

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

Nicola Voyle^a, Hamel Patel^a, Amos Folarin^a, Stephen Newhouse^a, Caroline Johnston^a, Pieter Jelle Visser^{b,c}, Richard J.B. Dobson^{a,d,e,1,*} and Steven J. Kiddle^{a,f,1,*} the EDAR and DESCRIPA study groups and the Alzheimer's Disease Neuroimaging Initiative^{2,†}

^a*MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK*

^b*Department of Neurology and Alzheimer Center, VU University Medical Center, Amsterdam, The Netherlands*

^c*Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA*

^d*NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK*

^e*Farr Institute of Health Informatics Research, UCL Institute of Health Informatics, University College London, London, UK*

^f*MRC 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 (A β) and tau) is ongoing, with the best markers currently being measurements of A β 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 A β 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 A β and tau outcome for the first time. A sub-study also considers plasma tau as markers of A β 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 A β 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 A β , and a combined A β 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 A β model when compared to the basic models (61.4%).

Conclusion: We see some evidence that a case/control PGRS is marginally more predictive of A β and tau pathology than the basic models. The search for predictive factors of A β and tau pathologies, above and beyond demographic information,

¹These authors are joint senior authors.

²Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://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](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

*Correspondence to: Richard J.B. Dobson and Steven J. Kiddle, SGDP Centre, IoPPN, 16 De Crespigny Park, Denmark Hill, London SE5 8AF, UK. Tel.: +44 020 7848 0924; Fax: +44 020 7848 0866; E-mails: richard.j.dobson@kcl.ac.uk (Richard J.B. Dobson), steven.kiddle@kcl.ac.uk (Steven J. Kiddle).

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 A β 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 (A β) plaques and phosphorylated tau tangles in the brain. Although diagnostic criteria for AD focus on pathological evidence of A β , 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 A β in the brain.

The search for a biomarker of AD pathologies (tau and A β) is ongoing with the best markers currently being measurements of tau and A β 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 (A β 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 A β could act as a filtering step to identify individuals with abnormal levels of A β 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 A β 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 A β 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 A β 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 A β and tau outcome; a more ‘AD-like’ phenotype.

This study aims to investigate the predictive ability of a case/control PGRS on CSF A β 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 DESCRIPA 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 A β) endpoint over tau and A β 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 A β 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 HumanOmniExpressExome 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 A β measurements in CSF. The analysis considered three endpoints: dichotomized A β , 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 A β and normal tTau, or abnormal A β and abnormal tTau. The distributions of tTau and A β 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 A β 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 A β 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 A β 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 A β 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 individual's APOE genotype was coded as 1 if at least one ϵ 4 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 (A β and tau N = 363; Total CSF burden N = 244) and tested in data from EDAR and DESCRIPA (A β and tau N = 250; Total CSF burden N = 135) and

Table 1
Sample sizes

PGRS analysis			
Outcome	ADNI 1 (N)	EDAR and DESCRIPA (N)	ADNI 2 (N)
Dichotomized amyloid or tTau	363	250	44
Total CSF burden	244	135	37
PGRS and plasma tau analysis			
Outcome	ADNI 1 (N)		
Dichotomized amyloid or tTau	323		
Total CSF burden	323		

ADNI 2 (A β 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 (A β 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 A β

Demographic	ADNI 1 (N = 363)			ADNI 2 (N = 44)			EDAR and DESCRIPA* (N = 250)		
	Normal CSF A β (N = 112)	Abnormal CSF A β (N = 251)	p-value	Normal CSF A β (N = 27)	Abnormal CSF A β (N = 17)	p-value	Normal CSF A β (N = 99)	Abnormal CSF A β (N = 151)	p-value
Median age [IQR]	74.4 [8.35]	75.3 [8.55]	0.825	67.8 [12.55]	75.2 [11.1]	0.049	66 [12]	69 [13]	0.136
Gender (%)			0.487			0.359			0.052
Female	36.6	40.6		51.9	35.3		38.4	51.7	
Male	63.4	59.4		48.1	64.7		61.6	48.3	
Median years in education [IQR]	16 [4.25]	16 [4]	0.748	16 [4]	16 [4]	0.825	11.5 [8]	10 [7]	0.133
Median MMSE [IQR]	29 [3]	26 [4]	<0.001	29 [2]	27 [3]	0.058	27.5 [4]	26 [5]	0.009
Diagnosis (%)			<0.001			0.247			0.001
Dementia	6.3	33.5		0	5.9		34.3	45.0	
MCI	40.1	51.0		74.1	82.4		43.4	41.1	
SCI	0	0		0	0		5.1	10.6	
CTL	53.6	15.5		25.9	11.7		17.2	3.3	
APOE status (%)			<0.001			0.225			<0.001
0	85.7	33.5		66.7	47.1		74.7	38.4	
1	14.3	66.5		33.3	52.9		25.3	61.6	

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 ϵ 4 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 A β 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 A β 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, A β , 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 A β 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 A β . 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

Demographic	ADNI 1 (N = 363)			ADNI 2 (N = 44)			EDAR and DESCRIPA* (N = 250)		
	Normal CSF tau (N = 203)	Abnormal CSF tau (N = 160)	p-value	Normal CSF tau (N = 30)	Abnormal CSF tau (N = 14)	p-value	Normal CSF tau (N = 150)	Abnormal CSF tau (N = 100)	p-value
Median age [IQR]	75.2 [8]	74.45 [10]	0.642	67.85 [10.7]	76.75 [12.68]	0.051	66.55 [13]	70.5 [12]	0.032
Gender (%)			0.234			>0.999			0.094
Female	36.5	43.1		46.7	42.9		42.0	53.0	
Male	63.5	56.9		53.3	57.1		58.0	47.0	
Median years in education [IQR]	16 [4]	16 [5]	0.064	16 [4]	16 [4]	0.7	10 [6]	9 [8]	0.032
Median MMSE [IQR]	28 [3]	26 [4]	<0.001	29 [1.75]	27 [3.75]	0.153	28 [4]	26 [4]	<0.001
Diagnosis (%)			<0.001			0.38			<0.001
Dementia	15.8	36.9		0	7.1		26.7	62.0	
MCI	45.8	50.0		76.7	78.6		47.3	34.0	
SCI	0	0		0	0		12.7	2.0	
CTL	38.4	13.1		23.3	14.3		13.3	2.0	
APOE status (%)			<0.001			0.191			0.001
0	63.5	31.9		66.7	42.9		59.3	43.0	
1	36.5	43.1		33.3	57.1		40.7	57.0	

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 $\epsilon 4$ 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

Outcome	Model	Accuracy	Sensitivity	Specificity	AUC ROC
EDAR and DESCRIPA					
tTau	Demographics only	0.584	0.570	0.593	0.625
A β	Demographics only	0.660	0.788	0.465	0.686
Total CSF burden	Demographics only	0.733	0.750	0.716	0.801
tTau	Demographics and PGRS	0.596	0.560	0.620	0.609
A β	Demographics and PGRS	0.616	0.768	0.384	0.692
Total CSF burden	Demographics and PGRS	0.719	0.702	0.702	0.748
ADNI 2					
tTau	Demographics only	0.636	0.571	0.667	0.655
A β	Demographics only	0.546	0.765	0.407	0.527
Total CSF burden	Demographics only	0.649	0.583	0.680	0.693
tTau	Demographics and PGRS	0.659	0.571	0.700	0.671
A β	Demographics and PGRS	0.614	0.647	0.593	0.590
Total CSF burden	Demographics and PGRS	0.649	0.583	0.680	0.683

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

(A β and tau). We modeled levels of A β 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 A β , or normal tau and A β . 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.

Several studies have focused on identification of blood biomarkers of A β [17, 32, 33]. However, few have achieved successful replication. The hypothesis tested in this study was that a combined tau and A β 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

Table 5
PGRS and plasma tau: Five fold cross-validation results

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

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 A β 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 A β . 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 A β 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 A β , 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 A β and tau.

CONCLUSIONS

This study has used data from the EDAR, DESCRIPA, and ADNI cohorts to investigate blood markers of A β 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 A β 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 A β 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, Alzheimers Association; Alzheimers 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.; Johnson and Johnson Pharmaceutical Research and Development LLC.; Lumosity; 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 Hospi-
tal, 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 Amster-
dam, VU University Medical Center Amsterdam,
The Netherlands

Rik Vandenberghe
University Hospital Leuven, Leuven, Belgium

Descripa

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 Med-
ical 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

- [1] Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, Dekosky ST, Gauthier S, Selkoe D, Bateman R, Cappa S, Crutch S, Engelborghs S, Frisoni GB, Fox NC, Galasko D, Habert MO, Jicha G, Nordberg A, Pasquier F, Rabinovici G, Robert P, Rowe C, Salloway S, Sarazin M, Epelbaum S, de Souza LC, Vellas B, Visser PJ, Schneider L, Stern Y, Scheltens P, Cummings JL (2014) Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *Lancet Neurol* **13**, 614-629.
- [2] Hardy J, Higgins G (1992) Alzheimer's disease: The amyloid cascade hypothesis. *Science* **256**, 184-185.
- [3] Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, Shaw LM, Vemuri P, Wiste HJ, Weigand SD, Lesnick TG, Pankratz VS, Donohue MC, Trojanowski JQ (2013) Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *Lancet Neurol* **12**, 207-216.
- [4] Johnson KA, Minoshima S, Bohnen NI, Donohoe KJ, Foster NL, Herscovitch P, Karlawish JH, Rowe CC, Carrillo MC, Hartley DM, Hedrick S, Pappas V, Thies WH (2013) Appropriate use criteria for amyloid PET: A report of the Amyloid Imaging Task Force the Society of Nuclear Medicine and Molecular Imaging and the Alzheimer's Association. *Alzheimers Dement* **9**, E1-E16.
- [5] Buck AK, Herrmann K, Stargardt T, Dechow T, Krause BJ, Schreyögg J (2010) Economic evaluation of PET and PET/CT in oncology: Evidence and methodologic approaches. *J Nucl Med Technol* **38**, 6-17.
- [6] Menéndez-González M (2014) Routine lumbar puncture for the early diagnosis of Alzheimer's disease: Is it safe? *Front Aging Neurosci* **6**, 1-2.
- [7] Bazenet C, Lovestone S (2012) Plasma biomarkers for Alzheimer's disease: Much needed but tough to find. *Biomark Med* **6**, 441-454.
- [8] Cruchaga C, Ebbert MTW, Kauwe JSK (2014) Genetic discoveries in AD using CSF amyloid and tau. *Curr Genet Med Rep* **2**, 23-29.
- [9] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thornton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Hollingworth P, Ramirez A, Hanon O, Fitzpatrick AL, Buxbaum JD, Campion D, Crane PK, Baldwin C, Becker T, Gudnason V, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Moron FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fiévet N, Huentelman MJ, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuinness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogava E, Gallacher J, George-Hyslop PS, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossù, P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Naranjo MCD, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannfelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM,

- Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RFAG, Passmore P, Montine TJ, Bettens K, Rotter JJ, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JSK, Boerwinkle E, Riemenschneider M, Boada M, Hiltunen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nöthen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458.
- [10] Lee SH, Harold D, Nyholt DR, ANZGene Consortium, International Endogene Consortium, Genetic and Environmental Risk for Alzheimer's disease Consortium, Goddard ME, Zondervan KT, Williams J, Montgomery GW, Wray NR, Visscher PM (2013) Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease multiple sclerosis and endometriosis. *Hum Mol Genet* **22**, 832-841.
- [11] Ridge PG, Mukherjee S, Crane PK, Kauwe JSK, Alzheimer's Disease Genetics Consortium (2013) Alzheimer's disease: Analyzing the missing heritability. *PLoS One* **8**, e79771.
- [12] Ruiz A, Heilmann S, Becker T, Hernández I, Wagner H, Thelen M, Mauléon A, Bellenguez C (2014) Follow-up of loci from the International Genomics of Alzheimer's Disease Project identifies TRIP4 as a novel susceptibility gene. *Transl Psychiatry* **4**, e358.
- [13] Ridge PG, Hoyt KB, Boehme K, Mukherjee S, Crane PK, Haines JL, Mayeux R, Farrer LA, Pericak-Vance MA, Schellenberg GD, Kauwe JSK (2016) Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol Aging* **41**, 200e13-200e20.
- [14] Gatz M, Pedersen NL, Berg S, Johansson B, Johansson K, Mortimer J, a Posner SF, Viitanen M, Winblad B, Ahlborn A (1997) Heritability for Alzheimer's disease: The study of dementia in Swedish twins. *J Gerontol A Biol Sci Med Sci* **52A**, M117-M125.
- [15] Pedersen NL, Gatz M, Berg S, Johansson B (2004) How heritable is Alzheimer's disease late in life? Findings from Swedish twins. *Ann Neurol* **55**, 180-185.
- [16] Dudbridge F (2013) Power and predictive accuracy of polygenic risk scores. *PLoS Genet* **9**, e1003348.
- [17] Cruchaga C, Kauwe JSK, Harari O, Jin SC, Cai Y, Karch CM, Benitez BA, Jeng AT, Skorupa T, Carrell D, Bertelsen S, Bailey M, McKean D, Shulman JM, De Jager PL, Chibnik L, Bennett DA, Arnold SE, Harold D, Sims R, Gerrish A, Williams J, Van Deerlin VM, Lee VMY, Shaw LM, Trojanowski JQ, Haines JL, Mayeux R, Pericak-Vance MA, Farrer LA, Schellenberg GD, Peskind ER, Galasko D, Fagan AM, Holtzman DM, Morris JC, Goate AM, GERAD Consortium, Alzheimer's Disease Neuroimaging Initiative, Alzheimer's Disease Genetic Consortium (2013) GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. *Neuron* **78**, 256-268.
- [18] Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA, Alzheimer's Disease Neuroimaging Initiative (2012) The Association between a polygenic Alzheimer score and cortical thickness in clinically normal subjects. *Cereb Cortex* **22**, 2653-2661.
- [19] Sleegers K, Bettens K, De Rooek A, Van Cauwenberghe C, Cuyvers E, Verheijen J, Struyfs H, Van Dongen J, Vermeulen S, Engelborghs S, Vandenbulcke M, Vandenbergh R, De Deyn PP, Van Broeckhoven C, BELNEU Consortium (2015) A 22-single nucleotide polymorphism Alzheimer risk score correlates with family history onset age and cerebrospinal fluid A β 42. *Alzheimers Dement* **11**, 1452-1460.
- [20] Visser PJ, Verhey FRJ, Boada M, Bullock R, De Deyn PP, Frisoni GB, Fratiglioni L, Hampel H, Jolles J, Jones R, Minthon L, Nobili F, Olde Rikkert M, Ousset PJ, Rigaud AS, Scheltens P, Soininen H, Spira L, Touchon J, Tsolaki M, Vellas B, Wahlund LO, Wilcock G, Winblad B (2008) Development of Screening Guidelines and Clinical Criteria for Predementia Alzheimer's Disease. *Neuroepidemiology* **30**, 254-265.
- [21] Barnett JH, Blackwell A, Scheltens P, Waldemar G, Johannsen P, Tsolaki M, Vandenbergh R, Wahlund LO, Verhey F, Visser PJ (2010) Cognitive function and cognitive change in dementia mild cognitive impairment and healthy aging: The EDAR study. *Alzheimers Dement* **6**, S127.
- [22] Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund LO, Freund-Levi Y, Tsolaki M, Minthon L, Wallin ÅK, Hampel H, Bürger K, Pirttila T, Soininen H, Rikkert MO, Verbeek MM, Spira L, Blennow K (2009) Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: A prospective cohort study. *Lancet Neurol* **8**, 619-627.
- [23] Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S, Risacher SL, Nho K, Huentelman MJ, Craig DW, Thompson PM, Stein JL, Moore JH, Farrer LA, Green RC, Bertram L, Jack Jr CR, Weiner MW (2010) Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims progress and plans. *Alzheimers Dement* **6**, 265-273.
- [24] Coleman JRI, Euesden J, Patel H, Folarin AA, Newhouse S, Breen G (2015) Quality control imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Brief Funct Genomics* **15**, 298-304.
- [25] Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flück P, Gabriel SB, Gibbs RA, Green ED, Hurles ME, Knoppers BM, Korbel JO, Lander ES, Lee C, Lehrach H, Mardis ER, Marth GT, McVean GA, Nickerson DA, Schmidt JP, Sherry ST, Wang J, Wilson RK, Boerwinkle E, Doddapaneni H, Han Y, Korchina V, Kovar C, Lee S, Muzny D, Reid JG, Zhu Y, Chang Y, Feng Q, Fang X, Guo X, Jian M, Jiang H, Jin X, Lan T, Li G, Li J, Li Y, Liu S, Liu X, Lu Y, Ma X, Tang M, Wang B, Wang G, Wu H, Wu R, Xu X, Yin Y, Zhang D, Zhang W, Zhao J, Zhao M, Zheng X, Gupta N, Gharani N, Toji LH, Gerry NP, Resch AM, Barker J, Clarke L, Gil L, Hunt SE, Kelman G, Kulesha E, Leinonen R, McLaren WM, Radhakrishnan R, Roa A, Smirnov D, Smith RE, Streeper I, Thormann A, Toneva I, Vaughan B, Zheng-Bradley X, Grocock R, Humphray S, James T, Kingsbury Z, Sudbrak R, Albrecht MW, Amstislavskiy VS, Borodina TA, Lienhard M, Mertes F, Sultan M, Timmermann B, Yaspo ML, Fulton L, Fulton R, Ananiev V, Belaia Z, Beloslyudtsev D, Bouk N, Chen C, Church D, Cohen R, Cook C, Garner J, Hefferon T, Kimelman M, Liu C, Lopez J, Meric P, O'Sullivan C, Ostapchuk Y, Phan L, Ponomarev S, Schneider V, Shekhtman E, Sirotkin K, Slotta D, Zhang H, Balasubramaniam S, Burton J, Danecek P, Keane TM, Kolb-Kococinski A, McCarthy S, Stalker J, Quail M, Davies CJ, Gollub J, Webster T, Wong B, Zhan Y, Campbell CL, Kong Y, Marcketta A, Yu F, Antunes L, Bainbridge M, Sabo

- A, Huang Z, Coin LJM, Fang L, Li Q, Li Z, Lin H, Liu B, Luo R, Shao H, Xie Y, Ye C, Yu C, Zhang F, Zheng H, Zhu H, Alkan C, Dal E, Kahveci F, Garrison EP, Kural D, Lee WP, Fung Leong W, Stromberg M, Ward AN, Wu J, Zhang M, Daly MJ, DePristo MA, Handsaker RE, Banks E, Bhatia G, del Angel G, Genovese G, Li H, Kashin S, McCarroll SA, Nemesh JC, Poplin RE, Yoon SC, Lihm J, Makarov V, Gottipati S, Keinan A, Rodriguez-Flores JL, Rausch T, Fritz MH, Stütz AM, Beal K, Datta A, Herrero J, Ritchie GRS, Zerbino D, Sabeti PC, Shlyakhter I, Schaffner SF, Vitti J, Cooper DN, Ball EV, Stenson PD, Barnes B, Bauer M, Keira Cheatham R, Cox A, Eberle M, Kahn S, Murray L, Peden J, Shaw R, Kenny EE, Batzer MA, Konkel MK, Walker JA, MacArthur DG, Lek M, Herwig R, Ding L, Koboldt DC, Larson D, Ye K, Gravel S, Swaroop A, Chew E, Lappalainen T, Erlich Y, Gymrek M, Frederick Willems T, Simpson JT, Shriver MD, Rosenfeld JA, Bustamante CD, Montgomery SB, De La Vega FM, Byrnes JK, Carroll AW, DeGortor MK, Lacroute P, Maples BK, Martin AR, Moreno-Estrada A, Shringarpure SS, Zakharia F, Halperin E, Baran Y, Cerveira E, Hwang J, Malhotra A, Plewczynski D, Radew K, Romanovitch M, Zhang C, Hyland FCL, Craig DW, Christoforides A, Homer N, Izatt T, Kurdoglu AA, Sinari SA, Squire K, Xiao C, Sebat J, Antaki D, Gujral M, Noor A, Ye K, Burchard EG, Hernandez RD, Gignoux CR, Haussler D, Katzman SJ, James Kent W, Howie B, Ruiz-Linares A, Dermitzakis ET, Devine SE, Min Kang H, Kidd JM, Blackwell T, Caron S, Chen W, Emery S, Fritsche L, Fuchsberger C, Jun G, Li B, Lyons R, Scheller C, Sidore C, Song S, Sliwerska E, Taliun D, Tan A, Welch R, Kate Wing M, Zhan X, Awadalla P, Hodgkinson A, Li Y, Shi X, Quitadamo A, Lunter G, Marchini JL, Myers S, Churchhouse C, Delaneau O, Gupta-Hinch A, Kretzschmar W, Iqbal Z, Mathieson I, Menelaou A, Rimmer A, Xifara DK, Oleksyk TK, Fu Y, Liu X, Xiong M, Jorde L, Witherspoon D, Xing J, Browning BL, Browning SR, Hormozdiari F, Sudmant PH, Khurana E, Tyler-Smith C, Albers CA, Ayub Q, Chen Y, Colonna V, Jostins L, Walter K, Xue Y, Gerstein MB, Abyzov A, Balasubramanian S, Chen J, Clarke D, Fu Y, Harmanci AO, Jin M, Lee D, Liu J, Jasmine Mu X, Zhang J, Zhang Y, Hartl C, Shakir K, Degenhardt J, Meiers S, Raeder B, Paolo Casale F, Stegle O, Lameijer EW, Hall I, Bafna V, Michaelson J, Gardner EJ, Mills RE, Dayama G, Chen K, Fan X, Chong Z, Chen T, Chaisson MJ, Huddleston J, Malig M, Nelson BJ, Parrish NF, Blackburne B, Lindsay SJ, Ning Z, Zhang Y, Lam H, Sisu C, Challis D, Evani US, Lu J, Nagaswamy U, Yu J, Li W, Habegger L, Yu H, Cunningham F, Dunham I, Lage K, Berg Jerspersen J, Horn H, Kim D, Desalle R, Narechania A, Wilson Sayres MA, Mendez FL, David Poznik G, Underhill PA, Coin L, Mittelman D, Banerjee R, Cerezo M, Fitzgerald TW, Louzada S, Massaia A, Ritchie GR, Yang F, Kalra D, Hale W, Dan X, Barnes KC, Beiswanger C, Cai H, Cao H, Henn B, Jones D, Kaye JS, Kent A, Kerasidou A, Mathias R, Ossorio PN, Parker M, Rotimi CN, Royal CD, Sandoval K, Su Y, Tian Z, Tishkoff S, Via M, Wang Y, Yang H, Yang L, Zhu J, Bodmer W, Bedoya G, Cai Z, Gao Y, Chu J, Peltonen L, Garcia-Montero A, Orfao A, Dutil J, Martinez-Cruzado JC, Mathias RA, Hennis A, Watson H, McKenzie C, Qadri F, LaRocque R, Deng X, Asogun D, Folarin O, Happi C, Omoniwa O, Strelau M, Tariyal R, Jallow M, Sisay Joof F, Corrah T, Rockett K, Kwiatkowski D, Kooner J, Tnh Hiê'n T, Dunstan SJ, Thuy Hang N, Fonnier R, Garry R, Kanneh L, Moses L, Schieffelin J, Grant DS, Gallo C, Poletti G, Saleheen D, Rasheed A, Brooks LD, Felsenfeld AL, McEwen JE, Vaydylevich Y, Duncanson A, Dunn M, Schloss JA (2015) A global reference for human genetic variation. *Nature* **526**, 68-74.
- [26] R Core Team (2014) R: A language and environment for statistical computing.
- [27] Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, R Core Team, Benesty M, Lescarbeau R, Ziem A, Scrucchi L (2015) caret: Classification and Regression Training. R Package Version 6.0-41, <http://CRAN.R-project.org/package=caret>
- [28] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77.
- [29] Sing T, Sander O, Beerenwinkel N, Lengauer T (2005) ROCr: Visualizing classifier performance in R. *Bioinformatics* **21**, 3940-3941.
- [30] Euesden J, Lewis CM, O'Reilly PF (2015) PRSice: Polygenic risk score software. *Bioinformatics* **31**, 1466-1468.
- [31] Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, Badarinarayan N, Morgan K, Passmore P, Holmes C, Powell J, Brayne C, Gill M, Mead S, Goate A, Cruchaga C, Lambert J C, Van Duijn C, Maier W, Ramirez A, Holmans P, Jones L, Hardy J, Seshadri S, Schellenberg GD, Amouyel P, Williams J (2015) Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* **138**, 3673-3684.
- [32] Baird AL, Westwood S, Lovestone S (2015) Blood-based proteomic biomarkers of Alzheimer's disease pathology. *Front Neurol* **6**, 236.
- [33] Voyle N, Kim M, Proitsi P, Ashton NJ, Baird AL, Bazenot C, Hye A, Westwood S, Chung R, Ward M, Rabinovici GD, Lovestone S, Breen G, Legido-Quigley C, Dobson RJB, Kiddle SJ (2016) Blood metabolite markers of neocortical amyloid- β burden: Discovery and enrichment using candidate proteins. *Transl Psychiatry* **6**, e719.
- [34] Mattsson N, Zetterberg H, Blennow K (2010) Lessons from multicenter studies on CSF biomarkers for Alzheimer's disease. *Int J Alzheimers Dis* **2010**, pii:610613.
- [35] García Barrado L, Coart E, Vanderstichele HM, Burzykowski T (2015) Transferring cut-off values between assays for cerebrospinal fluid Alzheimer's disease biomarkers. *J Alzheimers Dis* **49**, 187-199.
- [36] Zwan MD, Rinne JO, Hasselbalch SG, Nordberg A, Lleo A, Herukka SK, Soininen H, Law I, Bahl JMC, Carter SF, Fortea J, Blesa R, Teunissen CE, Bouwman FH, van Berckel BNM, Visser PJ (2015) Use of amyloid-PET to determine cutpoints for CSF markers: A multicenter study. *Am Acad Neurol* **86**, 50-58.
- [37] Wheeler HE, Aquino-Michaels K, Gamazon ER, Trubetskoy VV, Dolan ME, Huang RS, Cox NJ, Im HK (2014) Poly-omic prediction of complex traits: OmicKriging. *Genet Epidemiol* **38**, 402-415.