Orlovsky, I., Sperling, R. A. & Mormino, E. C. Optimizing the preclinical Alzheimer's cognitive composite with semantic processing: The PACC5. *Alzheimers Dement (N Y)* **3**, 668-677, doi:10.1016/j.trci.2017.10.004 (2017).
31 Jansen, W. J. *et al.* Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *JAMA* **313**, 1924-1938, doi:10.1001/jama.2015.4668 (2015).

32 Scheyer, O. *et al.* Female Sex and Alzheimer's Risk: The Menopause Connection. *J Prev Alzheimers Dis* **5**, 225-230, doi:10.14283/jpad.2018.34 (2018).
33 Jack, C. R., Jr. *et al.* NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535-562, doi:10.1016/j.jalz.2018.02.018 (2018).
34 Janelidze, S. *et al.* Cerebrospinal fluid p-tau217 performs better than p-tau181 as a biomarker of Alzheimer's disease. *Nature communications* **11**, 1683 (2020).
35 Petersen, R. C. Mild cognitive impairment as a diagnostic entity. *J Intern Med* **256**, 183-194, doi:10.1111/j.1365-2796.2004.01388.x
JIM1388 [pii] (2004).
36 Roberts, R. O. *et al.* The Mayo Clinic Study of Aging: design and sampling, participation, baseline measures and sample characteristics. *Neuroepidemiology* **30**, 58-69, doi:10.1159/000115751 (2008).
37 Gothlin, M., Eckerstrom, M., Rolstad, S., Wallin, A. & Nordlund, A. Prognostic Accuracy of Mild Cognitive Impairment Subtypes at Different Cut-Off Levels. *Dement Geriatr Cogn Disord* **43**, 330-341, doi:10.1159/000477341 (2017).
38 Petrazzuoli, F. *et al.* Brief Cognitive Tests Used in Primary Care Cannot Accurately Differentiate Mild Cognitive Impairment from Subjective Cognitive Decline. *J Alzheimers Dis*, doi:10.3233/JAD-191191 (2020).
39 Lopponen, M., Raiha, I., Isoaho, R., Vahlberg, T. & Kivela, S. L. Diagnosing cognitive impairment and dementia in primary health care -- a more active approach is needed. *Age Ageing* **32**, 606-612, doi:10.1093/ageing/afg097 (2003).
40 Valcour, V. G., Masaki, K. H., Curb, J. D. & Blanchette, P. L. The detection of dementia in the primary care setting. *Arch Intern Med* **160**, 2964-2968, doi:10.1001/archinte.160.19.2964 (2000).
41 Åstrand, R., Rolstad, S. & Wallin, A. *Cognitive Impairment Questionnaire (CIMP-QUEST) - reported topographic symptoms in MCI and dementia.* (2009).
42 Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Jr., Chance, J. M. & Filos, S. Measurement of functional activities in older adults in the community. *J Gerontol* **37**, 323-329 (1982).
43 Rosen, W. G., Mohs, R. C. & Davis, K. L. A new rating scale for Alzheimer's disease. *Am J Psychiatry* **141**, 1356-1364 (1984).
44 Shulman, K. I. Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry* **15**, 548-561 (2000).
45 Hansson, O. *et al.* CSF biomarkers of Alzheimer's disease concord with amyloid-beta PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimers Dement*, doi:10.1016/j.jalz.2018.01.010 (2018).
46 Palmqvist, S. *et al.* Earliest accumulation of beta-amyloid occurs within the default-mode network and concurrently affects brain connectivity. *Nature communications* **8**, 1214, doi:10.1038/s41467-017-01150-x (2017).
47 Landau, S. M. *et al.* Amyloid deposition, hypometabolism, and longitudinal cognitive decline. *Ann Neurol* **72**, 578-586, doi:10.1002/ana.23650 (2012).

Figures

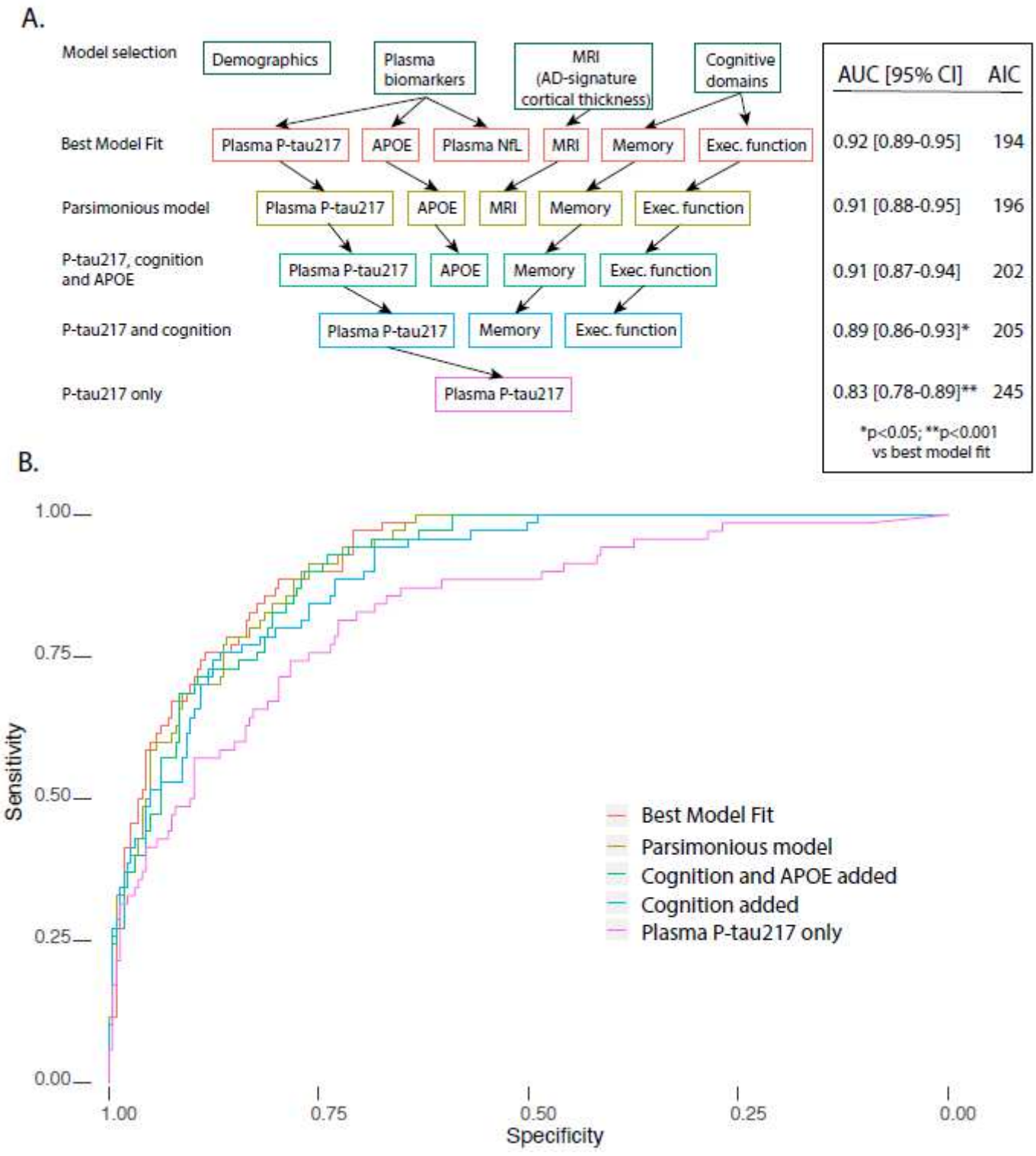

Figure 1

Model selection process and performance for predicting AD dementia within 4 years. A, The model selection process. Best Model Fit shows the data-driven model selection with the lowest AIC (i.e. the best model fit). The parsimonious model shows the model that had a similar model fit (ΔAIC <2) with as few

significant predictors as possible. In subsequent models, modalities were removed in a step-wise procedure. Model specifications including comparisons between all models are shown in Supplementary Table 1. B, ROC curves of the different models. Abbreviations: AD, Alzheimer's disease; AIC, Akaike Information Criterion; APOE, Apolipoprotein E genotype (number of ε4 alleles); AUC, Area under the ROC curve; Exec. function, Executive Function; MRI, Cortical thickness of a temporal AD-specific region; ROC, Receiver Operating Characteristic

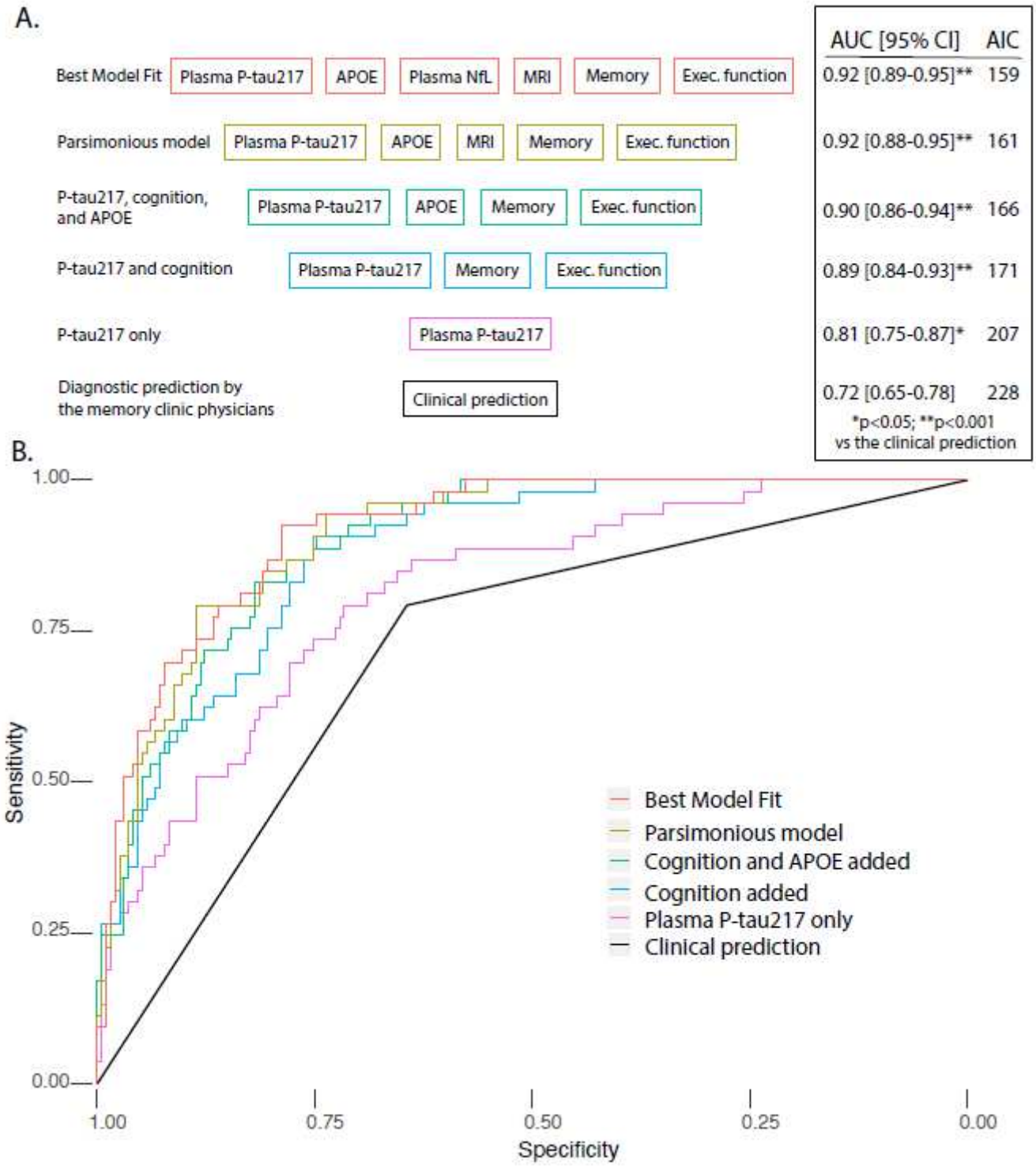

Figure 2

Comparison with the clinical prediction for predicting AD dementia within 4 years. A, Models from Fig. 1 compared to the clinical diagnostic prediction of memory clinic physicians at the baseline visit (blinded to fluid biomarker data). B, ROC curves of the different models. Note that comparison with the clinical prediction was performed on a subsample where the clinical prediction was available (n=285), hence the slightly different AUCs (and 95% CIs) compared with those shown in Fig. 1. Abbreviations: AD, Alzheimer's disease; AIC, Akaike Information Criterion; APOE, Apolipoprotein E genotype (number of ε4 alleles); AUC, Area under the ROC curve; Exec. function, Executive Function; MRI, Cortical thickness of a temporal AD-specific region; ROC, Receiver Operating Characteristic

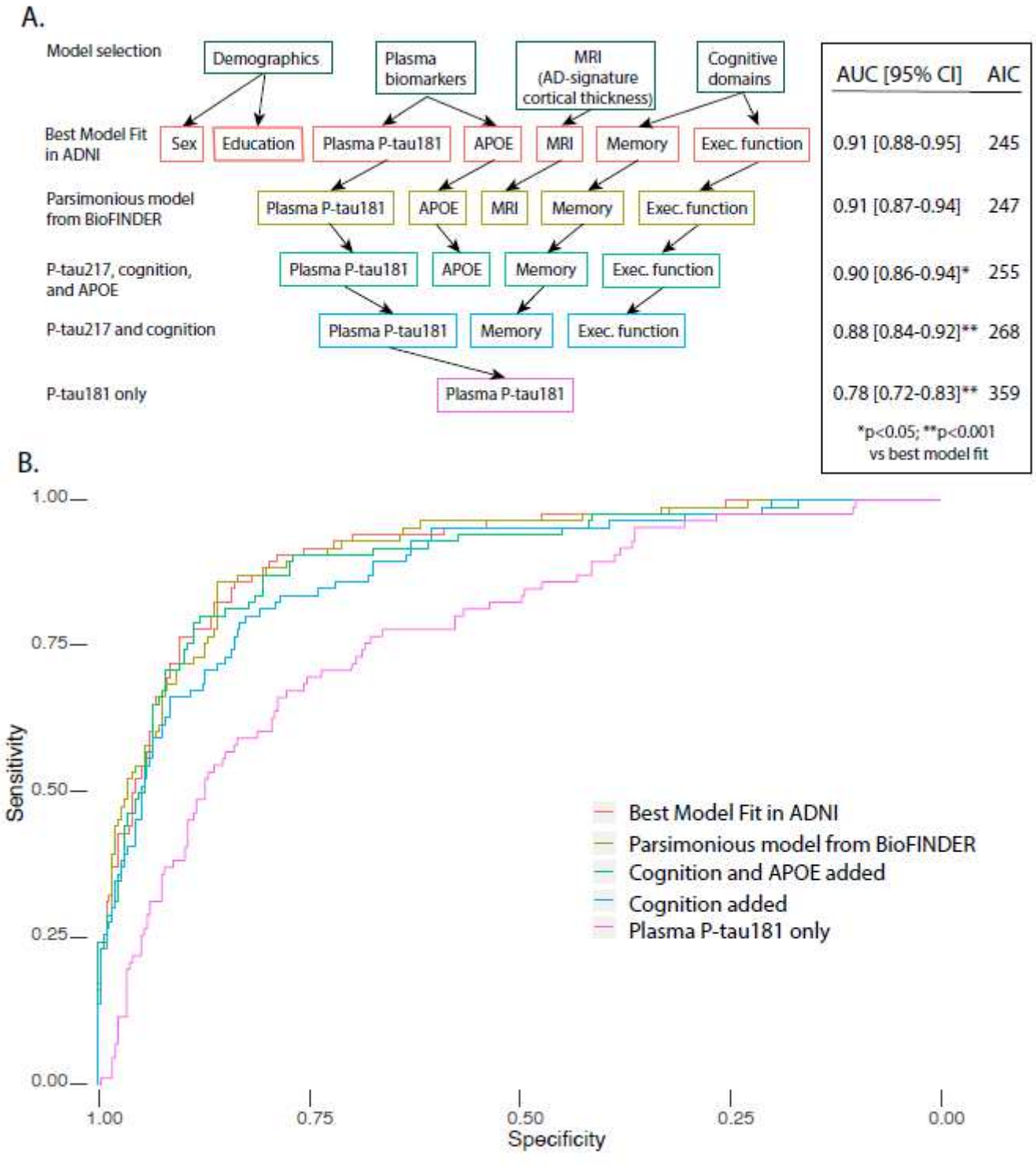

Figure 3

Model selection and performance in ADNI for predicting AD dementia within 4 years with comparisons using models selected in BioFINDER. A, Model selection process. Best Model Fit shows the data-driven model selection with the lowest AIC in the ADNI cohorts. This was compared to the models selected in BioFINDER and applied in ADNI. Model specifications are shown in Supplementary Table 8. B, ROC curves of the different models. Abbreviations: AD, Alzheimer's disease; AIC, Akaike Information Criterion; APOE, Apolipoprotein E genotype (number of ε4 alleles); AUC, Area under the ROC curve; Exec. function, Executive Function; MRI, Cortical thickness of a temporal AD-specific region; ROC, Receiver Operating Characteristic

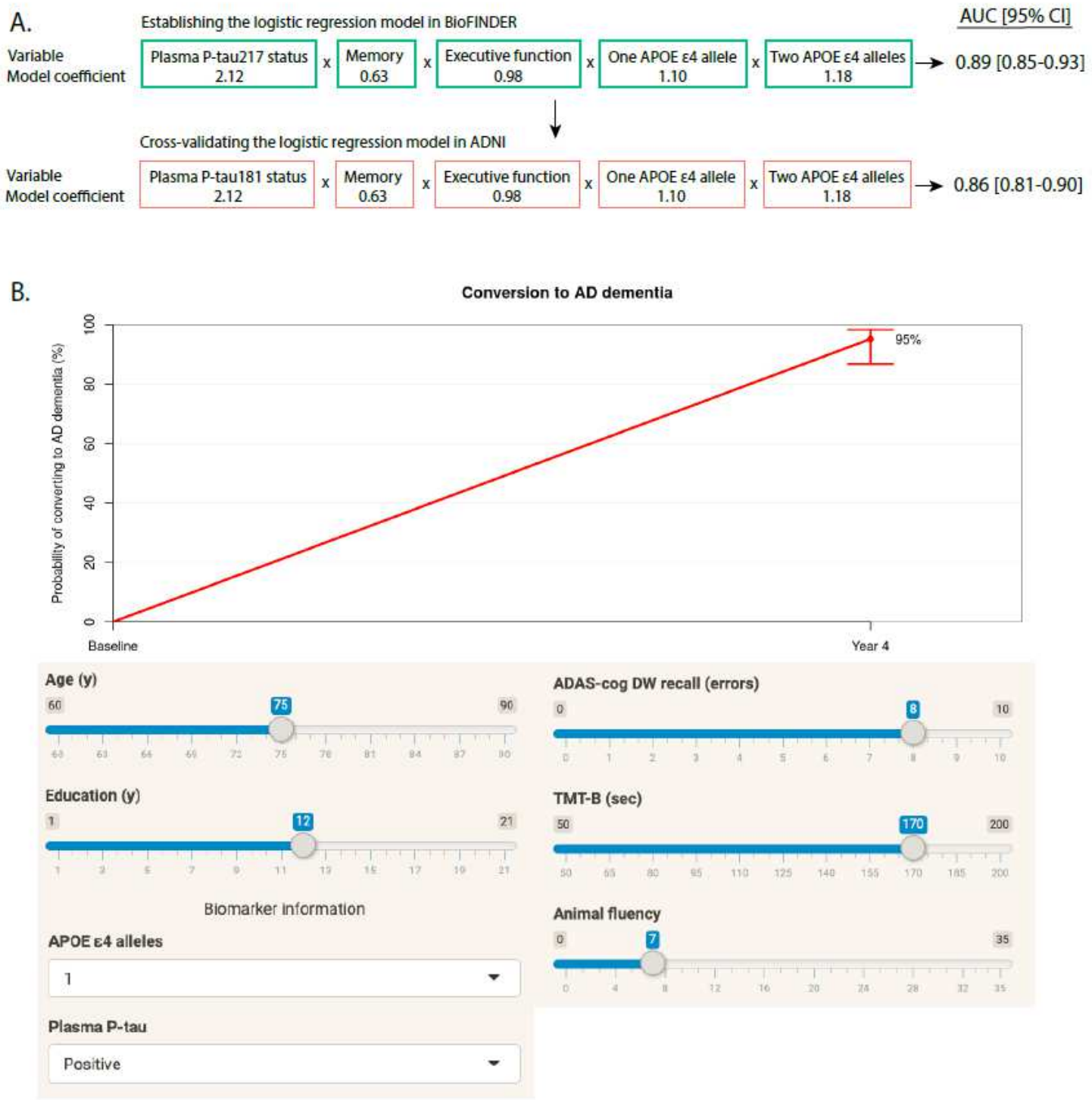

Figure 4

A. Cross-validation of the logistic regression model using plasma P-tau status instead of continuous plasma P-tau levels. Model coefficient were established in BioFINDER (AUC 0.89) and tested in ADNI (AUC 0.86). Z-scores have been inverted so that higher scores equal poorer results. B. Implementation of the logistic regression model at https://brainapps.shinyapps.io/PredictionADdementia/ where one can enter the raw cognitive test scores that constitute the z-scores, number of APOE ε4 alleles and plasma P-tau status (either from P-tau217 or P-tau181). Age and education are not part of the logistic regression

model, but used to calculate cognitive z-scores. The example shows that the risk of progressing to AD dementia within 4 years is 95% in a 75-year old individual with cognitive complaints that has 12 years of education, one APOE ε4 allele, abnormal plasma P-tau status and scores 8 errors on 10-word delayed recall (i.e., remembers 2 words), names 7 animals in one minute and completes the Trail-Making Test B in 170 seconds. Error bars represent the 95% prediction interval.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementsubmitted.pdf