Plasma biomarkers and genetics in the diagnosis and prediction of Alzheimer's disease

Joshua Stevenson-Hoare,[1] Amanda Heslegrave,[2,3] Ganna Leonenko,[1] Dina Fathalla,[1] Eftychia Bellou,[1] Lauren Luckcuck,[1] Rachel Marshall,[4] Rebecca Sims,[4] Bryan Paul Morgan,[1] John Hardy,[2,3] Bart de Strooper,[2] Julie Williams,[1] Henrik Zetterberg[2,3,5,6,7] and Valentina Escott-Price[1,4]

1 Dementia Research Institute, Cardiff University, UK

2 Dementia Research Institute, University College London, UK

3 Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London, UK

4 Division of Neuroscience and Mental Health, Cardiff University, UK

5 Hong Kong Center for Neurodegenerative Diseases, Hong Kong, China

6 Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden

7 Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden

Correspondence to: Valentina Escott-Price

Dementia Research Institute, Hadyn Ellis Building, Maindy Rd, Cardiff, CF24 4HQ, UK

E-mail: escottpricev@cardiff.ac.uk

Correspondence may also be addressed to: Henrik Zetterberg

Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden

E-mail: henrik.zetterberg@clinchem.gu.se

Running title: Plasma biomarkers and genetics in AD

Abstract

Plasma biomarkers for Alzheimer's disease-related pathologies have undergone rapid developments during the past few years, and there are now well-validated blood tests for amyloid and tau pathology, as well as neurodegeneration and astrocytic activation. To define Alzheimer's disease with biomarkers rather than clinical assessment, we assessed prediction of research-diagnosed disease status using these biomarkers and tested genetic variants associated with the biomarkers that may reflect more accurately the risk of biochemically defined Alzheimer's disease instead of the risk of dementia.

In a cohort of Alzheimer's disease cases (N=1439, mean age 68 years [SD=8.2]) and screened controls (N=508, mean age 82 years [SD=6.8]), we measured plasma concentrations of the 40 and 42 amino acid-long amyloid β fragments (Aβ40 and Aβ42, respectively), tau phosphorylated at amino acid 181 (P-tau181), neurofilament light (NfL), and glial fibrillary acidic protein (GFAP) using state-of-the-art Single molecule array (Simoa) technology. We tested the relationships between the biomarkers and Alzheimer's disease genetic risk, age at onset, and disease duration. We also conducted a genome-wide association study for association of disease risk genes with these biomarkers.

The prediction accuracy of Alzheimer's disease clinical diagnosis by the combination of all biomarkers, *APOE* and polygenic risk score reached AUC=0.81, with the most significant contributors being ε4, Aβ40 or Aβ42, GFAP and NfL. All biomarkers were significantly associated with age in cases and controls ($p<4.3x10^{-5}$). Concentrations of the Aβ-related biomarkers in plasma were significantly lower in cases compared with controls, whereas other biomarker levels were significantly higher in cases.

In the case-control genome-wide analyses, *APOE*-ε4 was associated with all biomarkers (p=0.011-$4.78x10^{-8}$), except NfL. No novel genome-wide significant SNPs were found in the case-control design; however, in a case-only analysis, we found two independent genome-wide significant associations between the Aβ42/Aβ40 ratio and *WWOX* and *COPG2* genes.

Disease prediction modelling by the combination of all biomarkers indicates that the variance attributed to P-tau181 is mostly captured by *APOE*-ε4, whereas Aβ40, Aβ42, GFAP and NfL biomarkers explain additional variation over and above *APOE*. We identified novel plausible genome wide-significant genes associated with Aβ42/Aβ40 ratio in a sample which is fifty times smaller than current genome-wide association studies in Alzheimer's disease.

Keywords: Plasma biomarkers; genome-wide association study; Alzheimer's disease

Abbreviations: Aβ = Amyloid beta; GFAP = glial fibrillary acidic protein; GWAS = genome wide association study; HWE = Hardy-Weinberg equilibrium; LD = linkage disequilibrium; MAF = minor allele frequency; MCI = mild cognitive impairment; MMSE = Mini Mental State Examination; NfL = neurofilament light chain; PC = principal component; PRS = polygenic risk score; P-tau = phosphorylated tau; SNP = single nucleotide polymorphism

Introduction

Alzheimer's disease is one of the greatest health challenges, affecting tens of millions of people worldwide. The clinical diagnosis of this disease is, however, often inaccurate; around 25% of people with clinical Alzheimer's disease do not have underlying pathology at autopsy, and many people who have not yet developed Alzheimer's disease-type dementia have incipient pathology, the prevalence of which increases with age[1]. Detecting Alzheimer's disease at the earliest possible stage remains essential to combating its effects and to further our understanding of this devastating illness. By diagnosing early, we can better understand how the disease progresses, plan and implement treatments earlier, and monitor response to drugs currently being trialled.

Aβ and tau pathology are the defining pathological features of Alzheimer's disease[2]. For many years, it has been possible to detect Alzheimer's disease pathology (amyloid aggregation, tau tangles and neurodegeneration) using imaging and cerebrospinal fluid (CSF) biomarkers. Although CSF and PET biomarkers of amyloid β and tau are highly accurate for detecting disease pathology[3], the costs, invasive nature, and low availability of the tools needed to detect these biomarkers hamper their feasibility for use in clinical diagnostic practice and for screening in clinical trials.

Assays for plasma Aβ fragments (ratio of amyloid $\beta_{1\text{-}42}$ (Aβ42) to amyloid $\beta_{1\text{-}40}$ (Aβ40)) reflect brain amyloidosis[4–7]; however, these assays have limitations, including the impact of substantial peripheral amyloid β production[8]. By contrast, CSF and plasma tau phosphorylated at threonine 181 (P-tau181) is a highly specific pathological marker of Alzheimer's disease that remains normal in other dementias[9,10]. GFAP and NfL are putative non-amyloid plasma-based biomarkers indicative of ongoing neuroinflammatory and neurodegenerative disease processes. Increased GFAP suggests abnormal activation and proliferation of astrocytes, for instance secondary to neuronal damage. It has been shown that GFAP levels in plasma and CSF are higher in Alzheimer's disease and correlate with cognitive impairment[11–13]. Plasma NfL is a marker of neuronal injury, increased in Alzheimer's disease[14], but this biomarker has low specificity, because increases are also reported in several other neurodegenerative disorders[13,15,16]. Thus, while NfL has potential as a monitoring biomarker, GFAP might be a valuable prognostic biomarker, predicting incident dementia[13]. Recent reports show that plasma P-tau181 concentration starts to increase around 15 years prior to clinical disease onset in familial Alzheimer's disease[17], and that plasma P-tau181 predicts disease neuropathology at least eight years prior to autopsy in sporadic disease[10].

Early disease prediction can be helped with genetic data as an individual's genetic makeup does not change over time and genetic data are precise and inexpensive to measure, however, the prediction accuracy by genetics is limited[18]. Biomarkers, in contrast to genetics, can only indicate the presence of Alzheimer's disease pathology after the disease has already been triggered, *i.e.*, a biomarker change marks the onset of a pathological process. Nevertheless, the prediction accuracy of, *e.g.*, P-tau181 and P-tau217 for discriminating Alzheimer's disease from other neurodegenerative diseases[19–21], when combined with *APOE* genotype, memory and executive function phenotypes, was reported to reach

AUC>90% in predicting the progression from mild cognitive impairment (MCI) to Alzheimer's disease in two relatively small samples of participants (N=340 and 543)[22].

Identifying genetic loci associated with biomarkers could aid understanding of the specific pathophysiological components underpinning these biomarkers. Genome-wide association studies (GWAS) of CSF biomarkers in AD case/control samples have found loci in genes *GEMC1* and *OSTN*[23] as well as more commonly reported loci such as the *TREM* cluster, *APOE*, *APOC*, and *TOMM40*[24]. However, these have also only focussed on small sets of biomarkers, typically P-tau181 and Aβ42. GWAS of blood plasma P-tau181 and NfL levels[25,26] have identified only loci within the *APOE* genomic region, and only for P-tau181. Investigation of the relationship between Alzheimer's disease PRS and plasma P-tau181[27] has revealed highly significant associations with PRS containing the *APOE* region ($p = 3x10^{-18}$-$7x10^{-15}$), and moderate association when *APOE* was excluded. GWAS studies for plasma Aβ40, Aβ42, and Aβ42/40 ratio in non-demented participants from population-based studies have identified GWAS significant variants in *APOE* and *BACE1* genes, and *APP*, *PSEN2*, *CCK*, and *ZNF397* genes in gene-based analysis[28].

The aims of this study are 1) to test the prediction ability of the biomarkers for clinical AD diagnosis in our cohort (over and above commonly used predictors such as *APOE*, age and AD PRS), and 2) to identify genetic loci associated with these plasma biomarkers. The latter may shed light on which SNPs associated with clinical Alzheimer's disease are also associated with plasma biomarkers. This could help to further refine the relevance of the AD GWAS genes to different biological processes, which the biomarkers represent. To that end, we measured plasma biomarkers in a sample of 1,439 early and late onset Alzheimer's disease cases (mean age 68 years [SD=8.0]) and 508 elderly screened controls (mean age 82 years [SD=6.7]). We used ultrasensitive Single molecule array (Simoa) assays to measure P-tau181, NfL, GFAP, Aβ40, Aβ42, and calculated the ratio of Aβ42/40. We then tested these biomarkers for association with the clinical diagnosis of AD and, in case samples, the relationship of the biomarkers with age at sample collection, age at onset and disease duration. To identify genetic loci associated with these biomarkers, we undertook a GWAS for P-tau181, NfL, Aβ40, Aβ42, ratio of Aβ42/40 and GFAP biomarkers in the largest case-control sample set to date.

Materials and methods

Alzheimer's Disease Cardiff Cohort

The Alzheimer's Disease Cardiff Cohort (ADCC) was collected between 2004 and 2020 using MRC, Moondance Foundation, and Health and Care Research Wales (HCRW) funding. The cohort collection used a standardised clinical and comprehensive neuropsychological assessment (validated by Holmes *et al.*[29]), see more details in Supplementary Section 1. AD diagnosis was not supported by any biochemical or imaging measures (*e.g.*, CSF or PET) due to the funds allocated to the study collecting the data.

We used plasma samples collected from 1,439 early and late onset sporadic Alzheimer's disease cases and 508 screened elderly controls. Information on age at assessment, sex, *APOE* genotype and genome-wide array genotyping was available for all 1947 samples. Within cases, information was also available

for N=1319 individuals on age at onset, and duration of disease was calculated for these samples. Details of the sample demographics are in Table 1.

Biomarkers

Biomarkers were tested for 1986 individual plasma samples from the ADCC. P-tau181 concentration was measured using the Simoa P-tau181 Advantage Kit, whilst Aβ40, Aβ42, NfL and GFAP concentrations were measured using the Simoa Human Neurology 4-Plex E (N4PE) assay (Quanterix, Billerica, MA). The measurements were performed in one round of experiments using one batch of reagents with the analysts blinded to diagnosis and clinical data. All measurements for all 5 analytes were above the limit of detection of the assays. Intra-assay coefficients of variation were below 10%. These data were then matched to phenotype information. Thirty-nine samples were removed at this stage based on missing/mismatching data for age and gender or due to ID duplication, leaving 1947 individuals for further analysis. Samples were excluded for each biomarker analysis on a case-by-case basis, based on outlier thresholds calculated using Median Absolute Deviation (MAD)[30]. This method is more robust to remote outliers than the mean and SD method, and copes better with skewed data due to its reliance on non-parametric measures of central tendency and variation. Pearson's correlations between biomarkers were calculated for the 1735 samples which had no outlier measurements for any biomarker. Details of biomarker distributions are in Table 1.

Genetics

Individuals for this analysis were included if both genetic and biomarker information were available, totalling 1,947 individuals in the final dataset. All individuals had information available on *APOE* genotype (ε2ε2 = 8, ε2ε3 = 145, ε2ε4 = 33, ε3ε3 =844, ε3ε4 = 620, ε4ε4 = 239. Quality control (QC) of the genetic data was performed for cases and controls together, the QC steps used are reported elsewhere[31,32] and in Supplementary Section 2. Genotyped data were aligned to human genome assembly GRCh37/hg19 and imputed via Michigan Imputation server using Minimac3[33] with the Haplotype Reference Consortium (HRC)[34] reference panel. Post-imputation QC used thresholds of MAF<5%, poor accuracy of imputation (INFO)<0.8, MISS>5%, and HWE $p \leq 10^{-6}$. This resulted in a final dataset containing 4,618,496 variants.

Statistical analysis

The association of biomarkers with age at onset and disease duration in cases, and with age at interview in cases and in controls (separately), was tested with linear regression where the biomarker was the

outcome variable, controlling for sex. For all following analyses the biomarkers were adjusted for age and standardised to have a mean of zero and standard deviation of one. The correlations between the biomarkers were assessed with Pearson's correlation.

The association of Alzheimer's disease case/control status by the biomarkers was tested using logistic regression, accounting for sex, *APOE* and PRS without the *APOE* region (chromosome 19:44.4-46.5Mb) using the glm() function in R. The most parsimonious model was derived with the backwards stepwise approach (step() function in R). The prediction accuracy was assessed by means of the area under the receiver operation curve (AUC), using auc() function in R.

The *APOE* region was represented by the number of ε2 and ε4 alleles which we used as two predictor variables. The PRS without *APOE* region (PRSnoAPOE) was used to account for the remaining genetic effect. For the PRS calculation we used the summary statistics from the largest *clinically assessed* late-onset case-control GWAS study on Alzheimer's disease available at the time of analysis (N=63,926)[35]. PRS were generated with the PLINK genetic data analysis toolset[36 36] for *p*-value threshold $p \leq 0.1$ on LD-clumped SNPs by retaining the SNP with the smallest *p*-value excluding variants with $r^2>0.1$ in a 1000-kb window, see details in[37]. Prior to analyses PRSnoAPOE was adjusted for five principal components and then standardised.

All statistical analyses were performed in R-statistical software (https://www.R-project.org/). The plots were generated using the *ggplot2* package with custom scripts generated in house.

The results of the biomarkers' association with the clinical/demographic characteristics are presented without correction for multiple testing, since these analyses are hypothesis-driven.

Genetic analysis

SNP-based association analyses were performed for each biomarker using linear regression model with PLINK. Association analyses of SNPs with the biomarkers were adjusted for age and sex, five principal components (PCs) and case-control status ("caseness"). The adjustment for caseness was introduced to reduce the variation due to potential differences in association pattern of biomarkers between cases and controls, whilst using all available samples to maintain the statistical power. In addition, association analyses for cases and controls were also conducted separately. Since the *APOE* region is not well covered by the Illumina arrays used to genotype the ADCC dataset, we tested association of the biomarkers with the number of directly genotyped *APOE*-ε4 alleles. PCs were computed using PLINK and the number of PCs was determined via visual inspection of the pairwise PC scatter plots. The GWAS significance level was set to the commonly accepted $p<5\times10^{-8}$. We did not further adjust this for the six biomarkers as the biomarker levels were measured in the same sample and are not independent.

To investigate further the variants of interest, we used Combined Annotation-Dependent Depletion (CADD) and RegulomeDB (RDB) scores for SNPs accessible within the Functional mapping and annotation of genetic associations (FUMA) on-line tool[38]. CADD is a tool for scoring the deleteriousness of single nucleotide variants as well as insertion/deletions variants in the human genome[39,40]. RDB[41] is a categorical score from 1a to 7 representing regulatory functionality of SNPs based on eQTLs and

chromatin marks. 1a is the highest score, indicating that the SNP has the most biological evidence to be a regulatory element.

We compared our GWAS biomarker association results to Alzheimer's disease genome-wide significant findings[35], assessing all SNPs in the ADCC GWAS within ±20kB of the GWAS-significant SNPs. The replication significance level was set to nominal significance level $p<0.05$.

To summarise the association results from all variants in a gene, accounting for number of variants and linkage disequilibrium (LD) between them, we used Multi-marker Analysis of GenoMic Annotation (MAGMA, v1.09b)[42]. For the gene-based analysis, we mapped a SNP to a gene (as defined by NCBI 37.3) if it resided within the gene boundaries. The LD between SNPs was estimated with the European reference panel in 1000 Genomes phase 3. The significance level for the gene-based analysis results was set to the commonly accepted $p<2.5\text{x}10^{-6}$.

For the pathway analyses, 10,271 gene sets were downloaded from Reactome, Biocarta, KEGG and Pathway Interaction Databases[32]. The pathway analyses were performed using the "competitive" option in MAGMA, assessing whether the genes in a gene set are more strongly associated with the phenotype than in other gene sets in the genome. We adopted the false discovery rate (FDR≤0.05) approach (p.adjust() function in R with method="fdr") to correct for multiple testing the results of the pathway analyses.

Data availability

GWAS summary statistics for the top results ($p\leq1\text{x}10^{-5}$) are listed in the main text of the paper and Supplementary Materials. Full GWAS summary statistics are available from the authors upon request.

Results

Biomarker results in relation to Alzheimer's disease, age at onset and disease duration

The correlation pattern between the biomarkers was similar for cases and controls, and agree with the results of Cullen *et al*.[43]. The correlation between Aβ42 and Aβ40 values was high (r=0.8 in cases and 0.7 in controls, $p<10^{-16}$). The lowest correlation was observed between P-tau181 and Aβ-related biomarkers, see Figure 1.

To assess whether the disease stage is captured by the biomarkers, we explored the relationship between biomarkers, age of onset and disease duration in cases. Table 2 summarises the results. In this case-only analysis, age at onset was strongly positively associated with Aβ40, Aβ42, GFAP and NfL (p-values $\leq 4.2\text{x}10^{-23}$), moderately with P-tau181 (p=0.0023), and negatively associated with Aβ42/Aβ40 ($p=4.8\text{x}10^{-4}$). The biomarkers GFAP, NfL and P-tau181 show significant increase in females as compared to males ($p=9.0\text{x}10^{-23}$, $1.4\text{x}10^{-7}$, and $2.1\text{x}10^{-8}$, respectively). This in part replicates the finding in Kumar-Singh *et al*.[44], who showed that age-of-onset of *PSEN1*-linked familial Alzheimer's disease correlated

negatively with Aβ42/Aβ40 but positively with Aβ40 levels. Longer disease duration was strongly associated with elevated levels of GFAP and NfL (p=2.9x10^{-6} and 1.2x10^{-12}, respectively) and moderately associated with increase of Aβ40 and P-tau181 levels (p=0.027 and 0.008, respectively).

In controls, all biomarkers were positively associated with age at interview (p-value ranked between 1.2x10^{-7} for Aβ42 and 1.9x10^{-30} for NfL), and negatively with the ratio Aβ42/Aβ40 (p=1.2x10^{-10}) (see Table 3), indicating that all biomarkers are sensitive to age, and will show less discrimination between AD cases and controls if AD cases with earlier onset (~65-68 years) are compared with elderly screened controls (see Supplementary Figure 1).

Next, we assessed the prediction accuracy of disease status in our sample. The prediction accuracy of the case-control status by sex and *APOE* genotype resulted in AUC=0.74 and R^2=0.21. All biomarkers were significantly associated with Alzheimer's disease status when tested separately (Table 4, Supplementary Figure 2). The prediction accuracies, however, were moderate with the highest prediction accuracy AUC=0.66 and 0.65 for Aβ42 and P-tau181, respectively.

The prediction accuracy of a model combining all biomarkers and genetics (*APOE*-ε4, *APOE*-ε2, PRS without *APOE* region) was AUC=0.81, R^2=0.29. The most parsimonious model that predicted the outcome with the same accuracy as above (derived using stepwise regression) included all predictors except Aβ42 and P-tau181 (*APOE*-ε4 B=1.3, p=2.02x10^{-24}; *APOE*-ε2 B=-0.45, p=0.011; PRSnoAPOE B=0.14, p=0.033; Aβ40 B=-0.62, p=6.6x10^{-18}; GFAP, B=0.29, p=3.9x10^{-4}; NfL B=0.45, p=4.6x10^{-8}; Aβ42/Aβ40 B=-0.20, p=0.003).

This model highlights the importance of all genetic predictors and the Aβ40, GFAP, and NfL biomarkers. The variance of Aβ42 was captured by Aβ40, as the correlation between these biomarkers was high. Indeed, when Aβ40 was dropped from the model, then Aβ42 became a significant predictor (B=-0.59, p=9.6x10^{-12}). In both models, the ratio of Aβ42/Aβ40 was significant but it changed its direction of effect depending on which marker was included (B=0.20, p=0.005, and B=-0.20, p=0.003, when Aβ42 or Aβ40 was included, respectively) P-tau181 was dropped from the model by the stepwise regression, however this should not be interpreted as P-tau181 being fully explained by the genetic predictors. In a model with only P-tau181 and genetics (APOE-ε4, APOE-ε2, PRSnoAPOE), P-tau181 remained highly significant over and above genetics (B=0.38, p=4.5x10^{-8}).

The model with all biomarkers but without genetic predictors had an accuracy of AUC=0.75 and explained variance of R^2=0.18. In this model, the same biomarkers as above showed significant association, with the addition of the P-tau181 biomarker (B=0.18, p=0.022), indicating that the P-tau181 signal may be explained by genetics, whereas the other significant biomarkers (Aβ-related, GFAP, and NfL) add to the prediction over and above genetics.

Genome-wide association study

We performed three sets of GWAS (cases only, controls only, all samples) in ADCC with the 5 biomarkers (Aβ40, Aβ42, NfL, P-tau181, GFAP) and the Aβ42/Aβ40 ratio as outcome measures. The top SNPs with an association p-value≤1x10^{-5} are presented in Supplementary Tables 1-6. In the case-control analysis,

APOE-ε4 was associated with all biomarkers (p=0.011 - 4.78x10^{-8}, Supplemental Tables 1-2,4-6), except NfL (in Supplemental Table 3).

We compared the GWAS we performed for biomarkers to the genome-wide significant SNPs from a large clinically assessed Alzheimer's disease GWAS study[35], see Supplemental Table 7. The strongest associations for the GWAS index *APOE* SNP (rs429358) were for P-tau181 and GFAP (p=0.001 and 0.002, respectively, Supplemental Table 7). Interestingly, SNPs in or near the *WWOX* gene were at least nominally associated with all biomarkers. The strongest association was found for GFAP (p=1.2x10^{-5}) for a SNP situated 2.7KB away from the GWAS index *WWOX* SNP.

The GWAS of the five biomarkers and the Aβ42/Aβ40 ratio in controls only and in all samples did not reveal any genome-wide significant loci. In the cases only GWAS, however, we observed two genome-wide significant loci for the Aβ42/Aβ40 ratio (Supplementary Table 6 and Figure 2). The lead SNPs for these loci lie within the intronic region of their respective genes (*COPG2* and *WWOX*), with the WWOX variant predicted to function as an enhancer.

The first genome-wide significant locus was a high LD region on chromosome 7 spanning from 130.2-130.4Mb and covering genes *COPG2* (chr7:130,146,080-130,353,598) and *TSGA13* (chr7:130,353,486-130,371,406) with the lead SNP rs17165066, (chr7:130,370,267, B=0.15, SE=0.026, p=8.9x10^{-9}). This SNP tags 50 other SNPs with r^2>0.8; see Manhattan plot (Figure 2) and LocusZoom plot (Figure 3A). Moreover, this region contains two SNPs (rs10264429 and rs375839317, MAF=0.06, 0.07, respectively) which are in high LD with the lead SNP (r^2=0.84 and 0.71, respectively) and have CADD scores = 13.6, 12.48, which are greater than the suggestive threshold for a SNP to be deleterious (CADD>12.37). The rs77696591 (MAF=0.06) intergenic variant is also tagged by the lead SNP (r^2=0.87) and has an RDB score=3a, *i.e.*, has "putatively functional impact on gene regulation". The lead SNP rs17165066 was not statistically significant in the clinically assessed AD GWAS[35].

The second genome-wide significant region was on chromosome 16 in the *WWOX* gene (chr16:78,133,327-79,246,564), that has also been linked to Alzheimer's disease by GWAS[35]. The lead SNP rs34946778 (chr16:78989116, B=0.15, SE=0.026, p=4.36x10^{-9}) was not statistically significant in the AD GWAS[35]. The linkage disequilibrium was r^2=0.0014 between the AD GWAS lead SNP (rs62039712) and the SNP identified in our study (rs34946778).

Finally, the number of *APOE*-ε4 alleles was associated with Aβ40 (B=-0.072, p=1.1x10^{-2}), Aβ42 (B=-0.015, p=6.3x10^{-7}), Aβ42/Aβ40 (B=-0.15, p=1.05x10^{-5}), GFAP (B=0.1, p=1.3x10^{-3}) and P-tau181 (B=0.18, p=4.7x10^{-8}), but not with NfL (p=0.40).

Discussion

We demonstrated that the prediction accuracy for Alzheimer's disease status by the combination of blood biomarkers, sex, *APOE* and PRS reaches AUC=0.81 (R^2=0.29) with the most significant contributors being *APOE*-ε4, Aβ40, and GFAP. This AUC value is lower than that reported in Palmqvist *et al.*[22] likely due to our controls being systematically older than cases, with the diagnostic accuracies for Alzheimer's

disease being decreased with age[45]. Note that Aβ42 becomes a highly significant predictor when Aβ40 is dropped from the model and vice versa, although a stepwise regression recommended dropping Aβ42 over Aβ40. The prediction accuracy by all biomarkers without genetic predictors was AUC=0.75, which is slightly higher than the accuracy by genetic predictors alone (AUC=0.73 in our sample). Interestingly, P-tau181 was not significant if genetic predictors were included in the model and became significant only when no genetic predictors were used, indicating that genetic factors, *APOE*-ε4 in particular, influence plasma P-tau181 levels. However, an advantage of P-tau181 as a biomarker over other predictors (*e.g.*, genetics) is that it is a relatively inexpensive blood biomarker and does not reveal any sensitive genetic information.

In controls, age at interview was positively associated with all biomarkers (*p*-value ranged between $1.2x10^{-7}$ for Aβ42 and $1.9x10^{-30}$ for NfL), and negatively associated with the ratio Aβ42/Aβ40 ($p=1.2x10^{-10}$), indicating that all biomarkers are sensitive to age or pre-clinical age-related neurodegenerative pathologies.

In case-only analyses, age at onset was significantly associated with all biomarkers, in particular, positively with Aβ40, Aβ42, GFAP, NfL and P-tau181 and negatively with the ratio Aβ42/Aβ40. In addition to age at onset, GFAP, NfL and P-tau181 were also associated with the disease duration, with similar effect sizes indicating that the associations can be attributed to age in general, rather than to a particular feature of the disease development and progression. These findings are in line with other recent studies. Chatterjee *et al*.[46] demonstrate that plasma GFAP levels are elevated in cognitively normal older adults at risk of Alzheimer's disease. Aschenbrenner *et al.*[47] conclude that NfL can be used to monitor both cognitive decline due to normal aging and dementia. Lantero Rodriguez *et al.*[10] report that the main increase in plasma P-tau181 occurred between eight and four years prior to death in patients with Alzheimer's disease neuropathology whereas patients without pathology and controls exhibited minor, although significant, increases in P-tau181 up until death.

The Aβ40 and Aβ42 results showing increasing concentration with age in both cases and controls support the earlier finding that Aβ40 and Aβ42 levels are increased before the onset of sporadic Alzheimer's disease[48–50]. It has also been shown that the biomarker distributions are more similar between subjects with and without Alzheimer's disease in elderly subjects than in young subjects[45]. When comparing cases and controls in our sample, we found that cases have lower concentrations of Aβ40 and Aβ42 in plasma, accounting for age. This might indicate that cases, despite early onset, are in the advanced stage of the disease (mean disease duration 5.3 years (SE=3.6) in our sample). An earlier study[50] showed that Aβ40 and Aβ42 levels are elevated in some patients before and during the early stages of Alzheimer's disease but decline thereafter. Our results show similar association patterns (lower Aβ42/Aβ40 is associated with increased age) to the recent report[7]-for participants of all ages and diagnoses who were enrolled in a longitudinal study of memory and aging. Another study[5] which included cognitively normal individuals, patients with mild cognitive impairment and patients with Alzheimer's disease, found no significant correlations between the biomarker values and age. A population-based study[51] reports results in a cohort where all individuals were born in the same week, but blood samples were collected within the testing period of 2.6 years. Within this very limited age range, Aβ42 (but not Aβ40) was significantly positively associated with age. Therefore Aβ42/Aβ40 was

also positively associated with age. In our study with a much wider age range, both Aβ42 and Aβ40 were significantly positively associated with age. The ratio Aβ42/Aβ40 was negatively associated with age because the increase in Aβ40 was greater than that in Aβ42 (Table 4, Supplementary Figure 2). To summarise the Aβ data, the biomarker is sensitive to age and potentially other clinical conditions and phenotypes unmeasured and unaccounted for in our and others' reports. Given this, interpretation of Aβ measurements in the absence of other clinical information is uncertain at best.

In addition, the biomarkers measuring Aβ40, Aβ42 and P-tau181 levels, also have complex trajectories as the disease develops, and this is all in the context of 80% Alzheimer's disease diagnostic accuracy. Counterintuitively, it seems that P-tau181 is largely a plaque amyloid marker[52]: it does not go up in progressive supranuclear palsy, it goes up in amyloid mice after onset of plaque pathology[53] (although it may also increase in tau-overexpressing mice[54]). Aβ, however, goes down when plaque deposition starts and *APOE* correlates with plaque number in a dose-dependent manner[55]. Thus, *APOE* and P-tau181 correlate positively because they both largely mark amyloid deposition. When P-tau181 increases, Aβ42/Aβ40 decreases because Aβ42 sticks to the amyloid plaques, preventing it from leaking into plasma or CSF. An advantage of using the Aβ42/Aβ40 ratio over the individual biomarkers is that the ratio normalises high vs low Aβ producers to each other and is a more reliable qualitative test for Aβ status in the brain than Aβ42 alone.

We found two independent genome-wide significant associations with the ratio of Aβ42/Aβ40 in the *COPG2* and *WWOX* genes in a case-only analysis (the lead SNPs in controls were not-significant). In the analysis, which included both cases and controls, these SNPs were not genome-wide significant despite the increased sample size compared to cases-only. The GWAS SNPs found in cases were not statistically significant in controls and had effect sizes in the opposite direction. This may indicate that there are genetic-protein associations that can only be identified when looking at disease-relevant groups (AD in this case).

COPG2 is a part of the coat protein complex I (COPI) which is responsible for retrograde transport from Golgi-to-endoplasmic reticulum. Genetic modulation of the COPI complex leads to changes in amyloid precursor protein processing and a decrease in the amyloid plaque burden in an Alzheimer's disease mouse model[56].

The WW domain-containing oxidoreductase gene (*WWOX*) maps to the ch16q23.1-23.2 region and encodes a 414-amino acid protein composed of two WW domains in its N-terminus and a central short-chain dehydrogenase/reductase domain[57]. In recent years, abundant evidence from multiple studies has causally linked *WWOX* loss of function with various central nervous system pathologies. *WWOX* dysfunction induced sequential aggregation of tau and amyloid β, and caused apoptosis[58]. The role of *WWOX/WOX1* in Alzheimer's disease pathology and in cell death signalling has previously been reported[59], as has its role in brain development and pathology[60].

In conclusion, our results demonstrate that the currently available plasma biomarkers reflect different aspects of Alzheimer's disease, some of which can be attributed to ageing in addition to the disease-specific features, while others are specifically related to disease progression mechanisms. Our study

shows that biomarker-based diagnosis is not perfect because the biomarker measurements in older controls are similar to those in younger clinically diagnosed Alzheimer's disease cases (which likely represents increased prevalence of pre-clinical Alzheimer's changes in older controls). Biomarkers, however, have the advantage of specificity over clinical assessments, which may confuse dementia subtypes due to phenotypic similarities. Therefore, blood plasma biomarkers can only be a useful tool for the assessment and prediction of Alzheimer's disease in the context of other genetic and/or clinical information. The idea that biomarkers alone might provide more accurate prediction for Alzheimer's disease remains to be fully validated. Longitudinal studies which use a combination of genetics, plasma biomarkers, brain imaging, and pathology confirmation to differentiate cases and controls could provide accurate analyses moving away from prediction of dementia towards prediction of Alzheimer's disease.

Funding

We thank the Dementia Research Institute [UKDRI supported by the Medical Research Council (UKDRI-3003), Alzheimer's Research UK, and Alzheimer's Society], Welsh Government, Joint Programming for Neurodegeneration (MRC: MR/T04604X/1), Dementia Platforms UK (MRC: MR/L023784/2), MRC Centre for Neuropsychiatric Genetics and Genomics (MR/L010305/1), The Moondance Foundation, VIB and KU Leuven (Methusalem grant), the European Union (grant no. ERC-834682 CELLPHASE_AD), the "Fonds voor Wetenschappelijk Onderzoek", the "Geneeskundige Stichting Koningin Elisabeth", Opening the Future campaign of the Leuven Universitair Fonds, the Belgian Alzheimer Research Foundation, the Dolby Foundation, and the Alzheimer's Association USA. B.D.S. is a holder of the Bax-Vanluffelen Chair for Alzheimer's disease. HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council (#2018-02532), the European Research Council (#681712), Swedish State Support for Clinical Research (#ALFGBG-720931), the Alzheimer Drug Discovery Foundation (ADDF), USA (#201809-2016862), the AD Strategic Fund and the Alzheimer's Association (#ADSF-21-831376-C, #ADSF-21-831381-C and #ADSF-21-831377-C), the Olav Thon Foundation, the Erling-Persson Family Foundation, Stiftelsen för Gamla Tjänarinnor, Hjärnfonden, Sweden (#FO2019-0228), the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860197 (MIRIADE), European Union Joint Program for Neurodegenerative Disorders (JPND2021-00694), and the UK Dementia Research Institute at UCL.

Competing interests

BDS has no direct conflict of interests with the results reported in this manuscript. He has however consulted for several major drug companies and is scientific founder of Augustin TX and Muna TX. He has a small amount of shares in Muna TX. HZ has served at scientific advisory boards and/or as a consultant for Abbvie, Alector, Annexon, Artery Therapeutics, AZTherapies, CogRx, Denali, Eisai, Nervgen, Novo Nordisk, Pinteon Therapeutics, Red Abbey Labs, Passage Bio, Roche, Samumed, Siemens Healthineers, Triplet Therapeutics, and Wave, has given lectures in symposia sponsored by Cellectricon, Fujirebio, Alzecure, Biogen, and Roche, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program (outside submitted work).

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Beach TG, Monsell SE, Phillips LE, Kukull W. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. J. Neuropathol. Exp. Neurol. 2012;
2. Jack CR, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. Lancet Neurol. 2013;
3. Jack CR, Holtzman DM. Biomarker modeling of alzheimer's disease. Neuron 2013;
4. Janelidze S, Stomrud E, Palmqvist S, et al. Plasma β-amyloid in Alzheimer's disease and vascular disease. Sci. Rep. 2016;
5. Nakamura A, Kaneko N, Villemagne VL, et al. High performance plasma amyloid-β biomarkers for Alzheimer's disease. Nature 2018;
6. Palmqvist S, Insel PS, Stomrud E, et al. Cerebrospinal fluid and plasma biomarker trajectories with increasing amyloid deposition in Alzheimer's disease. EMBO Mol. Med. 2019;
7. Schindler SE, Bollinger JG, Ovod V, et al. High-precision plasma β-amyloid 42/40 predicts current and future brain amyloidosis. Neurology 2019;
8. Citron M, Vigo-Pelfrey C, Teplow DB, et al. Excessive production of amyloid β-protein by peripheral cells of symptomatic and presymptomatic patients carrying the Swedish familial Alzheimer disease mutation. Proc. Natl. Acad. Sci. U. S. A. 1994;91(25):11993–11997.
9. Skillbäck T, Farahmand BY, Rosén C, et al. Cerebrospinal fluid tau and amyloid-β1-42 in patients with dementia. Brain 2015;
10. Lantero Rodriguez J, Karikari TK, Suárez-Calvet M, et al. Plasma p-tau181 accurately predicts Alzheimer's disease pathology at least 8 years prior to post-mortem and improves the clinical characterisation of cognitive decline. Acta Neuropathol. 2020;
11. Oeckl P, Halbgebauer S, Anderl-Straub S, et al. Glial fibrillary acidic protein in serum is increased in Alzheimer's disease and correlates with cognitive impairment. J. Alzheimer's Dis. 2019;
12. Elahi FM, Casaletto KB, La Joie R, et al. Plasma biomarkers of astrocytic and neuronal dysfunction in early- and late-onset Alzheimer's disease. Alzheimer's Dement. 2020;
13. Verberk IMW, Laarhuis MB, van den Bosch KA, et al. Serum markers glial fibrillary acidic protein and neurofilament light for prognosis and monitoring in cognitively normal older people: a prospective memory clinic-based cohort study [Internet]. Lancet Heal. Longev. 2021;2(2):e87–e95.Available from: http://dx.doi.org/10.1016/S2666-7568(20)30061-1
14. Mattsson N, Cullen NC, Andreasson U, et al. Association between Longitudinal Plasma Neurofilament Light and Neurodegeneration in Patients with Alzheimer Disease. JAMA Neurol. 2019;
15. Kuhle J, Kropshofer H, Haering DA, et al. Blood neurofilament light chain as a biomarker of MS

disease activity and treatment response. Neurology 2019;

16. Hansson O, Janelidze S, Hall S, et al. Blood-based NfL: A biomarker for differential diagnosis of parkinsonian disorder. Neurology 2017;

17. O'Connor A, Karikari TK, Poole T, et al. Plasma phospho-tau181 in presymptomatic and symptomatic familial Alzheimer's disease: a longitudinal cohort study. Mol. Psychiatry 2020;

18. Escott-Price V, Shoai M, Pither R, et al. Polygenic score prediction captures nearly all common genetic risk for Alzheimer's disease. Neurobiol. Aging 2017;

19. Janelidze S, Mattsson N, Palmqvist S, et al. Plasma P-tau181 in Alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to Alzheimer's dementia. Nat. Med. 2020;

20. Karikari TK, Pascoal TA, Ashton NJ, et al. Blood phosphorylated tau 181 as a biomarker for Alzheimer's disease: a diagnostic performance and prediction modelling study using data from four prospective cohorts. Lancet Neurol. 2020;

21. Palmqvist S, Janelidze S, Quiroz YT, et al. Discriminative Accuracy of Plasma Phospho-tau217 for Alzheimer Disease vs Other Neurodegenerative Disorders. JAMA - J. Am. Med. Assoc. 2020;

22. Palmqvist S, Tideman P, Cullen N, et al. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. Nat. Med. 2021;

23. Cruchaga C, Kauwe JSK, Harari O, et al. GWAS of cerebrospinal fluid tau levels identifies risk variants for alzheimer's disease. Neuron 2013;

24. Li J, Zhang Q, Chen F, et al. Genome-wide association and interaction studies of CSF T-tau/Aβ42 ratio in ADNI cohort. Neurobiol. Aging 2017;

25. Lord J, Zettergren A, Ashton NJ, et al. A genome-wide association study of plasma phosphorylated tau181. Neurobiol. Aging 2021;

26. Li JQ, Yuan XZ, Li HY, et al. Genome-wide association study identifies two loci influencing plasma neurofilament light levels. BMC Med. Genomics 2018;

27. Zettergren A, Lord J, Ashton NJ, et al. Association between polygenic risk score of Alzheimer's disease and plasma phosphorylated tau in individuals from the Alzheimer's Disease Neuroimaging Initiative. Alzheimer's Res. Ther. 2021;

28. Damotte V, Lee SJ, Chouraki V, et al. Plasma amyloid β levels are driven by genetic variants near APOE, BACE1, APP, PSEN2 : A genome- wide association study in over 12,000 non- demented participants . Alzheimer's Dement. 2021;

29. Holmes C, Cairns N, Lantos P, Mann A. Validity of current clinical criteria for Alzheimer's disease, vascular dementia and dementia with Lewy bodies. Br. J. Psychiatry 1999;

30. Jones PR. A note on detecting statistical outliers in psychophysical data. Attention, Perception, Psychophys. 2019;

31. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat. Genet. 2009;41(10):1088–1093.

32. Bellenguez C, Küçükali F, Jansen I, et al. New insights on the genetic etiology of Alzheimer's and related dementia. MedRxiv 2020;1–35.

33. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. Nat. Genet. 2016;

34. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation [Internet]. 2015.Available from: http://biorxiv.org/content/early/2015/12/23/035170.abstract

35. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nat. Genet. 2019;

36. Chang CC, Chow CC, Tellier LCCAM, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets [Internet]. Gigascience 2015;4(1):7.Available from: http://gigascience.biomedcentral.com/articles/10.1186/s13742-015-0047-8

37. Leonenko G, Baker E, Stevenson-Hoare J, et al. Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores. Nat. Commun. 2021;

38. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. 2017;

39. Rentzsch P, Schubach M, Shendure J, Kircher M. CADD-Splice—improving genome-wide variant effect prediction using deep learning-derived splice scores. Genome Med. 2021;

40. Rentzsch P, Witten D, Cooper GM, et al. CADD: Predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 2019;

41. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;

42. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLoS Comput. Biol. 2015;

43. Cullen NC, Leuzy A, Janelidze S, et al. Plasma biomarkers of Alzheimer's disease improve prediction of cognitive decline in cognitively unimpaired elderly populations. Nat. Commun. 2021;

44. Kumar-Singh S, Theuns J, Van Broeck B, et al. Mean age-of-onset of familial Alzheimer disease caused by presenilin mutations correlates with both increased Aβ42 and decreased Aβ40. Hum. Mutat. 2006;

45. Mattsson N, Roseń E, Hansson O, et al. Age and diagnostic performance of Alzheimer disease CSF biomarkers. Neurology 2012;

46. Chatterjee P, Pedrini S, Stoops E, et al. Plasma glial fibrillary acidic protein is elevated in

cognitively normal older adults at risk of Alzheimer's disease. Transl. Psychiatry 2021;

47. Aschenbrenner AJ, Gordon BA, Fagan AM, et al. Neurofilament Light Predicts Decline in Attention but Not Episodic Memory in Preclinical Alzheimer's Disease. J. Alzheimer's Dis. 2020;

48. Mehta PD, Pirttilä T, Mehta SP, et al. Plasma and cerebrospinal fluid levels of amyloid β proteins 1-40 and 1- 42 in Alzheimer disease. Arch. Neurol. 2000;

49. van Oijen M, Hofman A, Soares HD, et al. Plasma Aβ1-40 and Aβ1-42 and the risk of dementia: a prospective case-cohort study. Lancet Neurol. 2006;

50. Mayeux R, Honig LS, Tang M, et al. Plasma Abeta 40 and Abeta 42 and Alzheimer' s disease. Relation to age, mortality, and risk. Neurology 2003;

51. Keshavan A, Pannee J, Karikari TK, et al. Population-based blood screening for preclinical Alzheimer's disease in a British birth cohort at age 70. Brain 2021;

52. Kaeser SA, Häsler LM, Lambert M, et al. CSF p-tau increase in response to Aβ-type and Danish-type cerebral amyloidosis and in the absence of neurofibrillary tangles [Internet]. Acta Neuropathol. 2022;143(2):287–290.Available from: https://doi.org/10.1007/s00401-021-02400-5

53. Schelle J, Häsler LM, Göpfert JC, et al. Prevention of tau increase in cerebrospinal fluid of APP transgenic mice suggests downstream effect of BACE1 inhibition. Alzheimer's Dement. 2017;

54. Sankaranarayanan S, Barten DM, Vana L, et al. Passive immunization with phospho-tau antibodies reduces tau pathology and functional deficits in two distinct mouse tauopathy models. PLoS One 2015;

55. Nagy ZS, Esiri MM, Jobst KA, et al. Influence of the apolipoprotein E genotype on amyloid deposition and neurofibrillary tangle formation in Alzheimer's disease. Neuroscience 1995;

56. Bettayeb K, Hooli B V., Parrado AR, et al. Relevance of the COPI complex for Alzheimer's disease progression in vivo. Proc. Natl. Acad. Sci. U. S. A. 2016;

57. Bednarek AK, Laflin KJ, Daniel RL, et al. WWOX, a novel WW domain-containing protein mapping to human chromosome 16q23.3-24.1, a region frequently affected in breast cancer. Cancer Res. 2000;

58. Chang JY, Chang NS. WWOX dysfunction induces sequential aggregation of TRAPPC6AΔ, TIAF1, tau and amyloid β, and causes apoptosis. Cell Death Discov. 2015;

59. Teng CC, Yang YT, Chen YC, et al. Role of WWOX/WOX1 in Alzheimer's disease pathology and in cell death signaling. Front. Biosci. - Sch. 2013;

60. Kośla K, Kałuzińska, Bednarek AK. The WWOX gene in brain development and pathology. Exp. Biol. Med. 2020;

Figure legends

Figure 1 Pearson correlation between biomarkers in cases (A) and in controls (B).

Figure 2 Aβ42/Aβ40 case-only GWAS (N=1420 cases).

Figure 3 Genome-wide significant regions associated with Aβ42/Aβ40 in case-only analysis (N=1420 cases).

Table 1 Summary of demographics and plasma biomarker summary characteristics (mean [Standard Deviation], μ g/ml) in ADCC, post-outlier removal

	Controls (N=508)	Cases (N=1439)
Demographics		
Age	82.2 [6.72]	68.1 [8.03]
Sex M / F	221 / 287	748 / 691
Age at onset	N/A	62.4 [7.9]
Duration	N/A	5.3 [3.6]
Biomarkers		
Aβ40	140 [40.0]	94.5 [34.4]
Aβ42	7.50 [2.05]	5.00 [1.84]
GFAP	196 [85.3]	215 [103]
NfL	32.9 [13.7]	31.0 [13.9]
P-tau181	3.18 [1.54]	4.10 [1.90]
Aβ42/Aβ40	0.0556 [0.013]	0.0543 [0.014]

Values in brackets are standard deviation.

Table 2 Beta coefficients, standard errors, and *p*-values for linear regressions predicting biomarkers from age at onset and disease duration in Alzheimer's disease cases, controlling for age and sex

		Age at onset			Duration		
	N	**B**	**SE**	***p***	**B**	**SE**	***P***
Aβ40	1219	0.042	0.003	1.9×10^{-35}	0.016	0.007	0.027
Aβ42	1219	0.034	0.003	4.2×10^{-23}	0.013	0.007	0.077
GFAP	1301	0.034	0.003	7.1×10^{-24}	0.034	0.007	2.9×10^{-6}
NfL	1275	0.048	0.003	1.1×10^{-44}	0.050	0.007	1.2×10^{-12}
pTau-181	1309	0.011	0.003	0.0023	0.020	0.008	0.008
Aβ42/Aβ40	1215	-0.012	0.003	0.0005	-0.004	0.008	0.592

Table 3 Beta coefficients, standard errors, and *p*-values for linear regressions predicting biomarkers from age at interview in cases and controls, controlling for sex

	Cases (max N=1439)				Controls (max N=508)			
	N	**B**	**SE**	***p***	**N**	**B**	**SE**	***P***
Aβ40	1415	0.041	0.003	2.9×10^{-37}	492	0.064	0.006	1.2×10^{-22}
Aβ42	1417	0.034	0.003	4.6×10^{-25}	486	0.036	0.007	1.2×10^{-7}
GFAP	1394	0.034	0.003	8.4×10^{-28}	501	0.052	0.006	4.6×10^{-16}
NfL	1361	0.051	0.003	1.2×10^{-54}	478	0.074	0.006	1.9×10^{-30}
P-tau181	1389	0.014	0.003	4.3×10^{-5}	472	0.038	0.007	3.5×10^{-08}
Aβ42/Aβ40	1413	−0.010	0.003	0.0018	481	-0.044	0.007	1.2×10^{-10}

Table 4 Results of logistic regressions predicting Alzheimer's disease status from each biomarker, adjusted for age and sex (1302 cases and 421 controls after excluding the missing values list-wise)

	B	SE	*p*	R^2	AUC
Aβ40	−0.44	0.058	3.5×10^{-14}	0.05	0.63
Aβ42	−0.56	0.059	2.8×10^{-21}	0.08	0.66
GFAP	0.55	0.067	2.4×10^{-16}	0.07	0.64
NfL	0.47	0.066	1.1×10^{-12}	0.05	0.63
P-tau181	0.55	0.067	1.4×10^{-16}	0.07	0.65
Aβ42/Aβ40	−0.18	0.055	0.0009	0.01	0.56

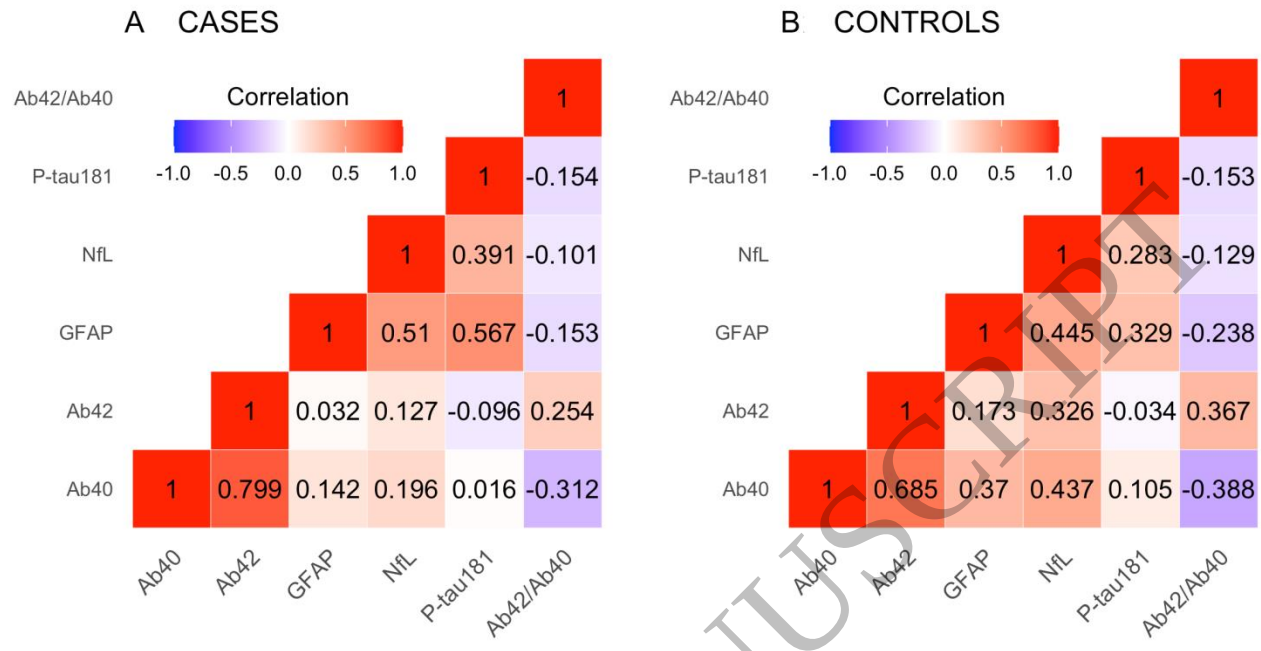

Figure 1
559x279 mm (0.0 x DPI)

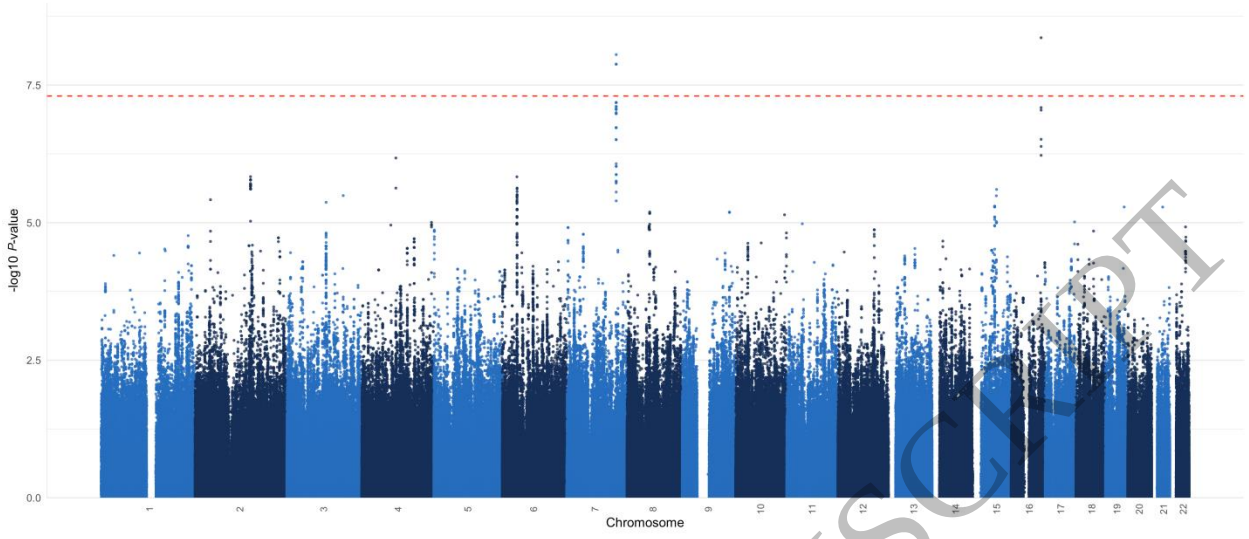

Figure 2
559x242 mm (0.0 x DPI)

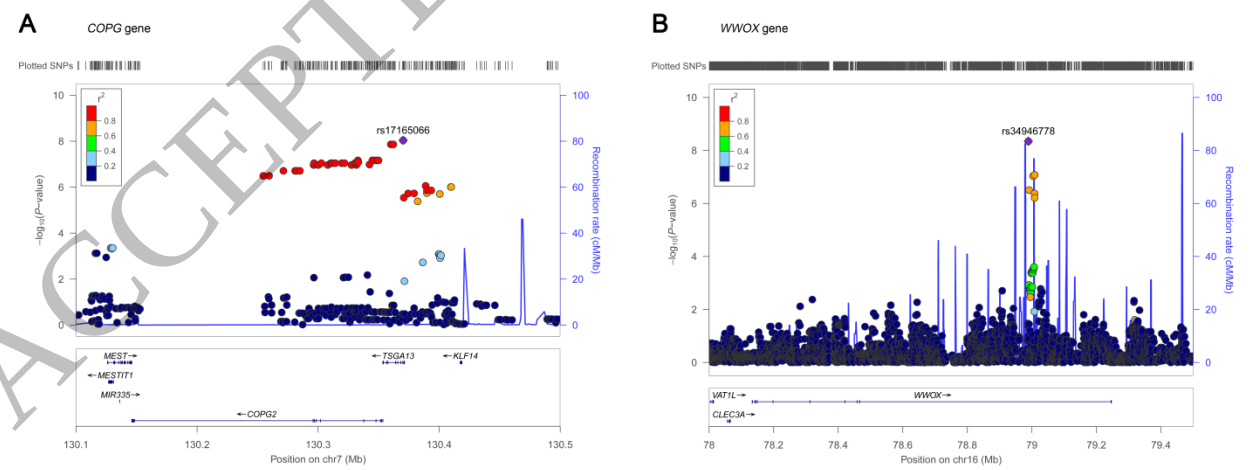

Figure 3
524x203 mm (0.0 x DPI)