Accurate risk estimation of β-amyloid positivity to identify prodromal Alzheimer’s disease: Cross-validation study of practical algorithms

Sebastian Palmqvist\textsuperscript{a,b,*}, Philip S. Insel\textsuperscript{a}, Henrik Zetterberg\textsuperscript{c,d,e,f}, Kaj Blennow\textsuperscript{c,d}, Britta Brix\textsuperscript{e}, Erik Stomrud\textsuperscript{a,h}, the Alzheimer’s Disease Neuroimaging Initiative\textsuperscript{1}, the Swedish BioFINDER study, Niklas Mattsson\textsuperscript{a,b}, Oskar Hansson\textsuperscript{a,h,**}

\textsuperscript{a}Clinical Memory Research Unit, Department of Clinical Sciences, Lund University, Malmö, Sweden
\textsuperscript{b}Department of Neurology, Skåne University Hospital, Lund, Sweden
\textsuperscript{c}Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden
\textsuperscript{d}Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden
\textsuperscript{e}Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, United Kingdom
\textsuperscript{f}UK Dementia Research Institute at UCL, London, United Kingdom
\textsuperscript{g}Euroimmun AG, Lübeck, Germany
\textsuperscript{h}Memory Clinic, Skåne University Hospital, Malmö, Sweden

Abstract

Introduction: The aim was to create readily available algorithms that estimate the individual risk of β-amyloid (Aβ) positivity.

Methods: The algorithms were tested in BioFINDER (n = 391, subjective cognitive decline or mild cognitive impairment) and validated in Alzheimer’s Disease Neuroimaging Initiative (n = 661, subjective cognitive decline or mild cognitive impairment). The examined predictors of Aβ status were demographics; cognitive tests; white matter lesions; apolipoprotein E (APOE); and plasma Aβ42/Aβ40, tau, and neurofilament light.

Results: Aβ status was accurately estimated in BioFINDER using age, 10-word delayed recall or Mini–Mental State Examination, and APOE (area under the receiver operating characteristics curve = 0.81 [0.77–0.85] to 0.83 [0.79–0.87]). When validated, the models performed almost identical in Alzheimer’s Disease Neuroimaging Initiative (area under the receiver operating characteristics curve = 0.80–0.82) and within different age, subjective cognitive decline, and mild cognitive impairment populations. Plasma Aβ42/Aβ40 improved the models slightly.

Discussion: The algorithms are implemented on http://amyloidrisk.com where the individual probability of being Aβ positive can be calculated. This is useful in the workup of prodromal Alzheimer’s disease and can reduce the number needed to screen in Alzheimer’s disease trials.

Keywords: Alzheimer’s disease; β-amyloid; Prediction; Diagnostic accuracy; Cerebrospinal fluid; Aβ42; Risk estimation; Position emission tomography; Plasma Aβ42/Aβ40

\textsuperscript{1}The data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

\textsuperscript{*}Corresponding author. Tel.: +4646177808.

\textsuperscript{**}Corresponding author. Tel.: +4640335435.

E-mail addresses: sebastian.palmqvist@med.lu.se (S.P.), oskar.hansson@med.lu.se (O.H.)

https://doi.org/10.1016/j.jalz.2018.08.014

1552-5260/© 2018 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer’s Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Introduction

β-Amyloid (A\textsubscript{β}) accumulation is believed to be the initial pathology of the most common type of neurological disease leading to dementia, Alzheimer’s disease (AD) [1]. Abnormal levels of A\textsubscript{β} are associated with longitudinal cognitive decline in healthy elderly [2] and progression to AD dementia in subjects with mild cognitive impairment (MCI) [3]. A verified A\textsubscript{β} status can be used to improve the accuracy of AD diagnostics and for including participants in trials of novel AD drugs, as currently used in several clinical trials [4]. Given the devastating symptoms of AD, the high number of affected people, and the tremendous costs for society (US$ 259 billion per year for dementia in the US alone), there will be a great pressure on the health care system to identify persons with abnormal A\textsubscript{β} deposition when disease-modifying AD treatments become available [5].

Brain A\textsubscript{β} can be detected in vivo either by performing a lumbar puncture (LP) and analyzing the levels of the peptide A\textsubscript{β}42 in cerebrospinal fluid (CSF) or by performing a positron emission tomography (PET) scan using a ligand that binds to A\textsubscript{β} fibrils (A\textsubscript{β} PET). There are no significant differences between the two methods in terms of accuracy for identifying AD [6,7], and they are used mostly not only in research but also in clinical practice at some specialized memory clinics. However, because these methods are invasive, costly, and not available in all health care settings, a screening process to select individuals for LP or PET testing, both in clinical practice and clinical treatment trials, would be very useful. Several studies on amyloid prediction tools or blood-based A\textsubscript{β} biomarkers exist, but due to lack of or failed validations, low accuracies, or the usage of advanced technology or extensive neuropsychological testing, none of them are currently being used in clinical or research settings, to the best of our knowledge [8–12].

In the present study, we aimed to develop algorithms that estimate the risk of being A\textsubscript{β} positive using readily available and noninvasive measures and tests. Nondemented subjects with either subjective or objective cognitive symptoms were examined to provide a clinically relevant target population. The models were developed in a training cohort and validated in an independent cohort. In a second step, we analyzed the added value of including the plasma biomarkers tau, neurofilament light (NFL), and the A\textsubscript{β}42/A\textsubscript{β}40 ratio.

2. Materials and methods

2.1. Participants of the training cohort (BioFINDER)

The Swedish BioFINDER study (Biomarkers For Identifying Neurodegenerative Disorders Early and Reliably) is a prospective study that focuses on identifying key mechanisms and improving clinical diagnostics of AD and other neurodegenerative disorders. Details about the Swedish BioFINDER study design have been published previously [12,13] and are available at http://biofinder.se. In the present study, we used the BioFINDER cohort of prospectively and consecutively included nondemented participants with cognitive complaints. They were enrolled between 2010 and 2015, mostly from primary care centers in the Southern part of Sweden. The inclusion/exclusion criteria are provided in the Supplementary Material. Based on the result of a comprehensive neuropsychological battery and the clinical assessment of a senior neuropsychologist and two physicians specialized in neurocognitive disorders, 54% of the 391 participants were classified as having MCI and 46% as having subjective cognitive decline [14].

2.2. Amyloid outcome measures in BioFINDER

A\textsubscript{β} was measured using 18\textsuperscript{F}-flutemetamol PET if available (n = 241), otherwise CSF A\textsubscript{β}42 was used (n = 150). The scanning [15] and processing [13] procedures have been described previously. The weighted mean standardized uptake value ratio (SUVR) from a global neocortical region of interest [16] relative to a composite reference region (white matter, cerebellum and brainstem [13]) was used to determine the A\textsubscript{β} status. The SUVR cutoff for A\textsubscript{β} positivity was determined using unbiased mixture modeling statistics, which is a well-validated method for determining such a cutoff [13,17,18]. The resulting cutoff for A\textsubscript{β} positivity was >0.738 SUVR.

LP and CSF handling followed a structured protocol [15]. CSF levels of A\textsubscript{β}42 were analyzed using INNOTEST ELISAs (Fujirebio Europe, Ghent, Belgium). The CSF A\textsubscript{β}42 cutoff for A\textsubscript{β} abnormality was determined using the optimized Youden’s Index against A\textsubscript{β} PET in BioFINDER (CSF A\textsubscript{β}42 < 552 ng/L; sensitivity 93%, specificity 84%).

2.3. Predictor variables of A\textsubscript{β} positivity

Different types of predictors were examined in the primary analysis, including demographics (age, education, and sex), apolipoprotein E (APOE) genotype, cognitive test scores, and white matter lesions. The cognitive tests were administered by experienced research nurses who were blinded to the A\textsubscript{β} status of the participants.

APOE genotypes were analyzed from blood samples, and the participants were stratified according to A\textsubscript{β} risk into the following groups (see reference [19] for rationale): (1) ε2/ε2 or ε2/ε3, (2) ε3/ε3, (3) ε2/ε4 or ε3/ε4, and (4) ε4/ε4. APOE ε3/ε3 was the reference category.

Episodic memory function was measured with the delayed recall part of the 10-word list from the Alzheimer’s Disease Assessment Scale–cognition [20]. Cognitive function was also assessed with the Mini–Mental State Examination (MMSE) [21]. Both the total score and the score from the orientation and memory parts of the test were used. The scores from the orientation and memory parts of the MMSE were used based on previous findings showing that
the orientation to time and place and the three-word delayed recall parts can differentiate MCI and dementia due to AD from other causes of cognitive impairment [22,23]. It consists of orientation to place (country, county/state, city, building/place, and floor), orientation to time (year, season, month, day of the week, and date), and three words that are being recalled after a short distraction task.

We also examined A Quick Test of Cognitive Speed (AQT)—color and form score, which is a sensitive test for attention and executive function to account for non-AD-specific cognitive impairment [24,25]. AQT was used alone and as a ratio with the delayed word recall test and MMSE orientation and memory.

Magnetic resonance imaging was performed on a 3-Tesla Siemens Tim Trio scanner (Siemens Medical Solutions, Erlangen, Germany). T2 FLAIR images were used for rating white matter lesions according to the ARWMC scale [26] to account for the impact of cerebrovascular pathology on cognitive impairment.

In a secondary analysis, we added the plasma biomarkers tau, the ratio of Aβ42/Aβ40 and NfL, which previously have been tested as AD biomarkers [27–29]. Plasma Aβ42 and Aβ40 levels were determined using the EUROIMMUN ELISAs (EUROIMMUN, Lubeck, Germany). The total levels of Aβ42 and Aβ40 were used to calculate the Aβ42/Aβ40 ratio. Plasma tau and NfL concentrations were measured on a Simoa HD-1 analyzer using the Human Total Tau kit (Quanterix, Lexington, MA) for tau and an in-house assay based on the same antibodies and standard protein as in the commercially available NF-light kit (UmanDiagnostics, Umeå, Sweden) for NfL [30]. All predictor variables were available in all patients, except for plasma NfL and tau (n = 346 of 391 participants).

2.4. Validation cohort—Alzheimer’s Disease Neuroimaging Initiative

A detailed study and data description of the Alzheimer’s Disease Neuroimaging Initiative (ADNI) as well as inclusion/exclusion criteria and MCI definitions can be found on www.adni-info.org and in the Supplementary Material. Only nondemented subjects with cognitive symptoms were selected, which included participants with early and late MCI and participants from the healthy control cohort who had significant memory concerns.

We included only participants with a complete data set of cognitive test, APOE, and Aβ data (Aβ PET or CSF Aβ42). This selection resulted in a population of 661 participants, of which 170 had plasma biomarker data.

Aβ status was based on (in order of preference) (1) Aβ PET using the ligand 11C-florbetapir, (2) Aβ PET using the ligand 18F-Pittsburgh Compound B (PiB), and (3) CSF Aβ42 measured using the multiplex xMAP luminex platform (Luminex Corp, Austin, TX, USA) with the INNOBIA AlzBio3 kit (Innogenetics, Ghent, Belgium) [31,32]. Predefined cutoffs for Aβ positivity were used for florbetapir (>1.11 SUVR) [33], 11C-Pittsburgh Compound B (>1.5 SUVR) [34], and Aβ42 (<192 ng/L) [32]. The methods for these three measures have previously been described [32–34].

Plasma Aβ42 and Aβ40 were measured using the INNOBIA plasma Aβ immunoassay kit (Fujirebio, Ghent, Belgium) on the Luminex 100 immunoassay platform (Luminex Corp) [35]. The total levels of Aβ42 and Aβ40 were used to calculate the Aβ42/Aβ40 ratio.

2.5. Statistical analysis

Group comparisons were done using the Mann-Whitney U test. In Table 1, we applied Bonferroni correction to adjust for multiple comparisons. P values were thus multiplied by 6 and a value of <0.05 was considered statistically significant. To predict Aβ positivity, the following variables from the training cohort (BioFINDER) were entered in a general linear model: age, gender, presence of APOE ε2/ε2 or ε2/ε3, presence of APOE ε2/ε4 or ε3/ε4, presence of APOE ε4/ε4 (APOE ε3/ε3 was not included because it was the reference variable), total MMSE score, the score from the orientation and delayed recall (memory) parts of the MMSE, the 10-word list delayed recall from Alzheimer’s Disease Assessment Scale–cognition (number of errors), years of education, AQT score, 10-word list delayed recall/AQT, MMSE orientation and memory/AQT, and degree of white matter lesions (ARWMC score). Using Aβ status as the dependent variable, the general linear model was fitted to the data using the least absolute shrinkage and selection operator (LASSO) [36]. The LASSO analysis uses a type of forward selection logistic regression that provides more robust predictors because it penalizes the absolute value of the coefficients and shrinks irrelevant coefficients to zero. The LASSO was only used for selecting predictor variables in BioFINDER (the training cohort), it could not be directly applied to the ADNI data (validation cohort) because not all BioFINDER variables were present in ADNI (ARWMC and AQT data). To increase the applicability of an Aβ risk model, we also used a reduced set of variables (but the same population) where we excluded the 10-word list delayed recall, AQT, and white matter lesions assessments because these measures are not always available in all settings. In a final step of Aβ risk analyses, we added plasma tau, plasma NfL, and the plasma Aβ42/Aβ40 ratio to the two LASSO models. The selected variables from the LASSO regression (variables with nonzero estimates) were entered in a logistic regression model to calculate the intercept, the coefficients, and the resulting area under the receiver operating characteristics curve (AUC). The Akaike Information Criterion (AIC) was used to assess the model fit in relation to its complexity (number of variables), where a drop of ≥2 indicated a statistically better model [37]. The best model was considered to be the one with the highest AUC and the lowest AIC. The logistic regression models from BioFINDER were then replicated in different settings.
Table 1
Characteristics of the training and validation cohorts

<table>
<thead>
<tr>
<th>Variables</th>
<th>BioFINDER (training cohort)</th>
<th>ADNI (validation cohort, plasma subset)</th>
<th>ADNI (validation cohort, total population)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A\beta^-$ &amp; $A\beta^+$ &amp; Total</td>
<td>$A\beta^-$ &amp; $A\beta^+$ &amp; Total</td>
<td>$A\beta^-$ &amp; $A\beta^+$ &amp; Total</td>
</tr>
<tr>
<td>N</td>
<td>197 (50%) &amp; 194 (50%) &amp; 391</td>
<td>66 (39%) &amp; 104 (61%) &amp; 170</td>
<td>311 (47%) &amp; 350 (53%) &amp; 661</td>
</tr>
<tr>
<td>SCD/MCI</td>
<td>55%/45% &amp; 36%/64% &amp; 54%</td>
<td>71.0 (56–89) &amp; 71.9 (57–83) &amp; 71.5 (56–89)</td>
<td>70.4 (7.4) &amp; 73.4 (6.9) &amp; 72.2 (55–91)</td>
</tr>
<tr>
<td>SMC/EMCI/LMCI</td>
<td>Age (range)</td>
<td>69.8 (60–80) &amp; 71.0 (60–80)</td>
<td>71.0 (56–89) &amp; 71.9 (57–83) &amp; 71.5 (56–89)</td>
</tr>
<tr>
<td>Sex (women)</td>
<td>49% &amp; 43% &amp; 46%</td>
<td>45% &amp; 44% &amp; 45%</td>
<td>45% &amp; 46% &amp; 302 (46%)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>12.1 (3.6) &amp; 11.4 (3.5) &amp; 11.8 (3.5)</td>
<td>16.4 (2.5) &amp; 16.2 (2.8) &amp; 16.3 (2.7)</td>
<td>16.4 (2.5) &amp; 16.2 (2.8) &amp; 16.3 (2.7)</td>
</tr>
<tr>
<td>MMSE (0–30 p)</td>
<td>28.2 (1.7) &amp; 27.4 (1.8) &amp; 27.8 (1.8)</td>
<td>28.6 (1.4) &amp; 27.6 (1.8) &amp; 28.0 (1.7)</td>
<td>28.6 (1.5) &amp; 27.7 (1.8) &amp; 28.1 (1.7)</td>
</tr>
<tr>
<td>MMSE orientation and delayed recall (0–13 p)</td>
<td>12.1 (1.0) &amp; 11.5 (1.4) &amp; 11.8 (1.2)</td>
<td>12.1 (0.9) &amp; 11.3 (1.5) &amp; 11.6 (1.3)</td>
<td>12.1 (1.1) &amp; 11.4 (1.5) &amp; 11.8 (1.4)</td>
</tr>
<tr>
<td>10-word list delayed recall (0–10 errors)</td>
<td>4.1 (2.5) &amp; 6.0 (2.5) &amp; 5.0 (2.6)</td>
<td>4.3 (2.1) &amp; 5.9 (2.6) &amp; 5.2 (2.5)</td>
<td>3.8 (2.3) &amp; 5.5 (2.7) &amp; 4.7 (2.6)</td>
</tr>
<tr>
<td>APOE ε2/ε2 or ε2/ε3</td>
<td>13% &amp; 7% &amp; 17% &amp; 2% &amp; 3% &amp; 13% &amp; 3% &amp; 8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOE ε3/ε3</td>
<td>63% &amp; 46%</td>
<td>59% &amp; 28% &amp; 40%</td>
<td>64% &amp; 33% &amp; 48%</td>
</tr>
<tr>
<td>APOE ε2/ε4 or ε3/ε4</td>
<td>22% &amp; 36%</td>
<td>20% &amp; 55% &amp; 41%</td>
<td>21% &amp; 51% &amp; 37%</td>
</tr>
<tr>
<td>APOE ε4/ε4</td>
<td>3% &amp; 11%</td>
<td>5% &amp; 16% &amp; 11%</td>
<td>2% &amp; 13% &amp; 8%</td>
</tr>
<tr>
<td>Plasma $A\beta_{42}/A\beta_{40}$ ratio</td>
<td>0.19 (0.06) &amp; 0.16 (0.03) &amp; 0.17 (0.05)</td>
<td>0.10 (0.05) &amp; 0.082 (0.05) &amp; 0.090 (0.05)</td>
<td>0.19 (0.06) &amp; 0.16 (0.03) &amp; 0.17 (0.05)</td>
</tr>
<tr>
<td>Plasma tau (pg/mL)</td>
<td>5.3 (2.3) &amp; 5.5 (2.7) &amp; 5.40 (2.5)</td>
<td>5.3 (2.3) &amp; 5.5 (2.7) &amp; 5.40 (2.5)</td>
<td></td>
</tr>
<tr>
<td>Plasma NfL (pg/mL)</td>
<td>24.0 (24) &amp; 26.7 (17) &amp; 25.4 (20.9)</td>
<td>24.0 (24) &amp; 26.7 (17) &amp; 25.4 (20.9)</td>
<td></td>
</tr>
<tr>
<td>WML (ARWMC scale, 0–27 p)</td>
<td>6.6 (5.7) &amp; 6.9 (5.4) &amp; 6.8 (5.6)</td>
<td>6.6 (5.7) &amp; 6.9 (5.4) &amp; 6.8 (5.6)</td>
<td></td>
</tr>
<tr>
<td>AQTL color-form (seconds)</td>
<td>79 (25) &amp; 85 (29) &amp; 82 (27)</td>
<td>79 (25) &amp; 85 (29) &amp; 82 (27)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: $A\beta$, β-amyloid; ADNI, Alzheimer’s Disease Neuroimaging Initiative; APOE, apolipoprotein E; ARWMC, age-related white matter changes; BioFINDER, Biomarkers For Identifying NeuroDegenerative Disorders Early and Reliably; EMCI, early MCI; LMCI, late MCI; MCI, mild cognitive impairment; MMSE, Mini–Mental State Examination; NfL, neurofilament light; SCD, subjective cognitive decline; SMC, significant memory concern; WML, white matter lesions.

NOTE. Data are given in mean values (standard deviation) if not otherwise specified. All $P$ values are Bonferroni corrected (multiplied by 6) to adjust for multiple comparisons. Within population comparisons (Aβ+ compared with Aβ−): $^{a}P < .05$; $^{b}P < .01$; $^{c}P < .001$. Comparison between ADNI and BioFINDER: $^{d}P < .05$; $^{e}P < .01$; $^{f}P < .001$. Comparison between total and plasma populations in ADNI: $^{g}P < .05$; $^{h}P < .01$; $^{i}P < .001$. *11 cognitively normal participants had progressed to MCI at the present study baseline, and these were approximated as EMCI.
subgroups in BioFINDER and in the independent ADNI cohort for a robust cross-validation. Equations for calculating the individual risk of being Aβ positive were derived from the estimates and intercepts in the different models. The statistics were performed using R, version 3.3 (R Foundation for Statistical Computing, Vienna, Austria, 2013), and SPSS for Mac, version 22 (SPSS Inc., Chicago, IL). The amyloid risk models were implemented online using a R Shiny (version 1.0.0) program.

3. Results

The characteristics of training (BioFINDER) and validation (ADNI) cohorts are described in the Supplementary Material and shown in Table 1.

3.1. Establishing the amyloid prediction models in BioFINDER

The different Aβ prediction models are illustrated in Fig. 1A and Supplementary Table 1. The selected variables from the LASSO regression were age, APOE ε2ε2/ε2ε3, APOE ε2ε4/ε3ε4, APOE ε4ε4, and the 10-word list delayed recall (see Fig. 1 legend for a complete list of examined variables). Hereafter, this is referred to as the “delayed recall” model. In a multivariable logistic regression, coefficients and intercept were established (Supplementary Table 1). Because a 10-year probability of 0.83 (95% CI 0.79–0.87) from the LASSO regression were then age, APOE ε2ε2/ε2ε3, APOE ε2ε4/ε3ε4, APOE ε4ε4, and MMSE orientation and memory. This is referred to as the “MMSE model.” In a logistic regression, this model had slightly less AUC than the delayed recall model (AUC 0.81, 95% CI 0.77–0.85), and a comparison of the AICs also favored the delayed recall model (ΔAIC 17).

Next, we reran the aforementioned LASSO analyses but also included the plasma biomarkers Aβ42/Aβ40, NfL, and tau. The selected variables from the analysis were age, APOE ε2ε2/ε2ε3, APOE ε2ε4/ε3ε4, APOE ε4ε4, the 10-word list delayed recall, and plasma Aβ42/Aβ40. This produced the best model with ΔAICs of -8 to -34 compared with the other models and the highest AUC of all models (0.85, 95% CI 0.81–0.89) (Fig. 1A; Supplementary Table 1). When excluding grading of white matter lesions, AQT, and 10-word list delayed recall from the LASSO model, plasma Aβ42/Aβ40 was again selected, in addition to age, APOE ε2ε2/ε2ε3, APOE ε2ε4/ε3ε4, APOE ε4ε4, and MMSE orientation and memory. The AUC from the

![Fig. 1. (A–B) Prediction of Aβ positivity in BioFINDER. Logistic regression analyses of the variables selected in the LASSO analysis. The results are shown as AUCs based on the probabilities from the models. Error bars represent the 95% CI of the AUC, where >0.5 in the lower bound indicates that the model significantly predicts Aβ positivity. All models were derived from the same 391 subjects. (A) shows the four multivariable amyloid risk models. The colors in (A) correspond to the color coding of the variables in (B) and show the added AUC in addition to the previous variable(s). A vertical dashed line has arbitrarily been added at AUC 0.80 for easier comparison between the models. The delayed recall, MMSE, and plasma models were derived from the different sets of variables (but from the same population) using the LASSO analysis as the selection method. The AIC shows the model fit in relation to its complexity (number of variables) where lower AIC equals a better model fit (a decrease of >2 indicates a significantly better model). Detailed data of each cumulative step are shown in Supplementary Table 1. (B) shows univariate analyses of the selected variables. Note that the different APOE variables show the performance of each specified APOE group in contrast to all other groups. The complete performance of APOE (divided into 2, 3, and 4 groups, respectively) is shown in Supplementary Fig. 2. Delayed recall model: Age, 10-word list delayed recall, APOE ε2ε2/ε2ε3, ε2ε4/ε3ε4, and ε4ε4. MMSE model: As above but with MMSE orientation and memory instead of delayed word recall.

List of predictor variables in the LASSO analysis: 10-word list delayed recall (from ADAS-cog), MMSE total score (0–30 p), MMSE orientation and memory (0–13 p), AQT (including ratios with the other cognitive measures), white matter lesions (ARWMC scale), presence of APOE ε2ε2/ε2ε3, presence of APOE ε2ε4/ε3ε4, and presence of APOE ε4ε4. In the reduced set of variables (for the MMSE model), white matter lesions, delayed recall, and AQT were excluded (but the same population was used). In the secondary analyses, plasma NfL, plasma Aβ42/Aβ40 ratio, and plasma tau were added to the two sets of variables (also using the same population). Abbreviations: Aβ, β-amyloid; AIC, Akaike Information Criterion; APOE, apolipoprotein E; ARWMC, age-related white matter changes; AUC, area under the ROC curve; BioFINDER, Biomarkers For Identifying NeuroDegenerative Disorders Early and Reliably; CI, confidence interval; LASSO, least absolute shrinkage and selection operator; MMSE, Mini–Mental State Examination; ROC, receiving operating characteristics.
logistic regression was 0.83 (95% CI 0.79–0.87), which was favorable compared with the MMSE model without plasma Aβ42/Aβ40 (ΔAUC 0.02 and ΔAIC −16). In univariate analyses of the selected variables from the LASSO regression, plasma Aβ42/Aβ40 had the highest accuracy (AUC 0.74, 95% CI 0.69–0.79) (Fig. 1B and Supplementary Table 1).

3.2. Replicating the models in ADNI

The BioFINDER models were replicated in both the ADNI subset where plasma Aβ42/Aβ40 values were available (n = 170) and in the total eligible ADNI population (n = 661), that is, the equations in Supplementary Fig. 1 were tested in the ADNI samples (a new model was not fitted in ADNI). The different replications are shown in Fig. 2 and described with exact data in Supplementary Table 2. When replicating the delayed recall model in ADNI, the AUC was 0.82 (95% CI 0.75–0.89) compared with 0.83 in BioFINDER. The AUC was 0.83 (95% CI 0.77–0.89) when replicating the delayed recall model plus plasma Aβ42/Aβ40 (AUC 0.85 in BioFINDER). The MMSE model had an AUC of 0.81 (95% CI 0.77–0.89), equal to its original performance in BioFINDER (AUC 0.81, 95% CI 0.77–0.85). Similar performance was seen when adding plasma Aβ42/Aβ40 (AUC 0.83, 95% CI 0.76–0.89, in ADNI compared with 0.83, 95% CI 0.77–0.85, in BioFINDER). In the total ADNI population (n = 661), both the delayed recall and MMSE models had AUC of 0.80 (95% CI 0.77–0.84 and 0.77–0.83, respectively). The performance of the models in the eight different subpopulations in BioFINDER and ADNI (Fig. 2 and Supplementary Table 2) was robust when tested within different age strata or within different groups of cognitive impairment (subjective cognitive decline, early MCI, and late MCI).

3.3. Calculating the individual risk of being amyloid positive

The models were implemented and published on http://amyloidrisk.com where the individual probability of being Aβ positive can be calculated, including a 95% CI of the predicted probability. The plasma models were not implemented on the website because we believe further research is needed in terms of assay standardization and preanalytical protocols. ROC curves with sensitivity and specificity for each amyloid risk probability is shown in Fig. 3A–D. The highest Youden index (sensitivity + specificity - 1) was produced using a cutoff of 56% probability of amyloid
positivity for the delayed recall model (sensitivity 71%, specificity 83%), 59% probability for the MMSE model (sensitivity 66%, specificity 83%), 43% for the delayed recall model plus plasma Aβ42/Aβ40 (sensitivity 85%, specificity 71%), and 50% for the MMSE model plus plasma Aβ42/Aβ40 (sensitivity 75%, specificity 77%).

4. Discussion

In this study, we have developed four different amyloid risk models based on consecutively recruited nondemented patients in BioFINDER (n = 391). The models, which included the predictors age, APOE genotype, and parts of the MMSE or a delayed recall test, could accurately predict Aβ positivity (AUCs 0.81–0.83) and were validated in an independent population (ADNI, n = 170–661) with similar accuracies. The addition of plasma Aβ42/Aβ40 to APOE, age, and brief cognitive testing increased the accuracy slightly.

There are several previous suggestions on how to estimate Aβ positivity based on MRI measures, neuropsychological tests, APOE genotypes, and blood-based biomarkers [8,9,12,38–41]. For example, we previously found that a combination of demographics, APOE, and longitudinal cognitive testing could be used to identify Aβ positivity in cognitively healthy controls [12]. Recently, age and APOE were examined as predictors of Aβ positivity in MCI and subjects without objective cognitive decline [42]. The AUCs in that study were lower (0.74–0.75), and no increase in AUC was seen when MMSE was added. This might be explained by how APOE was coded (only as ε4+/−) and that they used the total MMSE score, in contrast to the present study where we used four APOE groups based on their different contributing risks to Aβ accumulation [19] and the use of only AD-specific parts of the MMSE score (orientation and memory) [22,23].

A common limitation in many of the previous studies is that the Aβ prediction models have not been validated in
an independent population. In the present models, we only used biomarkers or measures that previously have been shown to either be associated with Aβ deposition or to predict future development of AD dementia [19,23,27,41], to reduce the risk of random inaccurate findings. The robustness of the models was confirmed by validating them in the independent ADNI population and in eight different subgroups (Fig. 2A–D). Note that the models performed well also in selected populations of individuals with only subjective cognitive symptoms (BioFINDER) and significant memory concerns or early MCI (ADNI), which may be of high interest in clinical trials of novel treatments. This also shows that the high accuracy of the models was not driven by the difference in cognitive status between subjective cognitive decline and MCI (BioFINDER) or early MCI and late MCI (ADNI).

The training (BioFINDER) and validation (ADNI) cohorts are different in many ways, which makes it more likely that the established models are indeed generalizable. The differences include, for example, geographic locations (Sweden and North America), education levels (lower in BioFINDER, high in ADNI), cognitive tests in different languages, and the patient selection process (consecutively recruited subjects referred to memory clinics in BioFINDER; selected enrollment in ADNI). Nonetheless, we want to mention potential limitations in these cohorts. The amyloidosis is to a large extent associated with late-onset AD, and the applicability in early-onset AD remains to be tested. The models need further validation in selected primary care populations with individuals who seek medical care due to cognitive complaints (i.e., tested in populations with lower prevalence of Aβ positivity). Finally, the models should be validated in populations where the prevalence of different APOE genotypes differs from the North European/North American populations used in the present study [43].

One popular aim has been to try to identify blood-based AD biomarkers. Plasma biomarker signatures of brain Aβ has, however, been difficult to replicate. Voyle et al. [8] recently performed a large attempt to validated 35 different plasma proteins that had predicted Aβ positivity in previous studies [38–40,44]. Unfortunately, none of the proteins were significantly associated with neocortical Aβ burden in the independent cohort. In the present study, we examined the additive effect of plasma AB42/AB10, NFL, and tau in our models because these biomarkers have been associated with AD [27–29]. Although levels of NFL were significantly higher in Aβ-positive individuals (Table 1), only plasma AB42/AB10 was an independent predictor of brain Aβ in addition to age, APOE genotype, and cognitive testing. Plasma AB42/AB10 was also the predictor with the highest accuracy in the univariate analysis (Fig. 1B). It increased the AUC in both the delayed recall and MMSE models (Fig. 1A and Supplementary Table 1) and increased the AUC when replicated in ADNI (Fig. 2C–D and Supplementary Table 2). However, the clinical relevance of such a small increase in AUC is limited. Also, assay-dependent differences, or possibly preanalytical factors, may have contributed to different levels in the cohorts (Table 1). This highlights the need for an optimal unified analysis method for plasma AB42/AB10. Promising results with very high accuracies have been seen using mass spectrometry [45,46], but unfortunately this is an advanced and time-consuming technique that cannot be implemented in primary care or large screening settings in the near future.

We propose that the presented models could be useful in mainly two settings, clinical AD trials and primary care. In clinical trials aimed at Aβ-positive subjects, amyloid risk models could reduce the number of unnecessary Aβ PET scans or LPs. In Fig. 4, we illustrate such a scenario using the delayed recall model. Here, we assume that 1000 Aβ-positive subjects are to be included in a clinical trial where Aβ PET is used to verify and assess the Aβ burden. An amyloid risk screening process in a population similar to the BioFINDER cohort could reduce the number of unnecessary (negative) Aβ PET scans by ~90% and reduce the costs by >3.5 million USD [12,47], when using a probability cutoff of >80% for undergoing an Aβ PET scan. In the trial scenario, the objective is thus to increase Aβ prevalence of the eligible population (high specificity). On the other hand, in a primary care workup of cognitive impairment or in a scenario where anti-Aβ drugs have become available, a high sensitivity may be preferred. Here, a probability threshold of around 30% would perhaps be more suitable to ensure a sensitivity of >90% (Fig. 3). To facilitate such a use of the risk models, we have implemented them on http://amyloidrisk.com where age, APOE genotype, and cognitive test score can be entered to calculate the individual probability of being Aβ positive. The website is only intended for research and education until further validation has been conducted, but we believe it can be a useful tool for deciding who should undergo further evaluation with LP or Aβ PET to verify the presence of Aβ pathology.

Acknowledgments

The authors thank Fredrik Kahn, MD, PhD, for valuable input and feedback. Work at the authors’ research center was supported by the European Research Council, the Swedish Research Council, the Marianne and Marcus Wallenberg foundation, the Strategic Research Area MultiPark (Multidisciplinary Research in Parkinson’s disease) at Lund University, the Swedish Brain Foundation, the Skåne University Hospital Foundation, the Swedish Alzheimer Association, and the Swedish federal government under the ALF agreement and F. Hoffmann-Roche Ltd. Doses of 18F-flutemetamol injection were sponsored by GE Healthcare. The ADNI data
collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer’s Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

S.P., P.S.I., N.M., and E.S. report no conflicts of interest. B.B. is a full-time employee of Euroimmun. K.B. has served as a consultant or at advisory boards for Alzheon, BioArctic, Biogen, Eli Lilly, Fujirebio Europe, IBL International, Merck, Novartis, Pfizer, and Roche Diagnostics. K.B. and H.Z. are co-founders of Brain Biomarker Solutions in Gothenburg AB, a GU Venture-based platform company at the University of Gothenburg. H.Z. has served at advisory boards of Eli Lilly and Roche Diagnostics and has received travel support from TEVA. O.H. has acquired research support (for the institution) from Roche, GE Healthcare, Biogen, AVID Radiopharmaceuticals, Fujirebio, and Euroimmun. In the past 2 years, he has received

Fig. 4. Cost analysis for clinical trials. In this scenario, we assumed the same prevalence of Aβ positivity as in the BioFINDER study and a trial design that required 1000 Aβ+ subjects to be enrolled and Aβ status verified using Aβ PET (with an approximate cost of 4000 USD per PET scan [12,47]). The y-axis shows the total Aβ PET cost, and the x-axis shows the probability of being Aβ positive according to the delayed recall model (this probability output is available on http://amyloidrisk.com). The line in the graph shows the study cost as a function of the Aβ probability, with the number needed to screen with the amyloid risk model on the left side (the prescreening process) and the number needed to undergo an Aβ PET scan to verify the Aβ status on the right side (normal trial screening process). The small costs for APOE genotyping (approximately 40 USD) and for administering the cognitive test were disregarded. The costs for verifying the Aβ status using LP and CSF Aβ42 analysis are approximately 10% of the costs for an Aβ PET scan (in Europe [47,48], so the costs on the y-axis can be divided by 10 for a CSF-based scenario. Abbreviations: Aβ, β-amyloid; APOE, apolipoprotein E; CSF, cerebrospinal fluid; LP, lumbar puncture; PET, positron emission tomography.
consultancy/speaker fees (paid to the institution) from Lilly, Roche, and Fujirebio.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jalz.2018.08.014.

RESEARCH IN CONTEXT

1. Systematic review: We reviewed publications of β-amyloid (Aβ) prediction using PubMed. There are previous prediction models, but they lack adequate accuracy, replicable results, readily available measures, and/or individual risk stratification.

2. Interpretation: Using just age, APOE genotype, and a brief cognitive test, we accurately predicted Aβ positivity in a training cohort (area under the receiver operating characteristics curve = 0.81–0.83, n = 391) and replicated the models in an independent validation cohort (area under the receiver operating characteristics curve = 0.80–0.82, n = 170–661). The individual probability of Aβ positivity can be calculated on http://amyloidrisk.com. This is useful, for example, in the primary care workup of prodromal Alzheimer’s disease or when screening participants in Alzheimer’s disease trials for selecting persons who should be further examined with amyloid PET or cerebrospinal fluid analysis.

3. Future directions: The models need to be replicated in populations with lower prevalence of Aβ positivity (e.g., primary care). The addition of plasma Aβ42/ Aβ40 seems to improve the models, but further standardization of assays and preanalytical protocols is needed.

References


