Plasma amyloid β 40/42 ratio predicts cerebral amyloidosis in cognitively normal individuals at risk for Alzheimer's disease

(Article begins on next page)

Plasma amyloid β 40/42 ratio predicts cerebral amyloidosis in cognitively normal individuals at risk for Alzheimer's disease.

Vergallo A, Mégret L, Lista S, Cavedo E, Zetterberg H, Blennow K, Vanmechelen E, De Vos A, Habert MO, Potier MC, Dubois B, Neri C, Hampel H; INSIGHT-preAD study group; Alzheimer Precision Medicine Initiative (APMI)

Abstract

INTRODUCTION: Blood-based biomarkers of pathophysiological brain amyloid β (Aβ) accumulation, particularly for preclinical target and large-scale interventions, are warranted to effectively enrich Alzheimer's disease clinical trials and management.

METHODS: We investigated whether plasma concentrations of the Aβ1-40/Aβ1-42 ratio, assessed using the single-molecule array (Simoa) immunoassay, may predict brain Aβ positron emission tomography status in a large-scale longitudinal monocentric cohort (N = 276) of older individuals with subjective memory complaints. We performed a hypothesis-driven investigation followed by a no-a-priori hypothesis study using machine learning.

RESULTS: The receiver operating characteristic curve and machine learning showed a balanced accuracy of 76.5% and 81%, respectively, for the plasma Aβ1-40/Aβ1-42 ratio. The accuracy is not affected by the apolipoprotein E (APOE) ε4 allele, sex, or age.

DISCUSSION: Our results encourage an independent validation cohort study to confirm the indication that the plasma Aβ1-40/Aβ1-42 ratio, assessed via Simoa, may improve future standard of care and clinical trial design.

Introduction

Blood-based biomarkers are easy to implement in clinical practice, as well as in research studies on large cohorts and when repeated sample collection is part of the study design, since blood is easy to collect, process and store. Blood-based biomarkers may provide researchers and physicians with tests to screen out those individuals who are very unlikely to have Alzheimer's disease (AD) pathology and, at the same time, increase the likelihood that subjects who are included for further evaluation with expensive or more invasive tests (e.g., positron emission tomography [PET] imaging or cerebrospinal fluid [CSF] assessment) actually have AD pathology. This, in turn, would represent a pivotal shift to proper multi-step diagnostic work-up, saving time and resources. From academic settings to pharmaceutical companies, blood-based biomarkers may allow for new clinical trial designs. Moreover, blood-based biomarkers may enter clinical laboratory practice to allow for a standardized stratification by risk of AD for preventive interventions once the latter will be available

Within multiple clinical contexts of use (COUs), blood-based biomarkers are already transforming the clinical practice in advanced fields of medicine, such as oncology and immunology. Indeed, biomarkers can support the identification of subjects likely to respond to a given drug (predictive outcome measures) and to experience adverse drug effects (safety outcome measures).

During the last few years, promising results have been obtained regarding the ability of plasma amyloid β (Aβ) to detect cerebral β-amyloidosis (Janelidze, Bateman, Nakamura and Nabers). The main findings of these + what Insight is adding.

A Professional Interest Area focused on Blood Based Biomarkers (BBB-PIA) − integrated in the Alzheimer's Association's International Society to Advance Alzheimer's Research and Treatment (ISTAART) − has been created to support the harmonization process of preanalytical and analytical protocols and emphasize the need for a biorepository of clinical reference samples enabling the assessment of clinical performances

Brain accumulation of Aβ and Aβ-induced neuronal and glial proteotoxicity is a prominent feature within the complex biological landscape of AD along its pathophysiological *continuum*.

Consequently, blood-based surrogate markers of cerebral amyloidosis have been increasingly pushed towards validation and establishment for several consolidated COUs.

It is well documented that cerebral Aβ enters the peripheral circulation through the blood brain barrier (BBB). Therefore, the peripheral (i.e. plasma) concentration of Aβ peptides reflects the brain deposition of Aβ in both physiological and pathophysiological conditions. Indeed, it has already been shown that Aβ metabolism undergoes significant changes over time in physiological conditions, i.e. both in ageing and diseases.

The temporal dynamics of plasma have a time-dependent fashion similar to that of CSF; indeed, longitudinal studies reported that plasma Aβ alterations over time show a peak during the early stages of AD and a plateau, or even a slight decrease, during the disease progression.

A growing body of clinical investigation have assessed the strength of the association between plasma Aβ and Aβ PET imaging (or CSF) using both a dichotomized status (positivity or negativity to Aβ) and quantitative measures, such as the standardized uptake value ratio (SUVR). Interestingly, Nakamura and colleagues reported that the $A\beta_{1-40}/A\beta_{1-42}$ ratio, assessed with mass spectrometry, displays an accuracy of approximately 90% when using Aβ PET imaging as a standard of truth in two pooled cohorts (either combined or alone) of cognitively normal, prodromal (mild cognitive impairment [MCI]), and mild dementia stage AD participants.

In the ESTHER cohort subjects clinically, but not biologically, diagnosed with AD were identified in a nested case control group in average 8 years before clinical onset throughout the assessment of plasma Aβ immune-infrared-assay.

Robust translational evidences have shown that disease-modifying therapies for neurodegenerative disorders, including AD, hare more likely to achieve efficacy if commenced in asymptomatic stages when the rate of neural networks is relatively limited and potentially restorable. In this context, it is essential investigating preclinical stages of AD where the biological mechanisms underlying the dynamic alterations of Aβ metabolism which arise from a complex cross-talk between neurons, immune cell, BBB components, and peripheral tissue, are more likely to be preserved.

The extent to which blood Aβ concentrations accurately reflect the neuroimaging-based evidence of extracellular aggregation and accumulation of Aβ in asymptomatic stages of AD is not fully elucidated yet. Such scientific background has motivated us to address this question through a validation study of plasma $A\beta_{1-42}$ and the ratio $A\beta_{1-40}/A\beta_{1-42}$ within a cohort (N = 300) of

individuals with subjective complaints of memory dysfunction (SMC), which is a condition at risk of developing AD. To this end, we investigated in the INSIGHT-preAD cohort the accuracy of $A\beta_{1-42}$ and of the ratio $A\beta_{1-40}/A\beta_{1-42}$ in predicting brain amyloidosis across three-time points and over a time period of 3 years follow-up.

The INSIGH-preAD cohort that is a standardized large-scale, observational, monocentric, university, expert-center cohort of cognitively intact individuals with SMC.

Our primary goal was to investigate whether the biological signature of brain amyloidosis may resist slight variations of the related surrogate marker levels in plasma. As secondary endpoints we aimed: I) at proving the superiority of the plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ on plasma $A\beta_{1-42}$ alone, II) the diagnostic performance of the plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ is not significantly different across sexes and *APOE* ε4 carrying status, III) showing that the predictive performance of plasma $A\beta$ is not influenced by neither age nor other plasma biomarkers reflecting distinctive pathophysiological mechanisms of AD.

Materials and methods

Study participants

We designed a mono-centric, cross-sectional and longitudinal study in the cohort of SMC recruited from the "INveStIGation of AlzHeimer's PredicTors in Subjective Memory Complainers" (INSIGHT-preAD) study, a French academic university-based cohort which is part of the Alzheimer Precision Medicine Initiative Cohort Program (APMI-CP). Participants were enrolled at the Institute of Memory and Alzheimer's disease (Institut de la Mémoire et de la Maladie d'Alzheimer, IM2A) at the Pitié-Salpêtrière University Hospital in Paris, France. The main goal of the INSIGHT-preAD study is to investigate the earliest preclinical stages of AD, including factors and biomarkers associated with clinical progression.

The INSIGHT-preAD study includes 318 cognitively normal Caucasian individuals, recruited from the community in the wider Paris area, France, aged 70 to 85, with SMC. The status of SMC was confirmed as follows: (i) participants gave an affirmative answer ("YES") to both questions: "Are you complaining about your memory?" and "Is it a regular complaint that has

lasted for more than 6 months?"; (ii) participants showed intact cognitive functions based on the Mini-Mental State Examination score (MMSE, ≥ 27), Clinical Dementia Rating scale (CDR = 0), and Free and Cued Selective Rating Test (FCSRT, total recall score ≥ 41). Aβ positron emission tomography (Aβ-PET) investigation was performed at baseline visit, as mandatory inclusion criterion. Thus, all subjects enrolled into the study have SMC and are stratified as either positive or negative for cerebral Aβ deposition. Time series data for plasma levels of A$\beta_{1\text{-}42}$ and A$\beta_{1\text{-}40}$ were collected at three-time points (within a 2-year follow-up) using an in house Single molecule array assay (see below). Additionally, four novel candidate surrogate markers of AD-related processes were longitudinally assessed in plasma, including YKL40, NFL, Tau, BACE1. At the point of study inclusion, several data are collected such as, demographic and clinical data, and Apolipoprotein E (*APOE*) genotype as well. Exclusion criteria included a history of neurological or psychiatric diseases. The study was conducted in accordance with the tenets of the Declaration of Helsinki of 1975 and approved by the local Institutional Review Board at the participating center. All participants or their representatives gave written informed consent for use of their clinical data for research purposes.

Blood sampling

Immunoassays for plasma biomarkers

All analyses for plasma Aβ42 and Aβ40 T-tau, NFL, YKL-40 were performed at the Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Sweden. The measurements of each biomarker were performed in one round of experiments, using the same batch of reagents, by board-certified laboratory technicians who were blinded to the clinical data.

Plasma Aβ42 and Aβ40 were analyzed using Simoa immunoassay (Quanterix, Lexington, MA, USA). For Aβ42, repeatability was 4.1% and intermediate precision 7.0% for an internal QC plasma sample with a concentration of 10.5 pg/mL, while for Aβ40, repeatability was 4.0% and intermediate precision 6.4% for an internal QC plasma sample with a concentration of 203 pg/mL.

Plasma T-tau was measured using the Human Total Tau 2.0 kit on the Simoa platform (Quanterix, Lexington, MA). For plasma T-tau, both repeatability and intermediate precision was

12.2% for an internal QC plasma sample with a concentration of 1.9 pg/mL. Plasma NFL concentrations were measured using an in-house Simoa method, as previously described in detail (Kuhle 2016). Repeatability was 9.6% and 10.6% and intermediate precision 14.6% and 11.6%, for two internal QC plasma samples with concentrations of 12.9 pg/mL and 107 pg/mL. Plasma YKL-40 was analyzed using a commercial available ELISA kit (R&D Systems, Minneapolis, MN, US), according to manufacturer instructions. Repeatability was 6.6% and 6.9% and intermediate precision 6.6% and 6.9%, for two internal QC plasma samples with concentrations of 14.100 pg/mL and 108.000 pg/mL.

Plasma BACE1 levels were measured using a research prototype ELISA, based on the commercially available ELISA for CSF measurements (EQ 6541-9601-L; Euroimmun AG, Lübeck, Germany). The design of the CSF ELISA was previously described. In summary, ADx401 (clone 5G7) coated plates are incubated simultaneously with sample (15µL; undiluted) and biotinylated detector mAb ADx402 (clone 10B8F1), during 3h at room temperature. For plasma measurements, the same protocol as for CSF analysis, as instructed on the kit insert, was followed. Additionally, the same material of the CSF kit was used, including the lyophilized, ready-to-use calibrators and run validation controls. The only modification involved the buffer of the biotinylated detector mAb, which was diluted in a buffer adapted for the plasma matrix, including a heterophilic blocker reagent. After analysis, BACE1 concentrations were calculated via intrapolation (5PL curve fit; log (X)) based on the calibrator curve. In parallel to the clinical plasma samples, which were blinded and randomized before testing, two reference samples from ADx were analyzed, by means of run validation. Intra-assay precision of this plasma research prototype was on average 2.1%CV and 3.2%CV (coefficient of variation), based on the two reference samples run in duplicate and in 10 runs. The inter-assay variability was 8.5%CV and 9.5%CV.

PET data acquisition and processing

All florbetapir-PET scans are acquired in a single session on a Philips Gemini GXL CT-PET scanner 50 (± 5) minutes after injection of approximately 370 MBq (333-407 MBq) of Florbetapir. PET acquisition consists of 3 x 5 minutes frames, a 128 x 128 acquisition matrix and a voxel size of 2 x 2 x 2 mm³. Images are then reconstructed using iterative LOR-RAMLA algorithm (10 iterations), with a smooth post-reconstruction filter. All corrections (attenuation,

scatter, and random coincidence) are integrated in the reconstruction. Lastly, frames are realigned, averaged and quality-checked by the CATI team. CATI is a French neuroimaging platform funded by the French Plan Alzheimer (available at http://cati-neuroimaging.com).

Reconstructed PET images are analyzed with a pipeline developed by CATI. A standard uptake value ratio (SUVR) with a threshold of 0.7918 has been used to categorize our population in $A\beta$ positive or $A\beta$ negative according to a method previously described. In addition, 12 brain region of interest (ROI) were extracted according CATI platform (cingulum posterior right and left, cingulum anterior right and left, frontal superior right and left, parietal inferior right and left, precuneus right and left, temporal mid right and left).

Statistical modeling

Comparison, univariate and diagnostic performance analysis at inclusion of study participants. Chi-square test was used to study categorical variables in the sample stratified for $A\beta$-PET burden. Independent sample T-tests (corrected for multiple comparisons) were carried out to compare continuous variables (clinical demographic data) in the sample stratified for brain amyloidosis. The samples size is large enough for the T-test to be valid even all variables are not normally distributed.

Univariate analyses were carried out to adjust comparisons of continuous variables (plasma $A\beta_{1-42}$, ratio $A\beta_{1-40}/A\beta_{1-42}$) across sex, *APOE* ε4 carrier status and $A\beta$-PET status. The evaluation of the performance of plasma $A\beta_{1-42}$ levels and of the ratio $A\beta_{1-40}/A\beta_{1-42}$ in predicting the individual $A\beta$-PET status was carried out by using Receiver Operating Characteristic curve (ROC) analysis. A confidence interval of 95% was chosen. The area under the curve (AUC), and the representative best values for the sensitivity (sensitivity*100), specificity (1-specificty*100) were used for evaluating the performance of the model. AUCs (plasma $A\beta_{1-42}$ and ratio $A\beta_{1-40}/A\beta_{1-42}$) were compared with each other (DeLong test). For each ROC, the best cut-off point was determined by Youden's index (YI), which optimizes biomarker performance when equal weight is given to sensitivity and specificity and represents the likelihood of a positive test result in subjects with the condition versus those without the condition. We elected to not apply any correction for random measurements error at our ROCs as we developed a rigid protocol thus limiting intra-assay variability To test the association between plasma biomarkers and global /

regional Aβ-PET SUVRs, at baseline, we used the Pearson's correlation coefficient (*r*). All significant *r* were adjusted for age, sex and *APOE* genotype (ε4 carriers versus non-carriers). All tests were two-tailed and p values <0.05 were considered statistically significant. The Bonferroni correction was applied to adjust the false discovery rate (FDR) for multiple comparisons. Data analysis was performed using SPSS 23.0$^{©}$ IBM software for MacOS.

Analysis of prediction performance for unseen time-series data. 264 subjects had no missing values at time-point 1, 212 subjects and 126 subjects had no missing value at time-point 2 and time-point 3, respectively. Only 66 subjects had no missing values for the three time-points and there was no significant difference in the distribution of plasma $A\beta_{1-42}$ and ratio $A\beta_{1-40}/A\beta_{1-42}$ between time-points. To test whether the brain amyloidosis status might be better explained by taking into account several plasma biomarkers reflecting distinctive pathophysiological mechanisms of AD (herein $A\beta_{1-42}$, $A\beta_{1-40}$, Tau, NFL, BACE1, YKL40) and whether the performance of such classifier(s) might be conserved across subjects and time-points, we applied a tree-based approach involving the use of random forest analysis (RFA) for variable selection and that of Classification and regression trees (CART) analysis for selecting optimal classifiers. Such analyzes were performed after over-sampling of time-point 1 data (i.e. to obtain a well-balanced training set) as well as on true data. To perform oversampling, we used a Synthetic Minority Over-Sampling Technique (SMOTE) as implemented in the R package DmwR (https://cran.rproject.org/web/packages/DMwR/index.html). The resulting data set is herein referred to as set SMOTE. These analyzes also involved the evaluation of predictive robustness in unseen data, that is a test set including 140 subjects at time-point 2 and 54 additional subjects at time-point 3, The corresponding data set is herein referred to as set test. RFA was performed by using the library randomForest for R (https://CRAN.R-project.org/package=randomForest). RFA included an evaluation of the predictive performance of the models where the numbers of predictors is sequentially reduced via a nested cross-validation procedure (N = XX?) with the function rfcv of the randomForest package. CART was performed using the library rpart for R (https://CRAN.R-project.org/package=rpart). Briefly, CART analysis was performed to accomplish an explicit model of the data -set containing the most important predictors resulting from RFA analysis. To test the influence of age as well as the concentrations of other AD-related pathophysiological mechanisms plasma candidate markers (NFL, T-Tau, YKL-40, and BACE1)

on the performance of the classifier, we used a logistic regression as implemented in R.

RESULTS

Univariate analysis identifies an association between plasma Aβ levels and brain amyloidosis at the time of subject inclusion

At time-point 1, 277 SMCs of the INSIGHT cohort had plasma Aβ1-40 and Aβ1-42 levels assessed, with a sex ratio (F/M) of 170/107. In this sample, 74 subjects had a positive Aβ-PET status while 54 subjects had at least a *APOE ε4* allele (see Table 1 for clinical and demographic baseline data). The two subgroups of subjects identified by Aβ-PET status (either positive or negative) differed for plasma $A\beta_{1-42}$ levels and ratio $A\beta_{1-40}/A\beta_{1-42}$. The Aβ-PET positive subjects showed lower plasma $A\beta_{1-42}$ levels (p< 0.001) and a smaller ratio $A\beta_{1-40}/A\beta_{1-42}$ (p = 0.004) compared to Aβ-PET-negative subjects (see Supplementary Figure 1A and B). We found a significant difference between *APOE ε4* carriers and non-carriers in terms of plasma $A\beta_{1-42}$ with the former showing lower concentrations of the marker (p: 0.013; see Supplementary Figure 2A and B) but not for the ratio $A\beta_{1-40}/A\beta_{1-42}$ concentrations. No sex differences were found in the total sample and in the two Aβ-PET status subgroups separately. No correlation (*r)* was found between $A\beta_{1-42}$ and the ratio $A\beta_{1-40}/A\beta_{1-42}$ levels (both alone and combined) and age. Baseline correlation analysis between global / regional Aβ-PET SUVR and plasma $A\beta_{1-42}$ showed a significant negative relationship in all test FDR: 0,0038) (Table 2). All results were controlled for age, sex first and even for *APOE* genotype to follow. When we carried out baseline diagnostic performance test, ROCs showed for $A\beta_{1-42}$ at the best cut-off point an AUC: 68.1 % with a sensitivity of 52.3% and a specificity of 79.7% (balanced accuracy of 66 %) while the ratio $A\beta_{1-40}/A\beta_{1-42}$ showed an AUC of 79.4% with sensitivity of 77.8% and a specificity of 71.2% (balanced accuracy of 74.5%), at the best cut-off point: 0.05550). The DeLong's test disclosed a statistically significant difference between the two ROCs (p = 0.006, see Figure 1).

RFA and CART analysis retain plasma Aβ ratio Aβ1-40/Aβ1-42 levels as the best predictive variable of brain amyloidosis

We tested whether plasma Aβ levels, in terms of Aβ$_{1-42}$, Aβ$_{1-40}$, the ratio Aβ$_{1-40}$/Aβ$_{1-42}$, might predict brain Aβ-PET status with a good diagnostic performance looking for the best classifier among the three. Moreover, we investigated whether such classifiers might hold true in unseen data. To this aim, we carried out RFA first and CART analysis to follow. These are two statistical approaches that have been successfully used for biomarkers-drive investigations in the field of ND, including AD. Lastly, we investigated the performance of the classifier across sexes and *APOE* ε4 carriers and not (see paragraph XX). For each subgroup, we considered only subjects with no missing value at time-point 1.

At RFA, the group with no missing value at time-point 1 showed an excess of Aβ-PET negative subjects, which might have led to week performance of the RFA. To overcome this limitation, we applied a Synthetic Minority Over-Sampling Technique (SMOTE) before performing RFA. SMOTE generated a data set in which the Aβ-PET negative subjects are under-sampled (retaining 136/195 true subjects) and Aβ-PET-positive subjects are over-sampled (adding 68 virtual subjects). RFA after over-sampling the dataset displayed that the bet variable that best explains cerebral Aβ deposition is plasma ratio Aβ$_{1-40}$/Aβ$_{1-42}$ whether the model (training set) show for the latter 93% specificity and 85% sensitivity, with a balanced accuracy of 89% (Figure 2A left panel). Of note, the predictive performance is preserved in the test running through the original data set (*i.e.* true subjects and biomarkers data for time-point 1 and 2) that proven a balanced accuracy of 76% (Figure 2A right panel).

Next, we used CART for analyzing both the over-sampled and original data training sets, using the two best explanatory variables (plasma Aβ$_{1-42}$ and ratio Aβ$_{1-40}$/Aβ$_{1-42}$ levels) retained by RFA. In the over-sampled data set, the resulting decision tree only retains the ratio Aβ$_{1-40}$/Aβ$_{1-42}$ (Figure 2B, left panel). In this sense, the value of ratio Aβ$_{1-40}$/Aβ$_{1-42}$ keep discriminating Aβ-PET-positive and Aβ-PET negative subjects with the same performance rate found when carrying out the ROC-AUC analysis (Figure 2B, see also Figure 1). The sensitivity was 85% and specificity was 78%, with a balanced accuracy of 81% (Figure 2B, two middle panels). Such performances are well preserved in the test set, showing a sensitivity of 68%, a specificity of 75%, and, a balanced accuracy of 71% (Figure 2B, right panel). After running CART throughout the original data (true subjects), the final output did not change compared to the over-sampled model since the resulting decision tree retained both the ratio Aβ$_{1-40}$/Aβ$_{1-42}$ and Aβ$_{1-42}$ levels (Figure 2C, left panel). The sensitivity was 85% and specificity was 78%, for a balanced accuracy of 81% (Figure 2C middle

panels). These performances are well preserved in the test set as sensitivity was 60% and specificity was 83%, with a balanced accuracy of 71% (Figure 2C, right panel).

Performance of the classifier across sexes and APOE subgroups

We carried out RFA and CART analysis across sexes and *APOE* ε4 carriers and not. For each subgroup, we considered only subjects with no missing value at time-point 1.

For women, the most important variables emerging from the RFA were the ratio $A\beta_{1-40}/A\beta_{1-42}$ and $A\beta_{1-42}$ (Supplementary Figure 3A). The CART analysis on the over-sampled data set (training set) provided a simple decision tree very close to the one resulting from the total sample (Supplementary Figure 3B) with a performance of the classifiers in predicting the Aβ-PET status in the training set displaying a good level (Supplementary Figure 3B decision tree diagram and left panel). However, this is not the case of the replication set (true subjects Supplementary Figure 3B right panel).

For men, the most important variables resulting from the RFA were the ratio $A\beta_{1-40}/A\beta_{1-42}$ and $A\beta_{1-42}$ (Supplementary Figure 3C decision tree diagram and left panel). The CART analysis on the over-sampled data set (training set) provided a simple decision tree very close to the one resulting from the total sample but with a different performance of the classifiers in predicting the Aβ-PET status in both the training and replication set displaying a good level (Supplementary Figure 3D decision tree diagram and right panel).

When looking at the classifier build with the whole population, the emerging differences is that there is two more false negative for the men of the replication set when with the classifier build with men only, there is two more false positive. It appeared that, in the test set, 31% of the women and 21% of the men are false positive ore false negative

When carrying out a confusion matrix for the combination of the CART models taking into account the sex for the test set (i.e. true subjects, time-point 2 and 3 dataset), we disclosed that the performance of the classifier was gender-independent. The confusion matrix provided a specificity of 77%, sensitivity of 64%, and a balanced accuracy of 70%.

When considering only *APOE* ε4 carrying status at time-point 1, we obtained a balanced data set between Aβ-PET+ and Aβ-PET- individuals. Thus, we did not perform oversampling within this

subgroup. For *APOE* ε4 carriers, the most important variables are still the ratio $A\beta_{1-40}/A\beta_{1-42}$ and $A\beta_{1-42}$ (Supplementary Figure 4A). The CART analysis provided a simple decision tree very close to thus obtained in the total sample, see (Supplementary Figure 4B decision tree diagram and left panel). The performance of the classifier in the training set is good as shown in Supplementary Figure 4B left panel), and it is still good to predict the Aβ-PET status of the subjects of the replication set decrease (Supplementary Figure 4B right panel).

In the *APOE* ε4 non-carriers subgroup, the dataset was very unbalanced in terms of Aβ-PET+ and Aβ-PET- individuals. Therefore, we performed the oversampling procedure. The most important variables resulting from RFA were the ratio $A\beta_{1-40}/A\beta_{1-42}$ and $A\beta_{1-42}$ (Supplementary Figure 4C). The CART analysis on the over-sampled dataset provided a more complicated decision tree (Supplementary Figure 4D, decision tree diagram and left panel). The performance of the classifier in the training set are good as shown Supplementary Figure 4D and left panel, and the performance to predict the Aβ-PET status of the subjects of the replication set remain good Supplementary Figure 4D right panel).

For the *APOE* ε4 non-carriers subgroup, it was not possible to build a model with only the true data at time-point 1.

As next step, we tested whether the classifier may perform better in *APOE* ε4 carriers versus non-carriers. For that purpose, we compared the performance of the classifier (Figure 2B, left panel) in the population carrying at least one *APOE* ε4 alleles and the population without *APOE* ε4 alleles in the replication set. We used a chi-square test as implemented in R to verify if the number of misclassified subjects (false positive or false negative) varied in an *APOE* ε4-dependent manner. We observed no significant difference between these two groups in terms of false positives/negatives. We performed RFA and CART analysis for the *APOE* ε4 positive and negative groups separately. The predictive features and performance to identify the Aβ-PET status of the subjects who carried at least one *APOE* ε4 allele were similar to those obtained with the model build for the entire population. Indeed, when we built a confusion matrix for the combination of the CART models taking into account the *APOE* status for the test set (*i.e.*, true subjects, time-point 2 and 3 dataset) we disclosed a specificity of 68%, sensitivity of 72%, and a balanced accuracy of 70%.

On the other hand, when we compared the performance of the classifier built with the whole population (Figure 1B left panel) and the classifier built with only *APOE* ε4 non-carriers at time

point one (Supplementary Figure 4D) to separate Aβ-PET positive and negative subjects without *APOE* ε4 alleles in the replication data-set, we observed that the performance of the classifier built with only *APOE* ε4 non-carriers collapsed much more than the performance of the classifier built with the whole population.

Together, these results suggest that distinguishing *APOE* ε4 carriers and no-carriers subjects does not improve the accuracy of CART classifiers (ratio $A\beta_{1-40}/A\beta_{1-42}$ and $A\beta_{1-42}$) for predicting brain amyloidosis.

Age effect and other markers

We tested whether the age or the concentrations in plasma of the major candidate biomarkers of AD-related pathophysiological mechanisms (i.e. T-Tau, NFL, BACE1, YKL40) may influence the performance of the classifier (subjects are either well classified as true positive or true negative or misclassified as false positive or false negative). To follow, we carried out a logistic regression which showed no significant association between age, the level of YKL40, NFL and T-Tau (p-value > 0.1). However, we disclosed an association between the concentrations of plasma BACE1 and the performance of the classifier (p-value = 0.09). Nevertheless, such a significance is likely to be driven by few exceptional values. To test whether the influence of plasma BACE1 is due to this few extremes value, we carried the analysis for only subject with plasma BACE1 level between 900 and 1300. Thus, we identified a dataset of 192 subjects at time-point 1 among who 46 individuals had a Aβ-PET positive status while 146 individuals had a Aβ-PET negative status. In the replication set (time-point 2 and 3) 145 subjects were identified among who 38 individuals had a Aβ-PET positive status while 107 had a Aβ-PET negative status. The classifier built with this data at time-point 1 was the same as the model built with the overall study population displaying a overlapping performance with the performance of this last classifier being not significantly influenced by BACE1 plasma level (p-value = 0.5).

Together, these results suggest that, in a larger population a model more accurate can be built for subjects with extreme (low or high) level of BACE1.

Discussion

There is an urgent need for biomarker-guided investigations carried out on asymptomatic individuals at risk for AD aimed initially, at expanding on the insights about the etiology and pathophysiology of AD, then at coping with the therapeutic challenges in the field.

There are not currently available data in literature about the predictive value of plasma Aβ for Aβ-PET status in cognitively intact individuals with SMC which is a condition at risk of AD. Moreover, there are not validation studies providing that there is no time-effect on the predictive value of plasma Aβ of cerebral amyloidosis. This is the first validation study of plasma Aβ ever conducted to explain brain Aβ dynamics in the INSIGH-preAD cohort that is a standardized large-scale, observational, monocentric, university, expert-center cohort of cognitively intact individuals with SMC.

Our primary goal was to investigate whether the biological signature of brain amyloidosis may resist to slight variations of the related surrogate marker levels in plasma. As secondary endpoints we aimed: I) at proving the superiority of the plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ on plasma $A\beta_{1-42}$ alone, II) the diagnostic performance of the plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ is not significantly different across sexes and *APOE* ε4 carrying status, III) showing that the predictive performance of plasma Aβ is not influenced by neither age nor other plasma biomarkers reflecting distinctive pathophysiological mechanisms of AD.

Consistent with our expected results, plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ concentrations showed good diagnostic performance in predicting cerebral deposition of Aβ assessed throughout the standard of truth we have selected, i.e. the Aβ-PET status (see Figure 1 and Figure 2A).

We initially carried out RFA and then CART analysis which final output suggests that plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ levels is the best predictor of cerebral Aβ deposition, even over multiple time points and regardless of *APOE* ε4 carrier status and sex.

We outline that plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ yielded 93% specificity, 85% sensitivity and 89% balanced accuracy when seeking predictive value of all markers setting Aβ-PET status as outcome variable. Such a good performance survives analysis when taking subjects at different time-points (see Figure 2A,B,C).

From a practical point of view, we found that, in a cohort of individuals at risk of AD, it was sufficient to assess two feasible plasma markers and the related ratio (plasma ratio $A\beta_{1-40}/A\beta_{1-42}$)

to reach a predictive test with an accuracy as good such to optimize the frequency of assessment of time-energy consuming tests as Aβ-PET scanning. Plasma $A\beta_{1-40}/A\beta_{1-42}$ ratio thus showed a robust performance of the prediction accounting for future multiple and long-term perspectives of its utilization in several COUs.

For instance, our finding is absolutely in line with the real-word perspective of blood-based biomarkers since it strongly supports that plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ would most likely play a role in initial screening to rule-out individuals that do not deserve Aβ-PET scanning rather than in confirmatory diagnosis cases.

We found that neither sex nor *APOE* ε4 carrier status influenced the performance of the classifier plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ (see Supplementary Figures 3A,B,C,D and 4A,B,C,D), suggesting that sex and *APOE* ε4 allele do not represent determinants for the validation of plasma Aβ as key biomarker in several COU for AD. Similar results have been obtained for CSF Aβ42 (PMID: 25162367). Thus, in the real world, there will not be the necessity to stratify subjects by the presence or not of *APOE* ε4 allele before assessing plasma Aβ markers. This finding has a huge and tremendous consequence on all potential COUs within which the marker may be utilized, especially concerning the therapeutic perspectives.

Our study shows that the diagnostic performance of plasma ratio $A\beta_{1-40}/A\beta_{1-42}$ is independent of the age of individuals as well as the concentrations in plasma of novel candidate surrogate markers, such as NFL and T-Tau that are supposed to chart neurodegeneration as well as YKL-40 which is the best candidate to *in-vivo* track neuroinflammation.

This, in turn, outlines a major biological point that the prediction of brain amyloidosis through a reliable blood-based window is not influenced by concurrent aberrant molecule pathways as those underlying neuroinflammation and neurodegeneration which are critical pathomechanistic alterations common to several brain proteinopathies.

Such finding confers a higher value to the performance of the classifier which can survive the real-word neurobiological heterogeneity of large-scale population settings.

Thus, we have accomplished a predictive model, robust to a variation of the condition (time-related fluctuations of plasma ratio $A\beta_{1-40}/A\beta_{1-42}$) in the measure of the predicative variable cerebral amyloidosis with Aβ-PET.

In our model the Aβ-PET load does not explain the levels of the plasma Aβ marker and its alterations over time. This finding is substantially consistent with the notion that orthogonal

factors (e.g. peripheral metabolism, blood brain barrier age-related changes, brain cellular and subcellular pathways) could all influence the dynamics of Aβ as well as other surrogate markers, in blood. Consequently, we propose plasma Aβ assessed using Simoa immunoassay as reliable and feasible blood-based signature indicative of AD-related amyloid pathophysiology.

FIGURE TITLES AND FOOTNOTES

Supplementary Figure 1A,B. Comparisons of subjects identified by Aβ-PET status (either positive or negative) for plasma concentrations of Aβ$_{1-42}$ and ratio Aβ$_{1-40}$/Aβ$_{1-42}$.
(A) Positive Aβ-PET scan individuals showed lower plasma Aβ$_{1-42}$ levels (p< 0.001)
(B) Positive Aβ-PET scan individuals showed smaller ratio Aβ$_{1-40}$/Aβ$_{1-42}$ (p< 0.001)

Supplementary Figure 2A,B. Comparisons of subjects identified by *APOE* ε4 carrier status (either carrier or non-carrier) for plasma concentrations of Aβ$_{1-42}$ and ratio Aβ$_{1-40}$/Aβ$_{1-42}$.
(A) Individuals carrying at least one ε4 allele showed lower plasma Aβ$_{1-42}$ levels (p< 0.0013)
(B) There were not statistical differences according to APOE ε4 carrier status for plasma concentrations of ratio Aβ$_{1-40}$/Aβ$_{1-42}$.

Figure 1. Receiver Operating Characteristic curve (ROC) analysis for the evaluation of the performance of plasma Aβ$_{1-42}$ levels and of the plasma Aβ$_{1-40}$/Aβ$_{1-42}$ ratio in predicting the individual Aβ-PET status
When we carried out baseline diagnostic performance test, ROCs showed for Aβ$_{1-42}$ at the best cut-off point a AUC: 68.1 % with a sensitivity of 52.3% and a specificity of 79.7% (balanced accuracy of 66 %) while the ratio Aβ$_{1-40}$/Aβ$_{1-42}$ showed an AUC of 79,4% with sensitivity of 77,8% and a specificity of 71,2% (balanced accuracy of 74.5%), at the best cut-off point: 0.05550). The DeLong's test disclosed a statistically significant difference between the two ROCs (p: 0,006)

Figure 2A,B,C. RFA and CART signatures retain plasma A$\beta_{1\text{-}40}$/A$\beta_{1\text{-}42}$ ratio as a predictive classifier of the brain amyloid status in the INSIGHT-pre AD cohort.

(**A**) RFA analysis identifies the ratio A$\beta_{1\text{-}40}$/A$\beta_{1\text{-}42}$ and the level of A$\beta_{1\text{-}42}$ as the sole explanatory variables explaining the amyloidosis status (i.e. positive, negative). The performance of the model are preserved between the training set (build from over-sampled data at time-point 1) and the test set (build using unseen data from true subjects at time-point 2, and time-point 3).

(**B**) CART analysis identifies the ratio A$\beta_{1\text{-}40}$/A$\beta_{1\text{-}42}$ as the only predictive variable. The threshold selected for the classification is the same as the one retained via the ROC-AUC analysis (see Figure 1). The performance of the model is preserved between the training set (build from over-sampled data at time-point 1) and the test set (same test set as in **A**).

(**C**) CART analysis identifies the ratio A$\beta_{1\text{-}40}$/A$\beta_{1\text{-}42}$ and A$\beta_{1\text{-}42}$ levels as predictive variables in true subjects. The performance of the model is preserved between the training set (build from data at time-point 1) and the test set (same test set as in **A**).

.

Supplementary Figure 3 A,B,C,D

Supplementary Figure 4 A,B,C,D