Genetic Risk as a Marker of Amyloid-β and Tau Burden in Cerebrospinal Fluid

Nicola Voylea, Hamel Patela, Amos Folarina, Stephen Newhousea, Caroline Johnstona, Pieter Jelle Visserb,c, Richard J.B. Dobsona,d,e,1,∗ and Steven J. Kiddlea,f,1,∗ the EDAR and DESCRIPA study groups and the Alzheimer’s Disease Neuroimaging Initiative2,†

aMRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK
bDepartment of Neurology and Alzheimer Center, VU University Medical Center, Amsterdam, The Netherlands
cHelen Wills Neuroscience Institute, University of California, Berkeley, CA, USA
dNIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK
eFarr Institute of Health Informatics Research, UCL Institute of Health Informatics, University College London, London, UK
fMRC Biostatistics Unit, Cambridge Biomedical Campus, Cambridge Institute of Public Health, Forvie Site, Robinson Way, Cambridge, UK

Handling Associate Editor: Gary Arendash

Accepted 28 September 2016

Abstract

Background: The search for a biomarker of Alzheimer’s disease (AD) pathology (amyloid-β (Aβ) and tau) is ongoing, with the best markers currently being measurements of Aβ and tau in cerebrospinal fluid (CSF) and via positron emission tomography (PET) scanning. These methods are relatively invasive, costly, and often have high screening failure rates. Consequently, research is aiming to elucidate blood biomarkers of Aβ and tau.

Objective: This study aims to investigate a case/control polygenic risk score (PGRS) as a marker of tau and investigate blood markers of a combined Aβ and tau outcome for the first time. A sub-study also considers plasma tau as markers of Aβ and tau pathology in CSF.

Methods: We used data from the EDAR*, DESCRIPA**, and Alzheimer’s Disease Neuroimaging Initiative (ADNI) cohorts in a logistic regression analysis to investigate blood markers of Aβ and tau in CSF. In particular, we investigated the extent to which a case/control PGRS is predictive of CSF tau, CSF amyloid, and a combined amyloid and tau outcome. The predictive ability of models was compared to that of age, gender, and APOE genotype (‘basic model’).

Results: In EDAR and DESCRIPA test data, inclusion of a case/control PGRS was no more predictive of Aβ and a combined Aβ and tau endpoint than the basic models (accuracies of 66.0%, and 73.3% respectively). The tau model saw a small increase in accuracy compared to basic models (59.6%). ADNI 2 test data also showed a slight increase in accuracy for the Aβ model when compared to the basic models (61.4%).

Conclusion: We see some evidence that a case/control PGRS is marginally more predictive of Aβ and tau pathology than the basic models. The search for predictive factors of Aβ and tau pathologies, above and beyond demographic information,
is still ongoing. Better understanding of AD risk alleles, development of more sensitive assays, and studies of larger sample size are three avenues that may provide such factors. However, the clinical utility of possible predictors of brain Aβ and tau pathologies must also be investigated.

**Beta amyloid oligomers in the early diagnosis of AD and as marker for treatment response**

**Development of screening guidelines and criteria for pre-dementia Alzheimer’s disease**

Keywords: Alzheimer’s disease, biomarker, blood, multi-modal, polygenic risk score, tau

**INTRODUCTION**

The hallmark pathologies of Alzheimer’s disease (AD) are amyloid-β (Aβ) plaques and phosphorylated tau tangles in the brain. Although diagnostic criteria for AD focus on pathological evidence of Aβ, tau levels in cerebrospinal fluid (CSF) are also considered [1]. Jack et al. provide a theoretical model for the progression of AD based on existing evidence that is consistent with the popular amyloid cascade hypothesis [2, 3]. This hypothesis states that the build up of tau is triggered by increasing levels of Aβ in the brain.

The search for a biomarker of AD pathologies (tau and Aβ) is ongoing with the best markers currently being measurements of tau and Aβ in the CSF and via positron emission tomography (PET) scanning. These methods are relatively invasive and often have high screening failure rates meaning a high proportion of individuals that are scanned have low levels of these pathologies. Additionally, PET scanning in particular is an expensive procedure costing around $3,000 per scan [4]. There are only approximately 2,380 (http://www.imvinfo.com, August 2016) PET scanners in the United States meaning access to facilities is limited [5]. The lumbar puncture needed to access CSF is also considered a high-risk procedure in many western countries [6]. To address these shortcomings, research is aiming to elucidate blood biomarkers of AD pathologies (Aβ and tau) [7]. One motivation for discovering a blood-based biomarker of AD pathology comes from clinical trial recruitment. For example, when recruiting into a trial of an anti-amyloid therapeutic, a blood-based biomarker of Aβ could act as a filtering step to identify individuals with abnormal levels of Aβ before performing a confirmatory PET scan or lumbar puncture. As such a test is likely to be cost-effective, we could reduce the cost of screening while also reducing the number of individuals subject to the invasive lumbar puncture and PET scanning procedures. So far, research into a blood marker has largely focused on Aβ brain burden, rather than tau pathology. However, some studies have investigated genetic markers of tau [8]. This study aims to further this research by investigating markers of a combined Aβ and tau outcome for the first time.

 Genome wide association studies (GWAS) of AD to date have identified over 20 risk loci explaining approximately 16–33% of genetic variability in the disease [9–13]. Compared to the predicted heritability of 50–70%, this is fairly low [13–15]. Modern technologies including next generation sequencing and the development of high coverage, exome chips are beginning to address this issue of missing heritability. Meanwhile polygenic risk scores (PGRS) are aiming to combine genetic risk from variants of lower effect size [16]. To date, PGRS in AD have only been trained on case/control endpoints as GWAS studies of pathological outcomes are relatively small (for example, N = 1,269 [17]). Studies have shown case/control PGRS are associated with AD-related phenotypes but few have investigated their predictive ability in test data [18, 19]. When we studied the predictive ability of a PGRS for Aβ burden, we saw that the case/control PGRS used was no more predictive than demographics (age, gender) and APOE genotype alone (Voyle et al., in submission). However, we hypothesize in the current study that the PGRS may be more predictive of a combined Aβ and tau outcome; a more ‘AD-like’ phenotype.

This study aims to investigate the predictive ability of a case/control PGRS on CSF Aβ and tau. A sub-study will also consider plasma tau as a predictor. We use individuals from the Alzheimer’s Disease Neuroimaging Initiative (ADNI), EDAR, and DESCRIPTA studies to investigate these associations and compare all models to those of demographics (age, gender) and APOE genotype [20, 21]. We hypothesize that blood markers of AD will perform better when predicting a combined pathology (tau and Aβ) endpoint over tau and Aβ individually. The combined endpoint should be more representative of an ‘AD-like’ phenotype.
MATERIALS AND METHODS

Cohorts

EDAR was a prospective, longitudinal study which aimed to examine and evaluate biomarkers of early AD and treatment response [21]. In particular, the study focused on Aβ oligomers and the effect of genetic variants on these oligomers. For more information see http://www.edarstudy.eu. Our access to samples and clinical and phenotypic information from the EDAR study was enabled by the European Medical Information Framework and has been previously described (Voyle et al. in submission).

DESCRIPA was also a prospective, multi-center study. It was coordinated by the European AD Consortium and focused on collecting data from non-demented subjects with the aim of developing screening guidelines and clinical criteria for AD. Further details of this study can be found in Visser et al. [20].

ADNI is a longitudinal cohort study aiming to validate the use of biomarkers in AD clinical trials and diagnosis. Data used in the preparation of this article were obtained from the ADNI database (http://adni.loni.usc.edu). The ADNI study 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 biological markers and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and AD. ADNI was approved by the institutional review boards of all participating institutions, and written informed consent was obtained for all participants. This study uses data from ADNI 1 and the ADNI 2 and ADNI GO sub-groups, referred to as ADNI 2 from here onwards.

Genetics

Samples from EDAR and DESCRIPA were genotyped on the Illumina HumanOmniExpressExome-8v1.2 BeadChip (N = 336) [22]. The data has previously been reported on by Voyle et al. (in submission). The HumanOmniExpressExome BeadChip has been optimized to tag SNPs that capture a large amount of common genetic variation. In total, the chip contains 960,919 markers of which over 273,000 correspond to functional exomic markers. ADNI 1 samples were run on the Human610-Quad BeadChip (N = 818), which has since been discontinued. The chip provides coverage of 924,000 randomly selected SNPs. ADNI 2 and ADNI GO samples were run on the Illumina HumanOmniExpress BeadChip (N = 432). This chip is similar to the HumanOmnExpressExome BeadChip used in the EDAR and DESCRIPA studies but does not include the exomic markers. In total, the chip contains 713,599 markers. Details of the genotyping protocols followed in ADNI are given elsewhere [23]. Details of the data processing are briefly outlined below.

The cohorts were subject to quality control and imputation separately, as described in Coleman et al. [24]. In short, the data was filtered to ensure a minor allele frequency of greater than 5% for all SNPs before removal of rare variants and subjects with high levels of missing data. SNPs that differed significantly (p < 0.00001) from the Hardy-Weinberg equilibrium were removed. The data was pruned for SNPs in linkage disequilibrium and for genetically similar individuals. Finally, the data was imputed using reference files from the 1000 Genomes Project [25].

CSF Measurements

This study focuses on total tau (tTau) and Aβ measurements in CSF. The analysis considered three endpoints: dichotomized Aβ, dichotomized tTau, and a binary representation of overall pathology. For the latter analysis, referred to as total CSF burden, each cohort was reduced to those individuals with normal Aβ and normal tTau, or abnormal Aβ and abnormal tTau. The distributions of tTau and Aβ were similar between cohorts in terms of shape but not in absolute value. Consequently, this work focused solely on a dichotomized outcome.

EDAR

Details of CSF collection and analysis can be found at http://www.edarstudy.eu. In brief, CSF measurements were collected using the Alzbio3 Luminex assay in one batch at the end of the study. CSF Aβ and tTau measurements were dichotomized at the previously published thresholds of 389 pg/ml and 98 pg/ml, respectively.

DESCRIPA

Details of CSF measurements in DESCRIPA have been described elsewhere [22]. In brief, CSF measurements were analyzed in one laboratory and collected using single-parameter ELISA kits.
(Innogenetics, Ghent, Belgium). CSF Aβ and tTau samples were dichotomized using the previously published thresholds of 550 pg/ml and 375 pg/ml, respectively.

**ADNI**
For ADNI, datasets used to extract CSF measures of Aβ and tau were chosen to maximize sample size. The dataset ‘UPENNBIOMK2’ was used for ADNI 1 and ‘UPENNBIOMK6’ for ADNI 2 and ADNI GO. Both datasets contain CSF measurements collected using the xMAP Luminex platform and Innogenetics immunoassay kits. The CSF measures of amyloid and tTau were dichotomized at the previously published thresholds (192 pg/ml and 93 pg/ml, respectively).

**Plasma Tau (ADNI 1 only)**
Plasma tau was investigated as a potential blood biomarker of Aβ and tau in a sub-study. ADNI 1 was the only cohort with such data available. Plasma tau was analyzed by the Single Molecule Array (SIMOA) technique and the Human total tau assay using a combination of monoclonal antibodies. Samples with a plasma tau concentration below the lower limit of quantification (<1.0 pg/ml) were removed (N = 36). Outliers, identified as values greater than 6 standard deviations from the mean, were removed (N = 2) and the data was subject to a natural logarithm transformation.

**Statistical analysis**
All statistical analysis was performed in R version 3.1.1 [26].

The three endpoints of interest (dichotomized amyloid, dichotomized tTau, and total CSF burden) were modeled using logistic regression models covarying for age, gender, and APOE genotype. An individuals APOE genotype was coded as 1 if at least one ε4 allele was present and 0 otherwise. Models were also built using only the demographic variables age, gender, and APOE genotype, for comparison. Throughout this study these models are referred to as ‘basic models’.

Two analyses were performed to study these three endpoints of interest:

1. **PGRS**:
   Models were built in ADNI 1 data (Aβ and tau N = 363; Total CSF burden N = 244) and tested in data from EDAR and DESCRIPA (Aβ and tau N = 250; Total CSF burden N = 135) and
   ADNI 2 (Aβ and tau N = 44; Total CSF burden N = 37).

2. **PGRS and plasma tau**: Models were built and tested in ADNI 1 using 5 fold cross-validation (Aβ and tau N = 323; Total CSF burden N = 219).

The sample sizes of each dataset used in these analyses are given in Table 1.

All models including a PGRS co-varied for the first three genetic principal components to account for population structure. The predictive ability of each model was quantified using accuracy, sensitivity, specificity and area under the receiver operating characteristic (ROC) curve [27–29].

**PGRS**
PGRS were created within each cohort using PRSice [30]. Effect sizes from stage 1 of the International Genomics of Alzheimer’s Project (IGAP) case/control GWAS were used as the weights to generate the risk score [9]. We used 0.5 as the \( p \)-value threshold for inclusion in the PGRS. This threshold showed the most significant association with case/control diagnosis in the large IGAP PGRS study [31]. The genetic region coding for APOE was removed from all scores and included as a separate covariate due to its strong influence. The PGRS was standardized by subtracting the mean and dividing by the standard deviation, per cohort. This aims to account for the scores including different SNPs due to availability on the different SNP chips. It is important to note that APOE genotype is included as a covariate in modeling and not in the PGRS. Therefore, the PGRS is only exploring variation above and beyond APOE. This is different to the focus of the IGAP PGRS study which included APOE within the PGRS [31]. We have chosen this approach to investigate
Table 2

Cohort demographics – Dichotomized Aβ

<table>
<thead>
<tr>
<th>Demographic</th>
<th>ADNI 1 (N = 363)</th>
<th>ADNI 2 (N = 44)</th>
<th>EDAR and DESCRIPA* (N = 250)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Abnormal</td>
<td>p-value</td>
<td>Normal</td>
</tr>
<tr>
<td>Median age [IQR]</td>
<td>74.4 [8.35]</td>
<td>75.3 [8.55]</td>
<td>0.825</td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>36.6</td>
<td>40.6</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>63.4</td>
<td>59.4</td>
<td></td>
</tr>
<tr>
<td>Diagnosis (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dementia</td>
<td>6.3</td>
<td>33.5</td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>40.1</td>
<td>51.0</td>
<td></td>
</tr>
<tr>
<td>SCI</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>53.6</td>
<td>15.5</td>
<td></td>
</tr>
<tr>
<td>APOE status (%)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Kruskal Wallis chi-squared was used to test between normal and abnormal groups for continuous demographic variables. Fishers exact was used to test between normal and abnormal groups for categorical demographic variables. APOE status is 1 if an individual’s genotype contains any e4 alleles, and 0 otherwise. IQR, Inter-quartile range; CSF, cerebrospinal fluid; MMSE, Mini-Mental State Exam; MCI, mild cognitive impairment; SCI, subjective cognitive impairment; CTL, control. *One individual has missing diagnosis, 2 have missing education information, and 3 have missing MMSE.

---

whether information from a SNP chip adds anything above APOE. Furthermore, generally, APOE is not well measured on SNP chips so APOE genotype is determined by targeted genotyping.

The datasets used for model testing (EDAR and DESCRIPA) were not included in the IGAP study. Furthermore, although some samples from ADNI were included in IGAP, the majority were only included in the training, not testing, stages.

**RESULTS**

**Cohort demographics**

Table 2 shows demographics against normal and abnormal CSF Aβ while Table 3 is against CSF tTau. Demographics for the sub-population used in the total CSF burden analysis are given in Supplementary Table 1. Demographics for the sub-group of ADNI 1 individuals with plasma tau measurements is given in Supplementary Table 2.

In ADNI 1 training data and EDAR/DESCRIPA test data, there is a significant difference between normal and abnormal Aβ in MMSE, diagnosis, and APOE genotype as we would expect. Similar associations are also seen with dichotomized tTau. The smaller sample size of the ADNI 2 test data (N = 44) is likely to have driven the lack of association in this group of individuals.

---

**Data analysis**

**Genetic risk**

The PGRS was not significant in any of the logistic regression models (p-values of 0.995, 0.929, and 0.796 for tTau, Aβ, and total CSF burden respectively). The inclusion of the PGRS marginally improved the predictive ability of tTau models over the basic models. The accuracy of the Aβ model was also marginally improved by inclusion of the PGRS in ADNI 2 test data. The models of total CSF burden had the highest accuracies (72% and 65%). See Table 4 for full results.

**Genetic risk and plasma tau**

When modeling CSF tTau no model outperformed the basic model at an accuracy of 66%. Similarly, when modeling total CSF burden the inclusion of a PGRS and plasma tau did not improve predictive ability above the basic model (77%). The only model to see an improvement over the basic model was when modeling Aβ. Inclusion of plasma tau measurements marginally improved accuracy from 71% to 74% and the area under the ROC curve from 0.658 to 0.697. See Table 5 for full results.

**DISCUSSION**

The aim of this study was to investigate blood biomarkers that may be predictive of AD pathologies...
Table 3
Cohort demographic - Dichotomized total tau

<table>
<thead>
<tr>
<th>Demographic</th>
<th>ADNI 1 (N = 363)</th>
<th>ADNI 2 (N = 44)</th>
<th>EDAR and DESCRIPA* (N = 250)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Abnormal</td>
<td>p-value</td>
</tr>
<tr>
<td></td>
<td>CSF tau (N = 203)</td>
<td>CSF tau (N = 160)</td>
<td></td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
<td>0.234</td>
</tr>
<tr>
<td>Female</td>
<td>36.5</td>
<td>43.1</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>63.5</td>
<td>56.9</td>
<td></td>
</tr>
<tr>
<td>Diagnosis (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dementia</td>
<td>15.8</td>
<td>36.9</td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>45.8</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>SCI</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>38.4</td>
<td>13.1</td>
<td></td>
</tr>
<tr>
<td>APOE status (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>63.5</td>
<td>31.9</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>36.5</td>
<td>43.1</td>
<td></td>
</tr>
</tbody>
</table>

Kruskal Wallis chi-squared was used to test between normal and abnormal groups for continuous demographic variables. Fisher’s exact was used to test between normal and abnormal groups for categorical demographic variables. APOE status is 1 if an individual’s genotype contains any e4 alleles, and 0 otherwise. IQR, inter-quartile range; CSF, cerebrospinal fluid; MMSE, Mini-Mental State Exam; MCI, mild cognitive impairment; SCI, subjective cognitive impairment; CTL, control. *One individual has missing diagnosis, 2 have missing education information, and 3 have missing MMSE.

Table 4
PGRS: Test data results

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Model</th>
<th>ADNI 1</th>
<th>ADNI 2</th>
<th>ADNI 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Accuracy</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>EDAR and DESCRIPA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tTau</td>
<td>Demographics only</td>
<td>0.584</td>
<td>0.570</td>
<td>0.593</td>
</tr>
<tr>
<td>Aβ</td>
<td>Demographics only</td>
<td>0.660</td>
<td>0.788</td>
<td>0.465</td>
</tr>
<tr>
<td>Total CSF burden</td>
<td>Demographics only</td>
<td>0.733</td>
<td>0.750</td>
<td>0.716</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tTau</td>
<td>Demographics and PGRS</td>
<td>0.596</td>
<td>0.560</td>
<td>0.620</td>
</tr>
<tr>
<td>Aβ</td>
<td>Demographics and PGRS</td>
<td>0.616</td>
<td>0.768</td>
<td>0.384</td>
</tr>
<tr>
<td>Total CSF burden</td>
<td>Demographics and PGRS</td>
<td>0.719</td>
<td>0.702</td>
<td>0.702</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AUC ROC, area under the receiver operating curve; PGRS, polygenic risk score.

Several studies have focused on identification of blood biomarkers of Aβ [17, 32, 33]. However, few have achieved successful replication. The hypothesis tested in this study was that a combined tau and Aβ endpoint would be closer to an AD case/control phenotype and consequently easier to predict. This study supports this hypothesis, although the improvements in accuracy are minimal. Further, this accuracy of prediction was achieved by the basic model and the two blood biomarkers investigated, a case/control PGRS (Aβ and tau). We modeled levels of Aβ and tau in CSF using a PGRS and a sub-study considered measurements of tau in blood plasma. We also studied a total CSF burden endpoint for individuals with abnormal tau and Aβ, or normal tau and Aβ. The results shown here highlight that a case/control PGRS and plasma tau do not substantially outperform demographics (age, gender) and APOE genotype. The highest model accuracies were seen when modeling the total CSF burden phenotype.
Table 5
PGRS and plasma tau: Five fold cross-validation results

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Model</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>tTau</td>
<td>Demographics only</td>
<td>0.656</td>
<td>0.691</td>
<td>0.628</td>
<td>0.659</td>
</tr>
<tr>
<td>tTau</td>
<td>Demographics and PGRS</td>
<td>0.638</td>
<td>0.665</td>
<td>0.617</td>
<td>0.641</td>
</tr>
<tr>
<td>tTau</td>
<td>Demographics and plasma tau</td>
<td>0.650</td>
<td>0.587</td>
<td>0.700</td>
<td>0.644</td>
</tr>
<tr>
<td>tTau</td>
<td>Demographics, PGRS and plasma tau</td>
<td>0.653</td>
<td>0.608</td>
<td>0.689</td>
<td>0.649</td>
</tr>
<tr>
<td>Aβ</td>
<td>Demographics only</td>
<td>0.709</td>
<td>0.787</td>
<td>0.530</td>
<td>0.658</td>
</tr>
<tr>
<td>Aβ</td>
<td>Demographics and PGRS</td>
<td>0.697</td>
<td>0.769</td>
<td>0.532</td>
<td>0.650</td>
</tr>
<tr>
<td>Aβ</td>
<td>Demographics and plasma tau</td>
<td>0.743</td>
<td>0.813</td>
<td>0.582</td>
<td>0.697</td>
</tr>
<tr>
<td>Aβ</td>
<td>Demographics, PGRS and plasma tau</td>
<td>0.725</td>
<td>0.813</td>
<td>0.523</td>
<td>0.668</td>
</tr>
<tr>
<td>Total CSF burden</td>
<td>Demographics only</td>
<td>0.772</td>
<td>0.742</td>
<td>0.816</td>
<td>0.779</td>
</tr>
<tr>
<td>Total CSF burden</td>
<td>Demographics and PGRS</td>
<td>0.763</td>
<td>0.735</td>
<td>0.804</td>
<td>0.769</td>
</tr>
<tr>
<td>Total CSF burden</td>
<td>Demographics and plasma tau</td>
<td>0.758</td>
<td>0.765</td>
<td>0.748</td>
<td>0.756</td>
</tr>
<tr>
<td>Total CSF burden</td>
<td>Demographics, PGRS and plasma tau</td>
<td>0.772</td>
<td>0.780</td>
<td>0.756</td>
<td>0.768</td>
</tr>
</tbody>
</table>

AUC ROC, area under the receiver operating curve; PGRS, polygenic risk score.

and plasma tau, did not improve model accuracy in the majority of cases.

Firstly, this could be explained by the PGRS being trained on a case/control endpoint. As individuals can often be misdiagnosed with AD the case/control phenotype can be misleading. GWAS are beginning to be large enough to detect risk SNPs associated with Aβ and tau. When sample sizes in these studies increase further they should be used to train PGRS. Intuitively, they may achieve improved predictive ability than models based on a case/control PGRS. Furthermore, due to the relatively small sample size of this study, the PGRS only utilizes common variation excluding loci such as TREM2. As larger studies become available the inclusion of such rarer variants in genetic risk scoring should be considered.

In the sub-study, it is interesting that plasma tau is no more predictive of CSF tau than age, gender, and APOE genotype. Furthermore, plasma tau achieves a higher accuracy when modeling CSF Aβ. The lack of ability to predict CSF tau indicates the need for further assay development to detect still more sensitive measurements. Furthermore, research into the permeability of the blood-brain barrier will help theorize as to how much pathology signal from the brain and CSF is likely to be seen in blood.

This study has shown the importance of replication in independent datasets; models that perform well in training data often do not replicate. It is particularly important to test replicability when standardized protocols for assays do not exist [34]. For example, this analysis highlights the difference between EDAR, DESCRIPA, and ADNI in the assays and cut-offs used to define high and low pathology burden. The CSF Aβ cut-off for high/low burden ranges between 192 pg/ml and 550 pg/ml. Efforts are being made to standardize such metrics for future research [35, 36].

It is important to point out the limitations of this study. Firstly, there are differences in sample collection methods, assays and data processing pipelines between ADNI 1, ADNI 2, EDAR, and DESCRIPA. In particular, the GWAS platforms used differ between the studies. The models in this study are trained in ADNI 1, which uses a slightly older Illumina chip (Human610-Quad) than ADNI 2, EDAR, and DESCRIPA. Although this means that the data from the other cohorts may not be fully utilized, it is unlikely to cause a lack of replicability. However, it is possible that some replicability is lost due to differences in sample collection methods and data processing. Furthermore, the use of the older SNP chip (Human610-Quad in ADNI 1) may have lead to reduced SNP content and sub-optimal imputation within the ADNI 1 cohort. It is also of note that the ADNI study was included in IGAP, effect sizes from which were used to create the PGRS. We believe that the benefit of a larger sample size for training outweighs any negative impact. Despite the use of well-characterized cohorts in this study, we must point out that the sample size considered is still relatively small. In future work, this could be addressed by the use of longitudinal aging studies instead of the case/control cohort studies used here.

Finally, we have shown that the multi-modal approach used in the sub-study investigating plasma tau, did not improve predictive ability above the basic model. We used a simple additive model and more complex methods such as OmicKriging may be useful in this setting [37]. Furthermore, the standard for measuring AD pathology, in particular Aβ, is through
PET imaging. Generally, PET imaging and CSF measurements are used interchangeably but any results should be replicated using imaging based outcomes. In combination with a pathological endpoint more closely related to an AD phenotype, such as the total CSF burden used here, the suggestions presented in this discussion could improve the predictive ability of proposed markers of Aβ and tau.

CONCLUSIONS

This study has used data from the EDAR, DESCRIPA, and ADNI cohorts to investigate blood markers of Aβ and tau. We see that a case/control PGRS is no more predictive of pathology than age, gender, and APOE genotype. A sub-study shows that model accuracy is not improved by the addition of plasma tau measurements. These results emphasize that the search for predictive factors of Aβ and tau, above and beyond demographic and APOE information, is still ongoing. Better understanding of AD risk alleles, development of more sensitive assays, and studies of larger sample size are three avenues that may provide such factors. However, the clinical utility of possible predictors of brain Aβ and tau must also be investigated.

ACKNOWLEDGMENTS

This work represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London. This study was supported by researchers at the National Institute for Health Research University College London Hospitals Biomedical Research Centre. The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115372, resources of which are composed of financial contribution from the European Union’s Seventh Framework Programme (FP7/2007-2013) and EFPIA companies’ in kind contribution. Steven Kiddle is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1). Nicola Voyle is funded by the Alzheimer’s Society. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

This work was also supported by awards to establish the Farr Institute of Health Informatics Research at UCL partners, London, from the Medical Research Council, Arthritis Research UK, British Heart Foundation, Cancer Research UK, Chief Scientist Office, Economic and Social Research Council, Engineering and Physical Sciences Research Council, National Institute for Health Research, National Institute for Social Care and Health Research, and Wellcome Trust (grant MR/K006584/1).

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: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; 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 and Development, LLC.; John-son and Johnson Pharmaceutical Research and Development LLC.; Lundbeck; Merck and Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (http://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.

Authors’ disclosures available online (http://j-alz.com/manuscript-disclosures/16-0707r1).

†DESCRIPA and EDAR

Magda Tsolaki
Aristotle University of Thessaloniki, Thessaloniki, Greece
Lars-Olof Wahlund, Yvonne-Freund-Levi
Karolinska Institutet, Karolinska University Hospital, Huddinge, Sweden

Frans Verhey
Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Alzheimer Center Limburg, Maastricht University, Maastricht, The Netherlands

Lucrezia Hausner
Heidelberg University, Central Institute of Mental Health, Mannheim, Germany

EDAR
Gunhild Waldemar Peter Johannsen, Danish Dementia Research Center, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Charlotte E. Teunissen
Neurochemistry Laboratory and Biobank, Dept. of Clinical Chemistry, Neuroscience Campus Amsterdam, VU University Medical Center Amsterdam, The Netherlands

Rik Vandenberghe
University Hospital Leuven, Leuven, Belgium

Descipa
Luiza Spiru
Carol Davila’ University of Medicine and Pharmacy, Bucharest, Romania

Åsa K. Wallin
Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University, Sweden

Marcel Olde-Rikkert
Department of Geriatrics, Radboud University Medical Centre, Nijmegen, The Netherlands

SUPPLEMENTARY MATERIAL
The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-160707.

REFERENCES


