



Featured Article

Discovery and validation of plasma proteomic biomarkers relating to brain amyloid burden by SOMAscan assay

Liu Shi^{a,*}, Sarah Westwood^a, Alison L. Baird^a, Laura Winchester^a, Valerija Dobricic^b, Fabian Kilpert^b, Shengjun Hong^b, Andre Franke^c, Abdul Hye^d, Nicholas J. Ashton^{d,e,f,g}, Angharad R. Morgan^h, Isabelle Bosⁱ, Stephanie J. B. Vos^j, Noel J. Buckley^a, Mara ten Kate^j, Philip Scheltens^j, Rik Vandenberghe^k, Silvy Gabel^{k,l}, Karen Meersmans^{k,l}, Sebastiaan Engelborghs^{m,n,o,p}, Ellen E. De Roeck^{m,n,o}, Kristel Slegers^{o,q}, Giovanni B. Frisoni^{r,s}, Olivier Blin^t, Jill C. Richardson^u, Régis Bordet^v, José L. Molinuevo^{w,x}, Lorena Rami^x, Anders Wallin^y, Petronella Kettunen^{y,z}, Magda Tsolaki^{aa}, Frans Verheyⁱ, Alberto Lleó^{bb}, Daniel Alcolea^{cc}, Julius Popp^{dd,ee}, Gwendoline Peyratout^{dd}, Pablo Martinez-Lage^{ff}, Mikel Tainta^{ff}, Peter Johannsen^{gg}, Charlotte E. Teunissen^{hh}, Yvonne Freund-Levi^{ii,jj,kk}, Lutz Frölich^{ll}, Cristina Legido-Quigley^{mm,nn}, Frederik Barkhof^{oo,pp}, Kaj Blennow^{e,qq}, Henrik Zetterberg^{e,qq,rr,ss}, Susan Baker^{tt}, B. Paul Morgan^h, Johannes Streffer^{n,uu}, Pieter Jelle Visser^{i,j,vv}, Lars Bertram^{b,ww}, Simon Lovestone^{a,xx}, Alejo J. Nevado-Holgado^a

^aDepartment of Psychiatry, University of Oxford, Oxford, UK

^bLübeck Interdisciplinary Platform for Genome Analytics, Institutes of Neurogenetics and Cardiogenetics, University of Lübeck, Lübeck, Germany

^cInstitute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany

^dMaurice Wohl Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, UK

^eDepartment of Psychiatry and Neurochemistry, University of Gothenburg, Mölndal, Sweden

^fWallenberg Centre for Molecular & Translational Medicine, University of Gothenburg, Gothenburg, Sweden

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

^hDementia Research Institute Cardiff, Cardiff University, Cardiff, UK

ⁱDepartment of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Alzheimer Centrum Limburg, Maastricht University, Maastricht, the Netherlands

^jAlzheimer Center, VU University Medical Center, Amsterdam, the Netherlands

^kUniversity Hospital Leuven, Leuven, Belgium

^lLaboratory for Cognitive Neurology, Department of Neurosciences, KU Leuven, Belgium

^mCenter for Neurosciences, Vrije Universiteit Brussel (VUB), Brussels, Belgium

ⁿReference Center for Biological Markers of Dementia (BIODEM), University of Antwerp, Antwerp, Belgium

^oInstitute Born-Bunge, University of Antwerp, Antwerp, Belgium

^pDepartment of Neurology, VUB University Hospital Brussels (UZ Brussel), Brussels, Belgium

S.L. has been named as an inventor on biomarker intellectual property protected by Proteome Sciences and Kings College London; in both cases not the subject of the present study. He has served on the advisory board of SomaLogic Inc. and has provided consultancy for Merck and Eisai within the past 3 years. He is currently an employee of Janssen-Cilag, UK. H.Z. has served at scientific advisory boards of Eli Lilly, Roche Diagnostics, Wave, CogRx, and Samumed, has received travel support from Teva, and is a cofounder of Brain Biomarker Solutions in Gothenburg AB (all unrelated to this study). J.L.M. has served at scientific advisory boards of Eli Lilly, Roche Diagnostics, Roche, Biogen, Lundbeck, Novartis, IBL, Axovant, Orzozon, and Merck (all unrelated to this study). A.L.B. has served at scientific advisory boards of Fujirebio Europe, Eli Lilly, Novartis, Nutricia, and Ot-

suka and is the inventor of a patent on synaptic markers in CSF (all unrelated to this study). L.F. has received research funding, consultancy fees, and speech honoraria from Allergan, Avid-Eli Lilly, Avanir, Avraham Pharmaceuticals, Axon Neuroscience, Axovant, Biogen, Boehringer Ingelheim, Eisai, Functional Neuromodulation, GE Health Care, Lundbeck, Merck Sharpe & Dohme, Novartis, Pfizer, Piramal Imaging, Roche, and Schwabe Pharma. J.P. has served at scientific advisory boards of Fujirebio Europe, Eli Lilly, and Nestlé Institute of Health Sciences, all unrelated to this study.

*Corresponding author. Tel.: +44 01865 618232; Fax: +44 01865 793101.

E-mail address: liu.shi@psych.ox.ac.uk

<https://doi.org/10.1016/j.jalz.2019.06.4951>

1552-5260/© 2019 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

^qNeurodegenerative Brain Diseases Group, Center for Molecular Neurology, VIB, Edegem, Belgium

^rUniversity of Geneva, Geneva, Switzerland

^sIRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy

^tAIX Marseille University, INS, Ap-hm, Marseille, France

^uNeurosciences Therapeutic Area, GlaxoSmithKline R&D, Stevenage, UK

^vUniversity of Lille, Inserm, CHU Lille, Lille, France

^wAlzheimer's disease & other cognitive disorders unit, Hospital Clinic-IDIBAPS, Barcelona, Spain

^xBarcelona Beta Brain Research Center, Universitat Pompeu Fabra, Barcelona, Spain

^yInstitute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden

^zNuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

^{aa}1st Department of Neurology, AHEPA University Hospital, Makedonia, Thessaloniki, Greece

^{bb}Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

^{cc}Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

^{dd}University Hospital of Lausanne, Lausanne, Switzerland

^{ee}Geriatric Psychiatry, Department of Mental Health and Psychiatry, Geneva University Hospitals, Geneva, Switzerland

^{ff}CITA-Alzheimer Foundation, San Sebastian, Spain

^{gg}Danish Dementia Research Centre, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

^{hh}Neurochemistry Laboratory, Department of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam University Medical Centers, Vrije Universiteit, Amsterdam, The Netherlands

ⁱⁱDepartment of Neurobiology, Caring Sciences and Society (NVS), Division of Clinical Geriatrics, Karolinska Institutet, and Department of Geriatric Medicine, Karolinska University Hospital Huddinge, Stockholm, Sweden

^{jj}Department of Psychiatry in Region Örebro County and School of Medical Sciences, Faculty of Medicine and Health, Örebro University, Örebro, Sweden

^{kk}Department of Old Age Psychiatry, Psychology and Neuroscience, Kings College London, London, UK

^{ll}Department of Geriatric Psychiatry, Zentralinstitut für Seelische Gesundheit, University of Heidelberg, Mannheim, Germany

^{mm}Kings College London, London, UK

ⁿⁿThe Systems Medicine Group, Steno Diabetes Center, Gentofte, Denmark

^{oo}Department of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherlands

^{pp}UCL Institutes of Neurology and Healthcare Engineering, London, UK

^{qq}Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden

^{rr}UK Dementia Research Institute at UCL, London, UK

^{ss}Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK

^{tt}Janssen R&D, Titusville, NJ, USA

^{uu}Janssen R&D, LLC, Beerse, Belgium

^{vv}Department of Neurobiology, Care Sciences and Society, Division of Neurogeriatrics, Karolinska Institutet, Stockholm, Sweden

^{ww}Department of Psychology, University of Oslo, Oslo, Norway

^{xx}Janssen-Cilag UK, formerly Department of Psychiatry, University of Oxford, Oxford, UK

Abstract

Introduction: Plasma proteins have been widely studied as candidate biomarkers to predict brain amyloid deposition to increase recruitment efficiency in secondary prevention clinical trials for Alzheimer's disease. Most such biomarker studies are targeted to specific proteins or are biased toward high abundant proteins.

Methods: 4001 plasma proteins were measured in two groups of participants (discovery group = 516, replication group = 365) selected from the European Medical Information Framework for Alzheimer's disease Multimodal Biomarker Discovery study, all of whom had measures of amyloid.

Results: A panel of proteins (n = 44), along with age and apolipoprotein E (APOE) ε4, predicted brain amyloid deposition with good performance in both the discovery group (area under the curve = 0.78) and the replication group (area under the curve = 0.68). Furthermore, a causal relationship between amyloid and tau was confirmed by Mendelian randomization.

Discussion: The results suggest that high-dimensional plasma protein testing could be a useful and reproducible approach for measuring brain amyloid deposition.

© 2019 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords:

Amyloid β; SOMAscan assay; Plasma proteomics; Replication; Causal relationship; Tau

1. Introduction

Anti-amyloid clinical trials for Alzheimer's disease (AD) have been unsuccessful to date [1]. The failure of such trials was partially caused by the fact that participants enrolled in

such trials present clinically with AD but lack amyloidosis [2]. Furthermore, most such trials have been conducted relatively late in the disease process and targeting treatment to earlier presymptomatic or prodromal stages of the disease might have more success [3]. As a consequence, it has

increasingly common to conduct clinical trials in preclinical or prodromal disease where AD pathology is confirmed using biomarkers [2,4]. Currently, the best characterized methods for measuring amyloid pathology are positron emission tomography (PET) imaging measures of brain amyloid deposition and cerebrospinal fluid (CSF) measurement of amyloid β ($A\beta$) levels [5]. However, these measures are challenging because of invasiveness, cost, and restricted availability [6,7].

Blood-based biomarkers have therefore been investigated as a less-invasive and potentially cost-effective option for early detection and monitoring of AD pathology. An increasing number of studies [8–12], including those by ourselves [13–19], have found that a range of proteins in plasma can reflect neocortical $A\beta$ burden and hence might act as biomarkers. Recent studies have added to these by demonstrating that measures of $A\beta$ in blood can predict brain amyloid with high accuracy [20–22], further indicating that blood proteins can be effective biomarkers for predicting brain pathology.

The failure of anti-amyloid clinical trials also provoked some to question the amyloid cascade hypothesis [23–25]. Broadly speaking, the amyloid cascade hypothesis states that over time, an imbalance in $A\beta$ production and/or clearance leads to gradual accumulation and aggregation of the peptide in the brain, initiating a neurodegenerative cascade that involves not only amyloid deposition but also inflammation, tau pathology, neuronal dysfunction, and loss [26,27]. Despite some reservations regarding this hypothesis prompted by the failure of clinical trials, amyloid still remains a central part of our understanding of the pathophysiology of the disease. Nonetheless, assessing the relationship between amyloid and other AD pathologies has become increasingly important.

With this in mind, we embarked on the study reported here, with two main objectives; first to identify a plasma molecular signature of amyloid pathology for potential use as a biomarker for preclinical or prodromal detection of disease and second to investigate the causal relationship between amyloid and other AD pathologies. We report here a replicating signature of disease reflecting brain amyloid load and other features of AD, including tau pathology, in blood. Then using Mendelian randomization (MR) [28], an approach more often used in genetic studies, we find a causal relationship between amyloid and tau while such reverse relationship was not found.

2. Methods

2.1. Participants: EMIF-AD Multimodal Biomarker Discovery study

The EMIF-AD Multimodal Biomarker Discovery (MBD) study is part of the European Medical Information Framework for Alzheimer's disease (EMIF-AD; <http://www.emif.eu>); a public-private precompetitive cross European

program seeking to make data, and in these case samples, reusable for secondary studies of neurodegeneration. The design of the EMIF-AD MBD study has been described previously [29]. Briefly, 1221 participants meeting specified and previously reported inclusion criteria were identified from 11 preexisting European cohorts ("parent cohorts") using the EMIF catalog (<https://emif-catalogue.eu>) and associated tools. Each parent cohort was approved by local medical ethics committees. As reported previously [29], a critical criterion for inclusion was that a measure of brain amyloid load was available, using either CSF $A\beta$ or amyloid PET scan. The classification of amyloid burden (high/low) in the EMIF-AD MBD cohort has been described previously [29]. Furthermore, $A\beta$ measurements were combined into a continuous variable using z-scoring, using the mean and standard deviation of the control subjects as a reference [29].

In addition to the primary outcome variable of amyloid load, another five markers in CSF were measured including total tau (t-tau), phosphorylated tau (p-tau), neurofilament light chain (NFL), neurogranin, and YKL-40 [29]. The levels of t-tau and p-tau in CSF were measured locally, and the local cutoff point was used to determine their status (high/low). Consequently, the p-tau and t-tau values were Z-scored with controls within each data set as a reference. The other three CSF markers (NFL, neurogranin, and YKL-40) were measured in a central laboratory (Gothenburg University, Sweden), and their status (high/low) was determined by the median value of each marker within the whole data set.

In addition, the following AD-related phenotypes were also measured: magnetic resonance imaging values of hippocampal volume; clinical assessments including baseline diagnosis, baseline Mini-Mental State Examination (MMSE) score, and mild cognitive impairment (MCI) conversion; and finally, apolipoprotein E (*APOE*) and genome-wide single nucleotide polymorphism (SNP) genotyping [29]. SNP assays were conducted using the "Global Screening Array" (Illumina, Inc.) using standard procedures from 250 ng of DNA extracted from whole blood [30]. Raw data processing and initial quality control (QC) were performed in GenomeStudio (v2.0.2) using GenTrain (v3.0) as clustering algorithm generating genotypes for 645,896 SNPs. Postprocessing was performed in PLINK (v1.9) using standard QC thresholds (i.e., sample genotyping efficiency > 95%, SNP genotyping efficiency > 98%, departures from Hardy Weinberg equilibrium at P value < .000005, and a minor allele frequency [MAF] of < 0.01). Eventually, this left 490,717 SNPs for analysis. Genotypes for reconstruction of the *APOE* ϵ 2/3/4 haplotypes were either generated on the Global Screening Array (rs7412) or by manual genotyping (rs429358) as described previously [31]. We classified individuals as *APOE* ϵ 4 carriers or noncarriers according to their genotype status at rs429358 (C-allele = ϵ 4).

The 1221 subjects selected for the EMIF-AD MBD cohort included 492 healthy controls (HCs), 527 MCI, and 202 AD participants. Among these participants, we included in the study all those that had plasma samples available (total

881). These samples were then divided into two separated groups, with each group of samples being processed independently. Group 1 was used as a discovery group and included 150 HCs (29.1%), 188 MCI (36.4%), and 178 AD (34.5%). Group 2 was used for validation and included 161 HCs (44.1%), 198 MCI (54.2%), and 6 AD (1.6%). Furthermore, we sought to maintain amyloid balance in controls and MCI individuals given that the objective is to identify biomarkers of amyloid pathology in preclinical stage. As a result, in this combined group of samples from people in the HC and MCI categories (total $n = 697$), there were 338 with and 359 without amyloid pathology.

2.2. SOMAscan assay

We measured proteins in plasma using the SOMAscan assay platform (SomaLogic Inc.). SOMAscan is an aptamer-based assay allowing for the simultaneous measurement and quantification of, in the version used here, 4001 proteins. The assay uses chemically modified nucleotides to transform a protein signal into a nucleotide signal that can be quantified using relative fluorescence on microarrays [32]. As noted previously, groups of samples were measured independently. Forty subjects were tested in both batches to normalize the data across assay runs.

2.3. Statistical analysis

All statistical analyses were completed using R (version 3.3.2). The Mann-Whitney U test was used to compare participants' characteristics between high- and low-amyloid categories in both groups. The association between proteins and amyloid Z-score was calculated by partial Spearman correlation, adjusting for five covariates including age, sex, *APOE* $\epsilon 4$ status, study of origin, and number of blood freeze-thaw cycles. False discovery rate correction was used to correct *P* values for multiple comparisons. Proteins that were differentially expressed at a significance level of $P < .05$ were further nominated for pathway analysis using the Metascape software (<http://metascape.org/gp/index.html#/main/step1>). Briefly, differentially expressed proteins were inputted as "protein list" and all 4001 proteins measured by SOMAscan assay were used as "background." This enrichment analysis was performed on the KEGG database.

2.4. Machine learning

Machine learning was used to identify optimal multivariate signatures, including both proteins and demographic data (age and *APOE* $\epsilon 4$) as input features, to differentiate between high and low $A\beta$ status in group 1. First, the effects of study and blood freeze-thaw cycle on proteins were removed by linear regression and the residuals were used for analysis. Subsequently, Lasso was used to select the "n" top input features that best differentiated high from low $A\beta$. Support vector machine classifiers were then built using from 1 to 100 features to predict the outcome under 10-fold cross-validation. Those

features that produced best performance were selected from group 1. To verify replication, the optimal classifier was then validated in the second and combined cohorts for distinguishing $A\beta$ status and other AD endophenotypes.

2.5. Mendelian randomization

MR was used to investigate the causal relationship between amyloid and tau pathology. MR uses genetic variants as an instrument when they are robustly associated with a risk factor (exposure), with the fundamental assumption that they are associated with the studied outcome through that exposure. As genetic variants cannot be altered and are static throughout an individual's lifetime, reverse causation can be excluded, rendering MR a powerful tool to examine causality between the exposure and outcome [28].

To explore the causal relationship between amyloid and tau (both t-tau and p-tau), we first performed a systematic review of the literature to search SNPs that were significantly associated with amyloid but not with tau or p-tau. Overall, 78 SNPs were found to be significantly associated with abnormal amyloid pathology measured by PET scan, CSF amyloid, or post-mortem measurement (Supplementary Table 1). Six SNPs were excluded including 3 which were also associated with t-tau and p-tau and 3 which were within 1 Mb of *APOE* (i.e., *APOC1* and *TOMM40*), given that *APOE* is strongly associated with both amyloid and tau. Of the remaining SNPs, 6 were not measured using the assay we used, leaving 66 SNPs for further analysis (Supplementary Table 1). The association of these 66 SNPs with the Z-score of amyloid, t-tau, and p-tau was calculated using PLINK (v1.7). Finally, based on the association, we used four MR approaches, including MR Egger linear regression, weighted median, inverse-variance weighted (IVW), and simple median [33], to investigate the causal relationship between amyloid and tau (both t-tau and p-tau). Each MR method is based on different assumptions. The Egger method is sensitive to SNP pleiotropy and allows the estimation of underlying bias by allowing a non-zero estimate for the intercept of the calculated ratio of β values [34]. The weighted median method uses a median of the individual causal estimate per SNP, which is calculated from the ratio estimates of outcome's regression coefficient divided by exposure [35]. The IVW method uses the same ratio estimates but incorporates IVWs into the final summary estimate [36]. Comparing estimates from all the methods shows the robustness of the overall analysis. Furthermore, to determine if there was any single SNP driving the relationship, we performed a leave-one-out analysis where the MR is performed removing a different SNP in each iteration [37].

A similar approach was performed to investigate the causal relationship between t-tau and amyloid. Using the same approach, we identified 21 SNPs that were significantly associated with the abnormal CSF t-tau level but not with amyloid (Supplementary Table 2). After checking the effect allele, eighteen SNPs passed QC and were used for MR analysis (Supplementary Table 2). The causal

Table 1
Demographics of the study population

Characteristics	Group 1 (n = 516)			Group 2 (n = 365)		
	High A β (n = 369)	Low A β (n = 147)	P value	High A β (n = 134)	Low A β (n = 231)	P value
A β Z-score	-1.39 \pm 0.50	0.49 \pm 0.57	<.001*	-1.22 \pm 0.40	0.52 \pm 0.67	<.001*
Age (yrs)	70.0 \pm 8.2	68.3 \pm 8.4	.04*	70.0 \pm 8.3	64.9 \pm 8.1	<.001*
Female sex, N (%)	210 (57)	62 (42)	.002*	70 (52)	139 (60)	.14
Male sex, N (%)	159 (43)	85 (58)	.002*	64 (48)	92 (40)	.14
APOE ϵ 4+, N (%)	244 (66)	43 (29)	<.001*	76 (57)	53 (23)	<.001*
APOE ϵ 4-, N (%)	125 (34)	104 (71)	<.001*	58 (43)	178 (77)	<.001*
MMSE	24.1 \pm 4.6	27.0 \pm 3.2	<.001*	26.9 \pm 2.7	28.2 \pm 1.8	<.001*

NOTE. Percentage of cases is shown in brackets for female and male sex, as well as APOE ϵ 4 carriers and noncarriers in each amyloid β (A β) category. P values compare each demographic across the high A β and low A β groups. Group 1 containing 516 participants was used for discovery, and group 2 of 365 participants for validation.

Abbreviations: APOE, apolipoprotein E; MMSE, Mini-Mental State Examination.

* $P < .05$.

relationship between p-tau and amyloid was not presented because few SNPs were found to be strongly associated with p-tau. Therefore, this approach was underpowered for making causal inferences.

3. Results

3.1. Subject demographics

The EMIF-MBD cohort comprises 1221 subjects of which 881 had samples available for this study. As described in the Methods section, these were divided into two groups for discovery and validation purposes. Table 1 shows the characteristics of the participants from two groups split by amyloid status (high/low). For both groups, the participants with high amyloid were older ($P = .04$ in group 1, $P < .001$ in group 2), the MMSE was lower ($P < .001$ in group 1, $P < .001$ in group 2), and the prevalence of APOE ϵ 4 carriers was higher ($P < .001$ in group 1, $P < .001$ in group 2). A significant difference in sex was found in group 1 ($P = .002$), whereas there was no difference in group 2 ($P = .14$).

3.2. Proteins significantly associated with amyloid Z-score in both groups

In this study, we sought to find plasma biomarkers related to amyloid pathology independent of diagnosis and so combined samples from study participants across the AD diagnostic spectrum (HC, MCI, and AD) and used measures of amyloid pathology as the principle outcome measure. Combining samples in this way reflects the operational need for a biomarker predicting brain amyloid pathology. Recent experience suggests that up to 20% of people with a diagnosis of AD in clinical trials had little or no brain amyloid pathology, demonstrating that clinical classification is an inadequate predictor of brain pathology. Moreover, a biomarker of brain pathology in all three states—HC, MCI, and AD—would have value. In HC, a marker of brain amyloid would help to identify preclinical AD, in MCI, to

identify prodromal AD, and in AD itself, to add to the accuracy of diagnosis.

Overall, 3676 and 3916 proteins passed QC measures in group 1 and group 2, respectively. Using partial correlation to analyze the association between proteins and amyloid Z-score, we found that in group 1, 301 proteins reached statistical significance ($P < .05$) and 5 of them reached false discovery rate ($q < 0.1$) (Supplementary Table 3), corrected for multiple comparisons. In group 2, 536 proteins reached statistical significance ($P < .05$) while none passed false discovery rate. Fig. 1A and B shows volcano plots of proteins from both groups with cutoff $P = .05$.

We then merged the lists of differentially expressed proteins obtained from both groups finding a total of 786 proteins associated with high versus low amyloid. This number of dysregulated proteins was significantly higher than that obtained by chance alone ($P < .001$). We then used the Metascape software, performed on the KEGG database, to assess the biological significance of these dysregulated proteins. Pathway analysis of the 786 differentially expressed proteins revealed 10 significantly enriched pathways: arginine biosynthesis ($P = .003$), inflammatory mediator regulation of TRP channels ($P = .007$), glycerophospholipid metabolism, ($P = .010$), citrate cycle ($P = .012$), Notch signaling ($P = .016$), RNA degradation ($P = .023$), type II diabetes mellitus ($P = .038$), peroxisome ($P = .038$), fatty acid degradation ($P = .041$), and NOD-like receptor signaling ($P = .043$) pathways.

3.3. Multiprotein classifier of amyloid normal/abnormal status and other AD endophenotypes

Having demonstrated that plasma proteins are significantly altered in samples from people with abnormal A β levels, we then sought to find a minimal signal from these plasma data that might act as a biomarker indicative of brain amyloid pathology. To do this we performed machine learning with 10-fold cross-validation in group 1 to identify the optimal multivariate signatures that differentiated between high- and low-A β individuals. We first used

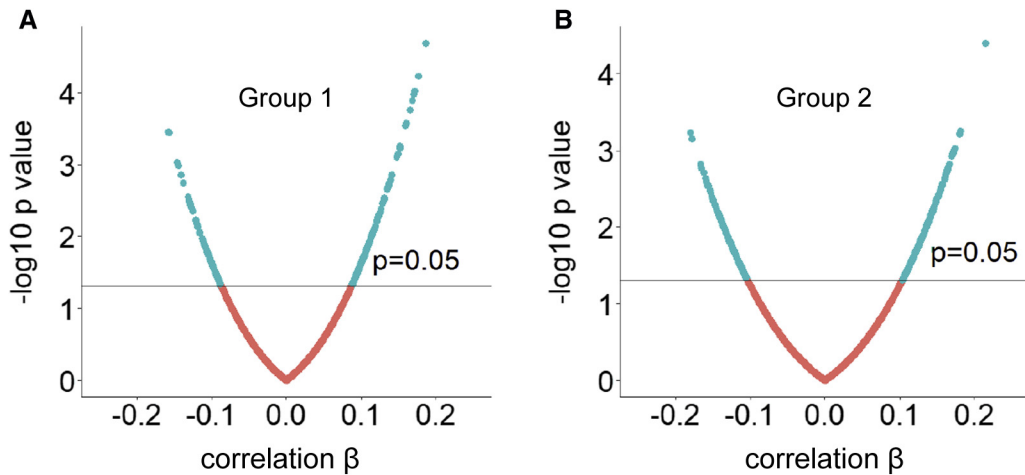


Fig. 1. Correlation of plasma proteins with amyloid Z-score by partial Spearman correlation. (A and B) Volcano plot of proteins obtained from group 1 and group 2, respectively.

demographic variables (age, sex, and education) with *APOE* $\epsilon 4$ as input features to predict amyloid burden. Results showed that a combination of age and *APOE* $\epsilon 4$ achieved the highest predictive value AUC of 0.70. Then, we added all measured proteins to these demographic variables and we found that a panel of 46 features achieved the highest predictive value AUC of 0.78. The input features automatically selected by the classifier were age, *APOE* $\epsilon 4$, and 44 proteins (Supplementary Table 4). The AUC in training and testing sets for different number of input features is shown in Supplementary Fig. 1. We then investigated the performance of this classifier on differentiating amyloid status in the second and independent group, maintaining the parameters obtained from group 1. The results showed that the same panel achieved an AUC of 0.68, demonstrating good replication (Fig. 2A).

We further investigated the performance of the same panel to predict other AD endophenotypes in the combined cohort (combining groups one and two), including discrim-

inating AD from HC and AD from MCI, differentiating status of t-tau, p-tau, MCI conversion, baseline MMSE score, hippocampal volume, and other three CSF biomarkers including NFL, neurogranin, and YKL-40. As shown in Fig. 2B, the panel predicted most of these AD phenotypes with an AUC of >0.6 , except for CSF NFL, neurogranin, and YKL-40. The sample size and AUC of each AD phenotype is shown in Supplementary Table 5.

3.4. Causal relationship between amyloid and tau pathology

We then sought to use these data to address the critically important question of causality in relation to AD pathological features. To investigate the causal relationship between amyloid and tau, we first identified SNPs associated with amyloid pathology from systematic review of the literature. We found 72 SNPs (Supplementary Table 1) significantly associated with abnormal amyloid but not with tau

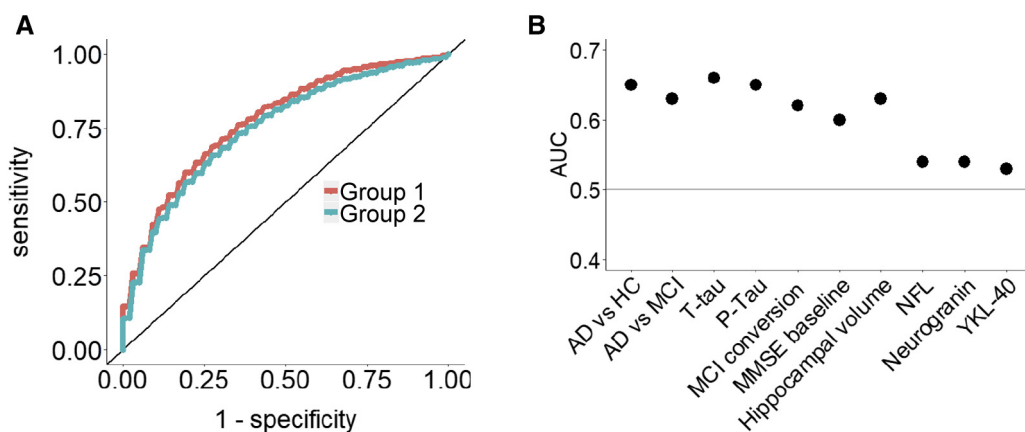


Fig. 2. (A) AUC of the 46 signatures to differentiate amyloid status in group 1 and group 2. (B) AUC of the 46 signatures to differentiate other AD endophenotypes in the combined cohort. Abbreviations: AUC, area under the curve; HC, healthy control; MCI, mild cognitive impairment; t-tau, total tau; p-tau, phosphorylated tau; MMSE, Mini-Mental State Examination; NFL, neurofilament light chain.

Table 2

MR estimates of the causal effect of amyloid on t-tau and p-tau as well as the causal effect of t-tau on amyloid

Study type	Methods	β	se	95% CI	<i>P</i> value
Amyloid SNPs with t-tau	MR Egger	0.85	0.15	0.54 to 1.15	<.001
	Weighted median	0.73	0.22	0.31 to 1.16	<.001
	IVW	0.86	0.11	0.65 to 1.08	<.001
Amyloid SNPs with p-tau	Simple median	0.81	0.50	-0.17 to 1.80	.110
	MR Egger	0.73	0.17	0.39 to 1.06	<.001
	Weighted median	0.62	0.23	0.17 to 1.06	.006
t-tau SNPs with amyloid	IVW	0.86	0.12	0.62 to 1.10	<.001
	Simple median	1.31	0.52	0.29 to 2.33	.014
	MR Egger	-0.45	0.36	-1.15 to 0.26	.23
	Weighted median	-0.31	0.29	-0.87 to 0.25	.27
	IVW	-0.02	0.20	-0.42 to 0.38	.93
	Simple median	-0.20	2.05	-4.22 to 3.81	.92

Abbreviations: β , β coefficient; CI, confidence interval; IVW, inverse-variance weighted; MR, Mendelian randomization; t-tau, total tau; p-tau, phosphorylated tau; SNP, single nucleotide polymorphism.

pathology of which 66 were assayed using the array used in this study (Supplementary Table 1) and hence available for MR analysis. Results showed that amyloid was associated with a moderate alteration in t-tau when using three MR approaches including the Egger method ($\beta = 0.85$, $se = 0.15$, 95% CI [0.54, 1.15], $P < .001$), the IVW method ($\beta = 0.86$, standard error of the effect size [se] = 0.11, 95% CI [0.65, 1.08], $P < .001$), and weighted median method ($\beta = 0.73$, $se = 0.22$, 95% CI [0.31, 1.16], $P < .001$), though not from simple median method ($\beta = 0.81$, $se = 0.50$, 95% CI [-0.17, 1.80], $P = .11$; Table 2, Supplementary Fig. 2). Altered amyloid was also associated with p-tau in all four MR approaches (Table 2, Supplementary Fig. 2). Results from leave-one-out analysis demonstrated that no single SNP was driving the majority of the association signal between amyloid and either t-tau or p-tau (Supplementary Fig. 3). In addition, we explored the causal relationship between amyloid and other three CSF markers including NFL, neurogranin, and YKL-40; no significant association between amyloid and these three markers was found (Supplementary Table 6).

We then reversed the direction of analysis and investigated whether there was a causal relationship between t-tau and amyloid. Twenty-one SNPs (Supplementary Table 2) were found to be significantly associated with the abnormal CSF t-tau level from systematic review, among which eighteen SNPs (Supplementary Table 2) passed QC and were used for MR analysis. Results showed no causal relationship between t-tau and amyloid from any of the four MR approaches (Table 2). We also investigated the causal relationship between t-tau and other three CSF markers (NFL, neurogranin, and YKL-40). No causal relationship between t-tau and NFL was found, whereas such

relationship was found between t-tau and neurogranin as well as between t-tau and YKL-40 (Supplementary Table 6).

4. Discussion

In this study, we used the SOMAscan assay to measure 4001 proteins in plasma in two groups of subjects from the EMIF-AD MBD study. Using machine learning, we identified a panel of 44 proteins that, along with age and *APOE* $\epsilon 4$, predicted measures of central amyloid with good performance in both the discovery group (AUC = 0.78) and the replication group (AUC = 0.68). The same panel also predicted other AD phenotypes, including CSF t-tau, CSF p-tau, MCI conversion, baseline MMSE score, and hippocampal volume. Moreover, a causal relationship between amyloid and t-tau and with p-tau was confirmed using a MR approach, whereas evidence for the reverse relationship between t-tau and amyloid was not found.

These results are in concordance with one of our previous studies [18], in which we used the same SOMAscan assay to measure plasma proteins in 58 cognitively healthy men. In this previous study, 667 proteins were significantly associated with CSF amyloid measurement, and of these 167 proteins (25%) overlapped with the differentially expressed proteins obtained in the present study. Furthermore, the number of overlapping proteins was significantly higher than expected from chance alone ($P = .01$), confirming that the SOMAscan assay may be a useful and powerful screening tool for identifying plasma proteins reflecting amyloid deposition in early, even preclinical, phases of disease and across the clinical AD spectrum as in the present study.

Pathway analysis of the 786 differentially expressed proteins revealed 10 significantly enriched pathways. When comparing these enriched pathways with AlzPathway, a comprehensive map of AD signaling and related pathways [38,39], we found two pathways common to both; Notch signaling and type II diabetes mellitus. This finding is interesting given that multiple lines of evidence have previously shown that Notch signaling and type II diabetes mellitus are associated with AD [40–43]. Our results add weight to this finding and indicate that proteins involved in these two pathways could be treatment targets of AD.

From the discovery phase of the present study, a panel of 46 features achieved a high AUC in discriminating high from low brain amyloid (Fig. 2A). Age and *APOE* $\epsilon 4$ contributed to the model, as in previous studies seeking biomarkers indicative of brain amyloid; unsurprisingly given that both age and *APOE* genotype are highly associated with amyloid pathology [44–48]. Despite this, the protein signature consisting of 44 proteins achieved a greater discrimination performance in the discovery group than age and *APOE* $\epsilon 4$ (0.78 vs. 0.70). In this study, as the sample collection was deliberately selected to overrepresent high amyloid individuals, the proportion of *APOE* $\epsilon 4$ carriers in the sample collection is higher than in the general population (47.2% vs. 10–20% [49]). In the real world of clinical trials

and indeed clinical practice, *APOE* is of less value in identifying people with likely AD pathology as most people in the early stages of disease are *APOE* $\epsilon 4$ negative. Because of this, an algorithm that relies on *APOE* genotype to identify people for inclusion in clinical trials is of limited value, and hence, the identification of additional markers, such as we report here, is even more valuable than a modest increase in AUC over a predictor that includes *APOE* would suggest.

This study is the largest we are aware of to report a plasma biomarker indicative of AD pathology both in terms of the number of proteins assayed and in samples size. Recently, Nakamura et al. (2018) [20] and Ovod et al. (2017) [21] reported methods to measure plasma $A\beta$ directly, and Ashton et al. (2019) [50] used high-resolution mass spectrometry to identify plasma proteins that correlate with altered central amyloid. In all three cases, these assays achieve higher measures of accuracy than we report. However, our study is considerably larger and conducted in a mixed cohort derived from multiple different sources. By virtue of this size and the likely heterogeneity of the sample characteristics, it is possible that the results we report here are likely to withstand the rigors of application in the real-world situations of multi-centre, multinational complex clinical trials. Moreover, the analytical platform we use here is clinic-ready and relatively low cost. Finally, using the panel we describe could have significant economic benefits. Based on the performance of our panel (AUC = 0.78, sensitivity = 0.84, specificity = 0.54), we obtain a likelihood ratio of 1.82, suggesting that the panel could improve approximately 15% in amyloid detection [51] relative to current practice. Given that screen failure rates for amyloid positive recruitment to clinical trials is approximately 70%, then to recruit 100 individuals there is a need to enroll 333 participants with a cost of approximately \$1.7 million (\$5000 per scan). A 15% improvement as we demonstrate would reduce the number needed to screen from 333 to 222, with a cost saving of \$0.55 million per 100 participants of a clinical trial. Therefore, as trials typically seek to recruit large numbers of participants, especially in late phase, we believe that despite the moderate improvement of AUC, these 44 proteins obtained from our large sample size study are strong candidates and worthy of further investigation.

The National Biomarker Development Alliance proposes 6 steps for any biomarker development (nbdabiomarkers.org/) including early discovery, translatable discovery, assay development, assay performance, biomarker quantification, and biomarker validation. Here, we have discovered and replicated 44 nominated biomarkers, the subsequent phase of research will be “translatable discovery,” indicating that we need to replicate these 44 proteins in even larger independent cohorts to further minimize overfitting and to obtain a smaller panel of proteins. One of the advantages of the platform that we used here, SOMAScan, is that it is a high-quality assay platform already in increasingly wide research use and appropriate for clinical implementation. In the short to medium term, the results we present here might be used in

prescreening processes to recruit people to clinical trials where screening includes determination of amyloid status by CSF or PET markers. Currently, the very high screen failure rate at this stage slows clinical development and adds substantially to the cost of clinical trials while also exposing large numbers of potential participants to relatively invasive tests. Our data suggest that this prescreen failure when seeking people with brain pathology could be reduced using a simple blood assay. In future years, an assay such as this could be used as part of a clinical workup directing at risk individuals to specific testing—such as by CSF measures or using PET imaging—before therapeutic intervention.

In addition to predicting amyloid status, the 44 proteins along with age and *APOE* $\epsilon 4$ could also predict other AD endophenotypes including discriminating AD from HC and AD from MCI, differentiating status of t-tau, p-tau, MCI conversion, baseline MMSE score, and hippocampal volume. These findings are in concordance with some recent studies [52,53] and further demonstrate the association of amyloid pathology with other pathologies of AD. Moreover, these findings are potentially useful in clinical trial settings especially as increasing number of clinical trials are directed at tau-related targets [1]. The biomarker panel that we report here could potentially be used in very large numbers of people to identify those likely to have both amyloid and tau pathology, even in preclinical phase.

Despite more than two decades of study, the mechanism of AD pathology is still disputed with serial trials failure provoking some to question the amyloid cascade hypothesis. Broadly speaking, this states that the generation of $A\beta$ leads, through a series of largely unknown steps, to tau pathology and hence neuronal dysfunction and dementia [26,27]. Using the data generated here we confirmed the relationship between amyloid and tau pathology using two approaches. First, the proteins selected from amyloid prediction also predict tau pathology (CSF t-tau and p-tau). Second, we used MR, a method more commonly used in genetic studies to go beyond association to causality, finding that higher amyloid levels drive a moderate increase in both t-tau and p-tau but that the reverse (t-tau driving amyloid pathology) is not true. Our results confirm the crucial role of amyloid in driving tau pathology marking the onset of preclinical AD, and consistent with previous findings [53–55].

Interestingly, no such association was found between amyloid and other three CSF biomarkers (NFL, neurogranin, and YKL-40), either from using proteins to discriminate their levels or from MR analysis. This suggests that these three biomarkers (NFL, neurogranin, and YKL-40) play different roles in AD independent of, or at least remote from, amyloid pathology. Although their role is not understood completely, changes in NFL are thought to reflect axonal damage [56], changes in neurogranin to reflect synaptic dysfunction [57], and changes in YKL-40 to provide a measure of neuroinflammation [58,59]. Our data suggest that amyloid is strongly and causally associated with tau pathology but not with synaptic dysfunction or neuroinflammation. Instead, we found that

tau is causally associated with neurogranin and YKL-40, indicating that synaptic dysfunction and neuroinflammation are consequences of tau-pathological processes. Interestingly, similar observations have been made and are in line with our findings. For example, in multiple transgenic models, amyloid pathology results in very little neuronal damage, synaptic loss, or indeed neuroinflammation, and where such pathological consequences do exist, they are abolished in the context of ablation of tau [60–62]. Although these findings are necessarily preliminary, they suggest that proteomic and perhaps other high-dimensionality molecular biomarker studies might also contribute to an understanding of pathological processes in disease.

We acknowledge that the sample origin is a limitation of this study. Given that both the discovery and replication samples are taken from the same meta-sample collection, it is not an independent validation in the traditional sense. However, because the EMIF-AD MBD study is a multi-centre study and collected samples from 11 centers across Europe, the meta-collection reflects the challenges faced in the real world of multisite and multinational clinical studies. Therefore, the replication obtained in this study in such a collection of samples has a higher probability of effective utility in practice.

In conclusion, our results strongly suggest that testing high-dimensional plasma protein levels is a useful and reproducible approach to measure central nervous system amyloid deposition. This can be potentially applied as a prescreen to preselect patients for further selection procedures for clinical trials, thus reducing the cost incurred to clinical trials by screen failure. Furthermore, we confirmed the causal relationship between amyloid and tau pathology from MR.

Acknowledgments

This research was conducted as part of the EMIF-AD project, which has received support from the Innovative Medicines Initiative Joint Undertaking under EMIF grant agreement no. 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Program (FP7/2007-2013), and EFPIA companies' in-kind contribution. The DESCRIPA study was funded by the European Commission within the 5th framework program (QLRT-2001-2455). The EDAR study was funded by the European Commission within the 5th framework program (contract # 37670). The Leuven cohort was funded by the Stichting voor Alzheimer Onderzoek (grant numbers #11020, #13007, and #15005). R.V. is a senior clinical investigator of the Flemish Research Foundation (FWO). J.S. is currently an employee of UCB, Braine-l'Alleud, Belgium. The San Sebastian GAP study is partially funded by the Department of Health of the Basque Government (allocation 17.0.1.08.12.0000.2.454.01.41142.001.H). The authors acknowledge the contribution of the personnel of the Genomic Service Facility at the VIB-U Antwerp

Center for Molecular Neurology. The research at VIB-CMN is funded in part by the University of Antwerp Research Fund. F.B. is supported by the NIHR biomedical research centre at UCLH. L.S. is funded by DPUK through MRC (grant no. MR/L023784/2) and the UK Medical Research Council Award to the University of Oxford (grant no. MC_PC_17215).

Authors' contributions: S.W., A.L.B., and S.L. contributed to study concept and design, and all authors contributed to sample collection/selection and/or interpretation of data. L.S. and A.J.N.-H. carried out data analysis and interpretation. L.S. drafted the manuscript, and all authors revised the manuscript.

Ethics statement: Written informed consent was obtained from all participants before inclusion in the study. The medical ethics committee at each site approved the study.

Data availability: The data sets generated and analyzed during the present study are available via the EMIF-AD Catalogue via submitted research questions, which have to be approved by the data-owners from each parent cohort

Supplementary Data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.jalz.2019.06.4951>.

RESEARCH IN CONTEXT

1. Systematic review: We reviewed the literature using PubMed sources, meeting abstracts, and presentations. Plasma proteins have been widely studied as candidate biomarkers to predict brain amyloid deposition, but most such studies are targeted to specific proteins or are biased toward high abundant proteins. Furthermore, the mechanism of AD pathology is still disputed, with serial trials failure provoking some to question the amyloid cascade hypothesis.
2. Interpretation: Our findings showed that testing plasma proteins using the high-dimensionality SOMAScan assay could be a useful and reproducible approach for indicating brain amyloid deposition as well as tau pathology. Furthermore, we confirmed the crucial role of amyloid pathology in AD process with a novel use of Mendelian randomization.
3. Future directions: Although it looks promising that a blood test could be used as a tool to preselect individuals for further clinical trials, the plasma biomarkers identified from this study need to be further validated in a larger and independent cohort.

References

- [1] Cummings J, Lee G, Ritter A, Zhong K. Alzheimer's disease drug development pipeline: 2018. *Alzheimers Dement (N Y)* 2018; 4:195–214.
- [2] Cummings J. Lessons Learned from Alzheimer Disease: Clinical Trials with Negative Outcomes. *Clin Transl Sci* 2018;11:147–152.
- [3] Golde TE, DeKosky ST, Galasko D. Alzheimer's disease: The right drug, the right time. *Science* 2018;362:1250–1251.
- [4] Cummings JL. Biomarkers in Alzheimer's disease drug development. *Alzheimers Dement* 2011;7:e13–44.
- [5] Cohen AD, Landau SM, Snitz BE, Klunk WE, Blennow K, Zetterberg H. Fluid and PET biomarkers for amyloid pathology in Alzheimer's disease. *Mol Cell Neurosci* 2019;97:3–17.
- [6] de Almeida SM, Shumaker SD, LeBlanc SK, Delaney P, Marquie-Beck J, Ueland S, et al. Incidence of post-dural puncture headache in research volunteers. *Headache* 2011;51:1503–1510.
- [7] Lista S, Faltraco F, Prvulovic D, Hampel H. Blood and plasma-based proteomic biomarker research in Alzheimer's disease. *Prog Neurobiol* 2013;101–102:1–17.
- [8] Olsson B, Lautner R, Andreasson U, Ohrfelt A, Portelius E, Bjerke M, et al. CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol* 2016; 15:673–684.
- [9] Khan AT, Dobson RJ, Sattler M, Kiddle SJ. Alzheimer's disease: are blood and brain markers related? A systematic review. *Ann Clin Transl Neurol* 2016;3:455–462.
- [10] Carmona P, Molina M, Toledano A. Blood-based biomarkers of Alzheimer's disease: diagnostic algorithms and new technologies. *Curr Alzheimer Res* 2016;13:450–464.
- [11] Baird AL, Westwood S, Lovestone S. Blood-Based Proteomic Biomarkers of Alzheimer's Disease Pathology. *Front Neurol* 2015;6:236.
- [12] Thambisetty M, Lovestone S. Blood-based biomarkers of Alzheimer's disease: challenging but feasible. *Biomark Med* 2010;4:65–79.
- [13] Thambisetty M, Tripaldi R, Riddoch-Contreras J, Hye A, An Y, Campbell J, et al. Proteome-based plasma markers of brain amyloid-beta deposition in non-demented older individuals. *J Alzheimers Dis* 2010;22:1099–1109.
- [14] Ashton NJ, Kiddle SJ, Graf J, Ward M, Baird AL, Hye A, et al. Blood protein predictors of brain amyloid for enrichment in clinical trials? *Alzheimers Dement (Amst)* 2015;1:48–60.
- [15] Westwood S, Leoni E, Hye A, Lyncham S, Khondoker MR, Ashton NJ, et al. Blood-based biomarker candidates of cerebral amyloid using PiB PET in non-demented elderly. *J Alzheimers Dis* 2016;52:561–572.
- [16] Kiddle SJ, Thambisetty M, Simmons A, Riddoch-Contreras J, Hye A, Westman E, et al. Plasma based markers of [11C] PiB-PET brain amyloid burden. *PLoS One* 2012;7:e44260.
- [17] Voyle N, Baker D, Burnham SC, Covin A, Zhang Z, Sangurdekar DP, et al. Blood Protein Markers of Neocortical Amyloid-beta Burden: A Candidate Study Using SOMAscan Technology. *J Alzheimers Dis* 2015;46:947–961.
- [18] Westwood S, Liu B, Baird AL, Anand S, Nevado-Holgado AJ, Newby D, et al. The influence of insulin resistance on cerebrospinal fluid and plasma biomarkers of Alzheimer's pathology. *Alzheimers Res Ther* 2017;9:31.
- [19] Shi L, Baird AL, Westwood S, Hye A, Dobson R, Thambisetty M, et al. A decade of blood biomarkers for Alzheimer's Disease Research: an evolving field, improving study designs, and the challenge of replication. *J Alzheimers Dis* 2018;62:1181–1198.
- [20] Nakamura A, Kaneko N, Villemagne VL, Kato T, Doecke J, Dore V, et al. High performance plasma amyloid-beta biomarkers for Alzheimer's disease. *Nature* 2018;554:249–254.
- [21] Ovod V, Ramsey KN, Mawuenyega KG, Bollinger JG, Hicks T, Schneider T, et al. Amyloid beta concentrations and stable isotope labeling kinetics of human plasma specific to central nervous system amyloidosis. *Alzheimers Dement* 2017;13:841–849.
- [22] Janelidze S, Stomrud E, Palmqvist S, Zetterberg H, van Westen D, Jeromin A, et al. Plasma beta-amyloid in Alzheimer's disease and vascular disease. *Sci Rep* 2016;6:26801.
- [23] Herrup K. The case for rejecting the amyloid cascade hypothesis. *Nat Neurosci* 2015;18:794–799.
- [24] Karran E, De Strooper B. The amyloid cascade hypothesis: are we poised for success or failure? *J Neurochem* 2016;139:237–252.
- [25] Ricciarelli R, Fedele E. The Amyloid Cascade Hypothesis in Alzheimer's Disease: It's Time to Change Our Mind. *Curr Neuropharmacol* 2017;15:926–935.
- [26] Barage SH, Sonawane KD. Amyloid cascade hypothesis: Pathogenesis and therapeutic strategies in Alzheimer's disease. *Neuropeptides* 2015;52:1–18.
- [27] Lemere CA, Masliah E. Can Alzheimer disease be prevented by amyloid-beta immunotherapy? *Nat Rev Neurol* 2010;6:108–119.
- [28] Grover S, Del Greco MF, Stein CM, Ziegler A. Mendelian Randomization. *Methods Mol Biol* 2017;1666:581–628.
- [29] Bos I, Vos S, Vandenberghe R, Scheltens P, Engelborghs S, Frisoni G, et al. The EMIF-AD Multimodal Biomarker Discovery study: design, methods and cohort characteristics. *Alzheimers Res Ther* 2018;10:64.
- [30] Hong S, Dobricic V, Smith RG, Küçükali F, Kilpert F, Bos I, et al. Genomics and epigenomics analyses in the EMIF-AD multimodal biomarker discovery study. *Alzheimer's Dement J Alzheimer's Assoc* 2018;14:P204.
- [31] Ten Kate M, Redolfi A, Peira E, Bos I, Vos SJ, Vandenberghe R, et al. MRI predictors of amyloid pathology: results from the EMIF-AD Multimodal Biomarker Discovery study. *Alzheimers Res Ther* 2018; 10:100.
- [32] Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* 2010;5:e15004.
- [33] Yavorska OO, Burgess S. Mendelian Randomization: an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol* 2017;46:1734–1739.
- [34] Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512–525.
- [35] Angrist JD, Imbens GW. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *J Am Stat Assoc* 1995;90:431–442.
- [36] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* 2010;1:97–111.
- [37] Hemani G, Zheng J, Wade KH, Laurin C, Elsworth B, Burgess S, et al. MR-Base: a platform for systematic causal inference across the phenotype using billions of genetic associations. *BioRxiv* 2016:78972.
- [38] Ogishima S, Mizuno S, Kikuchi M, Miyashita A, Kuwano R, Tanaka H, et al. AlzPathway, an updated map of curated signaling pathways: towards deciphering Alzheimer's Disease pathogenesis. *Methods Mol Biol* 2016;1303:423–432.
- [39] Ogishima S, Mizuno S, Kikuchi M, Miyashita A, Kuwano R, Tanaka H, et al. A map of Alzheimer's disease-signaling pathways: a hope for drug target discovery. *Clin Pharmacol Ther* 2013; 93:399–401.
- [40] Hartmann D, Tournoy J, Saftig P, Annaert W, De Strooper B. Implication of APP secretases in notch signaling. *J Mol Neurosci* 2001; 17:171–181.
- [41] Kopan R, Goate A. Aph-2/Nicastrin: an essential component of gamma-secretase and regulator of Notch signaling and Presenilin localization. *Neuron* 2002;33:321–324.
- [42] Bosco D, Fava A, Plastino M, Montalcini T, Pujia A. Possible implications of insulin resistance and glucose metabolism in Alzheimer's disease pathogenesis. *J Cell Mol Med* 2011;15:1807–1821.
- [43] Zhao W-Q, Townsend M. Insulin resistance and amyloidogenesis as common molecular foundation for type 2 diabetes and Alzheimer's disease. *Biochim Biophys Acta* 2009;1792:482–496.

- [44] Jack CR Jr, Wiste HJ, Weigand SD, Rocca WA, Knopman DS, Mielke MM, et al. Age-specific population frequencies of cerebral β -amyloidosis and neurodegeneration among people with normal cognitive function aged 50–89 years: a cross-sectional study. *Lancet Neurol* 2014;13:997–1005.
- [45] Rowe CC, Ellis KA, Rimajova M, Bourgeat P, Pike KE, Jones G, et al. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol Aging* 2010; 31:1275–1283.
- [46] Morris JC, Roe CM, Xiong C, Fagan AM, Goate AM, Holtzman DM, et al. APOE predicts amyloid-beta but not tau Alzheimer pathology in cognitively normal aging. *Ann Neurol* 2010;67:122–131.
- [47] Jack CR Jr, Wiste HJ, Weigand SD, Knopman DS, Vemuri P, Mielke MM, et al. Age, sex, and APOE epsilon4 effects on memory, brain structure, and beta-amyloid across the adult life span. *JAMA Neurol* 2015;72:511–519.
- [48] Jansen WJ, Ossenkoppele R, Knol DL, Tijms BM, Scheltens P, Verhey FR, et al. Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *Jama* 2015;313:1924–1938.
- [49] Singh PP, Singh M, Mastana SS. APOE distribution in world populations with new data from India and the UK. *Ann Hum Biol* 2006; 33:279–308.
- [50] Ashton NJ, Nevado-Holgado AJ, Barber IS, Lynham S, Gupta V, Chatterjee P, et al. A plasma protein classifier for predicting amyloid burden for preclinical Alzheimer's disease. *Sci Adv* 2019;5. eaau7220.
- [51] McGee S. Simplifying likelihood ratios. *J Gen Intern Med* 2002; 17:646–649.
- [52] Ciarmiello A, Tartaglione A, Giovannini E, Riondato M, Giovacchini G, Ferrando O, et al. Amyloid burden identifies neuropsychological phenotypes at increased risk of progression to Alzheimer's disease in mild cognitive impairment patients. *Eur J Nucl Med Mol Imaging* 2019;46:288–296.
- [53] Jacobs HIL, Hedden T, Schultz AP, Sepulcre J, Perea RD, Amariglio RE, et al. Structural tract alterations predict downstream tau accumulation in amyloid-positive older individuals. *Nat Neurosci* 2018;21:424–431.
- [54] Khan UA, Liu L, Provenzano FA, Berman DE, Profaci CP, Sloan R, et al. Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nat Neurosci* 2014; 17:304–311.
- [55] Pooler AM, Polydoro M, Maury EA, Nicholls SB, Reddy SM, Wegmann S, et al. Amyloid accelerates tau propagation and toxicity in a model of early Alzheimer's disease. *Acta Neuropathol Commun* 2015;3:14.
- [56] Gentil BJ, Tibshirani M, Durham HD. Neurofilament dynamics and involvement in neurological disorders. *Cell Tissue Res* 2015; 360:609–620.
- [57] Lista S, Hampel H. Synaptic degeneration and neurogranin in the pathophysiology of Alzheimer's disease. *Expert Rev Neurother* 2017; 17:47–57.
- [58] Baldacci F, Lista S, Cavedo E, Bonuccelli U, Hampel H. Diagnostic function of the neuroinflammatory biomarker YKL-40 in Alzheimer's disease and other neurodegenerative diseases. *Expert Rev Proteomics* 2017;14:285–299.
- [59] Querol-Vilaseca M, Colom-Cadena M, Pegueroles J, San Martin-Paniello C, Clarimon J, Belbin O, et al. YKL-40 (Chitinase 3-like I) is expressed in a subset of astrocytes in Alzheimer's disease and other tauopathies. *J Neuroinflammation* 2017;14:118.
- [60] Morris M, Knudsen GM, Maeda S, Trinidad JC, Ioanoviciu A, Burlingame AL, et al. Tau post-translational modifications in wild-type and human amyloid precursor protein transgenic mice. *Nat Neurosci* 2015;18:1183–1189.
- [61] Seward ME, Swanson E, Norambuena A, Reimann A, Cochran JN, Li R, et al. Amyloid-beta signals through tau to drive ectopic neuronal cell cycle re-entry in Alzheimer's disease. *J Cell Sci* 2013; 126:1278–1286.
- [62] Ittner LM, Ke YD, Delerue F, Bi M, Gladbach A, van Eersel J, et al. Dendritic function of tau mediates amyloid-beta toxicity in Alzheimer's disease mouse models. *Cell* 2010;142:387–397.