

RESEARCH

Open Access



A plasma protein signature associated with cognitive function in men without severe cognitive impairment

Kanika Mehta^{1,2}, Mohammadreza Mohebbi^{1,3}, Julie A. Pasco^{1,4,5,6}, Lana J. Williams^{1,6}, Sophia X. Sui¹, Ken Walder¹, Boon Lung Ng⁷ and Veer Bala Gupta^{1*}

Abstract

Background A minimally invasive blood-based assessment of cognitive function could be a promising screening strategy to identify high-risk groups for the incidence of Alzheimer's disease.

Methods The study included 448 cognitively unimpaired men (mean age 64.1 years) drawn from the Geelong Osteoporosis Study. A targeted mass spectrometry-based proteomic assay was performed to measure the abundance levels of 269 plasma proteins followed by linear regression analyses adjusted for age and *APOE* $\epsilon 4$ carrier status to identify the biomarkers related to overall cognitive function. Furthermore, two-way interactions were conducted to see whether Alzheimer's disease-linked genetic variants or health conditions modify the association between biomarkers and cognitive function.

Results Ten plasma proteins showed an association with overall cognitive function. This association was modified by allelic variants in genes *ABCA7*, *CLU*, *BDNF* and *MS4A6A* that have been previously linked to Alzheimer's disease. Modifiable health conditions such as mood disorders and poor bone health, which are postulated to be risk factors for Alzheimer's disease, also impacted the relationship observed between protein marker levels and cognition. In addition to the univariate analyses, an 11-feature multianalyte model was created using the least absolute shrinkage and selection operator regression that identified 10 protein features and age associated with cognitive function.

Conclusions Overall, the present study revealed plasma protein candidates that may contribute to the development of a blood-based screening test for identifying early cognitive changes. This study also highlights the importance of considering other risk factors in elucidating the relationship between biomarkers and cognition, an area that remains largely unexplored.

Keywords Cognitive function, Alzheimer's disease, Proteomic analysis, Genotyping, Risk factors

*Correspondence:

Veer Bala Gupta

veer.gupta@deakin.edu.au

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

There is a substantial rise in the number of people living with Alzheimer's disease (AD), which is one of the leading causes of mortality worldwide. High prevalence and associated disability have necessitated a focus on investigating its preclinical phase, where biochemical changes are thought to begin 15–20 years prior to the presentation of clinical symptoms [1]. This is known as the asymptomatic stage where a screening test might be useful in identifying individuals at risk of developing AD. Currently, the most promising sets of biomarkers for identifying preclinical AD constitute amyloid β ($A\beta$) or tau positron emission tomography (PET) imaging and their measurement in the cerebro spinal fluid (CSF) [1–3]. However, these approaches are limited in their use as population screening tools and unlikely to attract high participation rates as PET scans are expensive and only available in specialised centres, while the collection of CSF involves an invasive lumbar puncture procedure that carries the risk of adverse effects. The use of blood biomarkers instead circumvents the above issues and is potentially a more useful screening tool for the general population. Despite this, there is a paucity of studies investigating blood biomarkers associated with cognitive function among non-demented individuals as the majority focus on the identification of dementia biomarkers. There is also a risk of misclassification with the existing plasma biomarkers for amyloid pathology, for example, the $A\beta_{42}/A\beta_{40}$ ratio, precluding their implementation in routine clinical settings [4]. Another challenge with the use of blood biomarkers in clinical settings is the lack of reproducibility, which possibly could be due to the high complexity and wide dynamic range of protein abundances in the blood as well as limited sensitivity offered by immunoassays. However, advancements in the area of targeted mass spectrometry have resulted in highly sensitive and specific multiplexed methods that can accurately quantify low abundant protein markers [5].

The lack of reproducible results can also be due to the multifactorial nature of AD pathology [6]. AD is a complex condition that involves an interplay between several genetic and environmental factors, which are often not considered in biomarker studies [7]. For instance, genetic predisposition has a substantial role in the development of AD, and individuals who are carriers of allelic variants linked to AD risk may have an altered biomarker profile as compared to those who do not harbour the risk alleles. We hypothesised that genetic variants previously linked to AD risk may modify the relationship observed between protein markers and cognitive function. We also expected some of the modifiable physical and mental health conditions associated with AD to have a similar impact on the relationship between protein markers and

cognition. Several studies including systematic reviews and meta-analyses have postulated depression and poor bone health to be risk factors for AD or general cognitive decline [8–12]. In light of AD's complex aetiology, it is important to account for genetic variants and health conditions related to AD to obtain reliable estimates of biomarker levels. Therefore, the primary aim of our study was to identify blood protein markers associated with cognitive function among a population cohort of 448 healthy male participants using a high-throughput mass spectrometric platform. Furthermore, we investigated two-way interactions to see whether AD-linked genetic variants or health conditions modify the association between biomarkers and cognitive function.

Methods

Study cohort, assessment procedures and sample collection

The present study analysed data and blood samples collected from men participating in the Geelong Osteoporosis Study (GOS), an ongoing, prospective population-based study. In brief, age-stratified samples of men and women were selected at random from electoral rolls for the Barwon Statistical Division in south-eastern Australia [13]. The inclusion criteria were a listing on the electoral rolls for the Barwon Statistical Division and residence in the area for a minimum of six months. A total of 1540 men were recruited from 2001 to 2006 (67% participation) and returned for follow-up 5 and 15 years post-recruitment. This study includes a cross-sectional analysis of data and blood samples collected from 448 men during the 15-year follow-up phase (2016–2020). Further information on participant selection is provided in Additional file 1: Fig. S1. As cognitive testing using the CogState Brief Battery (CBB) is currently underway for GOS women, their data could not be included in the present study. Participants were mostly Caucasian (~ 98%). They provided information on their lifestyle and demographic characteristics in addition to undergoing mental and physical health assessments.

Cognitive function was evaluated using a computer-based neuropsychology battery, the CBB, which has been described previously [14–16]. The CBB requires participants to respond to stimuli cards as a part of detection (DET), identification (IDN), one-card learning (OCL) and one-back (OBK) tasks that assess cognitive performance across four domains, namely psychomotor function, visual identification, recognition memory and working memory, respectively. Both a practice trial and a real test were included for each task. The tasks were completed by participants in a quiet room accompanied by a researcher. For the tasks DET, IDN and OBK, scores were calculated by measuring the time (milliseconds) taken

to answer correctly, which was then normalised using a \log_{10} transformation. For the OCL task, scores were calculated based on the accuracy of participant response and normalised using an arcsine square-root transformation. Furthermore, scores for the overall cognitive function (OCF, unitless) were determined by combining the primary measures in the four domains. The present analysis only includes scores from the overall cognitive function, for which higher scores indicate better performance. In addition, participants attended the Mini-Mental State Examination (MMSE) that assessed their overall cognitive function [17]. However, due to inherent limitations associated with MMSE, only OCF scores from the CBB were used in the analysis [18].

Details on sociodemographic variables such as education, smoking and domestic partnership status were acquired from self-reports. Education was defined as a nominal factor based on secondary education completion. Similarly, domestic partnership status was defined as living with a partner (coded "1") or not (coded "0"). Participants who reported smoking at least one cigarette per day were defined as current smokers. The Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, Non-Patient Edition (SCID-I/NP) was used to determine the current mood disorders, as described previously [19]. Bone mineral density (BMD; g/cm^2) was measured at the total hip, femoral neck and lumbar spine (posterior-anterior projection, L2–L4) using dual-energy X-ray absorptiometry (GE Lunar, Prodigy Pro, Madison, WI, USA) [20, 21]. Furthermore, participants reported whether they suffered a major osteoporotic fracture in the last 20 years that included hip, vertebral, wrist or humerus fractures and confirmed using radiology reports as previously described [22]. Fractures that occurred in a motor vehicle accident were not considered. Blood plasma samples were collected after overnight fasting in EDTA tubes at the Australian Clinical Labs and stored at -80°C until use. Cognitive assessments and other health measurements were conducted by trained technicians who were not involved in blood biomarker measurements or data analysis. A schematic representation of the workflow is shown in Fig. 1.

Liquid chromatography-mass spectrometry (LC-MS)-based proteomic analysis

The plasma samples were shipped on dry ice to the University of Victoria – Genome British Columbia Proteomics Centre (BC, Canada) where they were analysed for a panel of 269 proteins. These are common plasma proteins that belong to physiological processes such as inflammation, lipid transport, signalling, oxidative stress and immune response. A complete list of plasma protein

markers included in the assay is provided in Additional file 1: Table S1. A peptide-based targeted quantitation was conducted using multiple reaction monitoring assays following the Clinical Proteomic Tumor Analysis Consortium guidelines for assay development (<https://assays.cancer.gov/>). Tryptic peptides were selected to serve as molecular surrogates for the 269 target proteins according to a series of peptide selection rules (e.g. unique sequence, devoid of oxidisable residues and previous detectability in plasma samples). To help compensate for a matrix-induced suppression or variability in LC-MS performance, $^{13}\text{C}/^{15}\text{N}$ -labelled peptide analogues were used as internal standards. All peptides were synthesised via Fmoc chemistry, purified through reversed-phase high-performance liquid chromatography and characterised via amino acid analysis and capillary zone electrophoresis. A detailed protocol has been described previously [23]. The plasma proteolytic digests were analysed using a triple quadrupole mass spectrometer (Agilent 6495) in the positive ion mode. The data were visualised and examined using the Skyline Quantitative Analysis software (version 21.0.9.139, University of Washington). This involved peak inspection to ensure accurate selection, integration and uniformity in terms of peak shape and retention time. After defining a small number of criteria (i.e. $1/x$ regression weighting, $< 20\%$ deviation in the QC's level's accuracy), a standard curve was used to calculate the peptide concentration in $\text{fmol}/\mu\text{L}$ of plasma.

DNA extraction and genotyping

Total genomic DNA was isolated from buffy coats using the QIAamp® DNA Mini Kit (Qiagen, Hilden, Germany) as per the manufacturer's instructions. The DNA samples were genotyped for the following 11 single nucleotide polymorphisms (SNPs): rs429358 (apolipoprotein E; *APOE* $\epsilon 4$), rs7412 (apolipoprotein E; *APOE* $\epsilon 2$), rs744373 (bridging integrator 1; *BIN1*), rs11136000 (clusterin; *CLU*), rs3764650 (ATP-binding cassette subfamily A member 7; *ABCA7*), rs3818361 (complement receptor 1; *CRI*), rs3851179 (phosphatidylinositol-binding clathrin assembly protein; *PICALM*), rs3865444 (cluster of differentiation 33; *CD33*), rs610932 (membrane spanning 4-domains A6A; *MS4A6A*), rs6265 (brain-derived neurotrophic factor; *BDNF*) and rs9349407 (*CD2*-associated protein; *CD2AP*) at the Australian Genome Research Facility, Brisbane, using the Agena Bioscience MassARRAY® platform. These SNPs are associated with late-onset AD and were selected based on a thorough literature review including the meta-analysis results from the AlzGene database [24]. The carrier status was defined by the presence of at least one copy of the risk allele. No

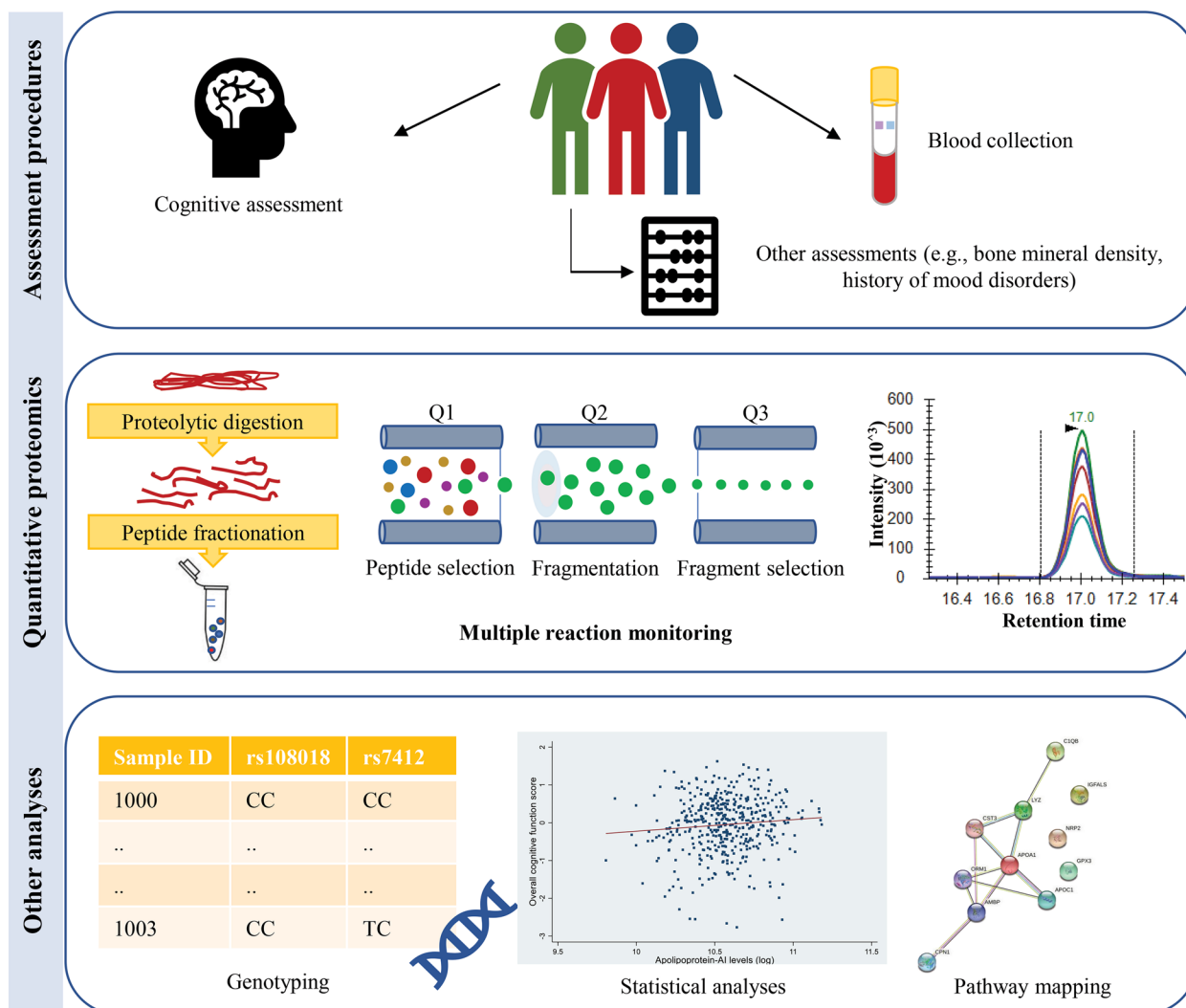


Fig. 1 Schematic representation of the workflow

departure from the Hardy-Weinberg equilibrium was detected. The genetic data analysed in this study are provided in Additional file 1: Table S2.

Statistical analyses

Data cleaning steps were applied to the proteomic data prior to the analyses. For each protein, concentration values outside the limit of quantitation were treated as missing, following which protein analytes with greater than 10% missing values were dropped [3, 25]. For the remaining 125 proteins with more than 90% of data available, missing values were imputed by assigning the lower limit of quantitation/2 value [3]. Furthermore, the protein concentrations were transformed to a natural logarithm to achieve normal distribution. Individual linear regression analyses were conducted to investigate the association between overall cognitive function

and protein concentrations. Age and *APOE* ε4 carrier status were included as confounders in the regression model. As this was an exploratory analysis, correction for multiple testing was not applied.

Next, two-way interactions between proteins that showed a significant association with cognitive function and nine AD-linked genetic polymorphisms [rs744373 (*BIN1*), rs11136000 (*CLU*), rs3764650 (*ABCA7*), rs3818361 (*CR1*), rs3851179 (*PICALM*), rs3865444 (*CD33*), rs610932 (*MS4A6A*), rs6265 (*BDNF*) and rs9349407 (*CD2AP*)] were explored (in separate regression analyses) to see whether the latter affects the relationship between cognition and protein markers. The interaction analyses were adjusted for age and *APOE* ε4 status. Similar interaction analyses were performed to investigate whether comorbidities such as mood disorders and bone loss affect the relationship observed

between cognitive function and protein markers. All statistical analyses were performed using Stata/SE 17.0.

Multianalyte model building and functional mapping of proteins

While the univariate analysis described above revealed associations between a single protein and cognitive function, classifiers and machine learning methods enable the identification of a panel of features that relate to the condition only when considered in combination with one another. Thus, a linear multianalyte model was built using the least absolute shrinkage and selection operator (LASSO) method. A total of 127 features were selected that included 125 protein markers, age and *APOE* ϵ 4 carrier status. A 10-fold cross-validation method was used to develop the model. The optimal regularisation parameter (λ) was estimated through the 10-fold cross-validation process to maximise the out-of-sample R^2 and minimise the mean prediction error. The analysis was performed using Stata/SE 17.0.

Furthermore, complex interactions among proteins identified through LASSO were investigated using bioinformatic tools such as NetworkAnalyst 3.0 (<http://www.networkanalyst.ca/>) [26] and STRING version 11.5 (<https://string-db.org/>) [27]. The list of UniProt protein

accession numbers was uploaded to generate the protein-protein interaction networks. In order to identify the perturbed pathways, pathway mapping was performed using Reactome.org (www.reactome.org) [28].

Results

Participant characteristics are presented in Table 1. The study participants had a mean age of 64.1 years (SD 13.3), and more than three-quarters had completed secondary education (75.5%) and were living with a partner (81.5%). The average MMSE score of the participants was 28.6 (SD 1.7), and only nine had an MMSE score less than 24, which suggests cognitive impairment. Among them, one participant scored 19, another scored 21, two scored 22 and the remaining five scored 23 on MMSE. Overall, 25.9% of all participants carried the *APOE* ϵ 4 risk allele.

Association between plasma protein levels and overall cognitive function

Preliminary data cleaning steps resulted in 125 proteins that were analysed. Of these, 10 proteins, namely apolipoprotein A-I, apolipoprotein C-I, apolipoprotein M, carboxypeptidase N catalytic chain, complement C1q subcomponent subunit B, complement C1q subcomponent subunit C, glutathione peroxidase 3, lysozyme C,

Table 1 Demographic characteristics of the study participants ($n = 448$). Data are presented as mean (SD) or n (%)

| Variable | | Mean (SD)/ n (%) |
|---|-----------------------------------|--------------------|
| Age [years], mean (SD) | | 64.1 (13.3) |
| Education ^a , n (%) | Secondary education completed | 338 (75.5) |
| | Secondary education not completed | 109 (24.3) |
| Domestic partnership status, n (%) | Living with a partner | 365 (81.5) |
| | Not living with a partner | 83 (18.5) |
| Current smoker, n (%) | Yes | 30 (6.7) |
| | No | 418 (93.3) |
| <i>APOE</i> ϵ 4 carriage ^a , n (%) | Yes | 116 (25.9) |
| | No | 323 (72.1) |
| Current mood disorder ^a , n (%) | Yes | 22 (4.9) |
| | No | 37 (8.3) |
| Femoral neck BMD [g/cm^2], mean (SD) | | 1.0 (0.1) |
| Total hip BMD [g/cm^2], mean (SD) | | 1.1 (0.1) |
| Spine BMD [g/cm^2], mean (SD) | | 1.3 (0.2) |
| Occurrence of a major osteoporotic fracture in the last 20 years, n (%) | Yes | 33 (7.4) |
| | No | 415 (92.6) |
| MMSE, mean (SD) | | 28.6 (1.7) |
| CBB-IDN, mean (SD) | | 2.7 (0.1) |
| CBB-DET, mean (SD) | | 2.5 (0.1) |
| CBB-OBK, mean (SD) | | 2.9 (0.1) |
| CBB-OCL, mean (SD) | | 1.0 (0.1) |
| CBB-OCF, mean (SD) | | -0.01 (0.7) |

^aExcludes missing values

neuropilin-2/cystatin-C and alpha-1-microglobulin were found to be significantly associated with cognitive function (Table 2). Among these, the biggest effect size was observed for the protein apolipoprotein A-I; with every 1% rise in plasma apolipoprotein A-I levels, the average cognitive function score increased by 0.004 units.

Interaction between protein markers and genetic variants (linked to AD) for predicting cognitive function

Four genetic variants: rs3764650 (*ABCA7*), rs11136000 (*CLU*), rs6265 (*BDNF*) and rs610932 (*MS4A6A*) that have been previously associated with AD risk, modified the relationship between protein markers and cognitive function (Table 3 and Fig. 2). For instance, individuals with increasing plasma apolipoprotein C-I levels showed a higher score for cognitive function; however, among those carrying the *ABCA7* risk allele, increasing apolipoprotein C-I levels were associated with poorer cognitive function (Fig. 2A). Among the risk allele carriers, with every 1% rise in plasma apolipoprotein C-I levels, the average cognitive function score decreased by 0.004 units. A similar pattern was observed for apolipoprotein C-I's interaction with the risk allele belonging to the *CLU* gene (Fig. 2B). Increasing plasma apolipoprotein C-I levels were associated with a steep rise in cognitive function among individuals who did not carry the risk allele; however, this was greatly diminished among carriers, for whom a minimal change in cognitive function scores was observed with increasing apolipoprotein C-I levels.

For the interaction between protein complement C1q subcomponent subunit C and *BDNF* risk allele, a weak negative association was observed between the protein level and cognitive function among non-carriers

(Fig. 2C). Although the directionality of association remained the same among risk allele carriers, the decline in cognitive function was greater.

Interaction between protein markers and health conditions such as mood disorders and bone health for predicting cognitive function

Six proteins: carboxypeptidase N catalytic chain, glutathione peroxidase 3, apolipoprotein A-I, apolipoprotein C-I, complement C1q subcomponent subunit B and complement C1q subcomponent subunit C, showed a significant interaction with mood disorders or bone health-related variables for predicting overall cognitive function (Table 4 and Fig. 3). An overall negative interaction was observed between carboxypeptidase N catalytic chain and bone mineral density at the femoral neck as shown in Table 4. A closer inspection of the predictive margin plots revealed that for femoral neck BMD values up to 1.0, a positive association between carboxypeptidase N catalytic chain and cognitive function existed (Fig. 3A). However, for BMD values beyond 1.0, increasing carboxypeptidase N catalytic chain was associated with a decline in cognitive function. An opposite trend was seen for the relationship between apolipoprotein A-I levels and cognitive function as modified by spine BMD (Fig. 3C). For spine BMD values up to 1.1, there was a negative association between apolipoprotein A-I levels and cognitive function that changed to a positive association for higher BMD values. Other bone health-related variables included total hip BMD and major osteoporotic fractures in the last 20 years, which also modified the association between some of the protein markers and cognitive function. In addition to variables relating to

Table 2 Results from the linear regression analysis investigating the association between plasma protein levels and overall cognitive function

| Protein | <i>B</i> _{coeff} | 95% CI | t-value | p-value | ^a Partial eta-squared |
|---|---------------------------|----------------|---------|---------|----------------------------------|
| Apolipoprotein A-I | 0.38 | 0.11, 0.65 | 2.80 | 0.005 | 0.02 |
| Apolipoprotein C-I | 0.18 | 0.04, 0.32 | 2.46 | 0.014 | 0.01 |
| Apolipoprotein M | 0.23 | < 0.01, 0.47 | 1.97 | 0.050 | 0.01 |
| Carboxypeptidase N catalytic chain | 0.22 | < 0.01, 0.44 | 2.01 | 0.046 | 0.01 |
| Complement C1q subcomponent subunit B | -0.24 | -0.47, -0.01 | -2.05 | 0.040 | 0.01 |
| Complement C1q subcomponent subunit C | -0.22 | -0.44, -0.01 | -2.01 | 0.045 | 0.01 |
| Glutathione peroxidase 3 | 0.28 | 0.04, 0.51 | 2.29 | 0.022 | 0.01 |
| Lysozyme C | -0.18 | -0.36, -0.01 | -2.06 | 0.040 | 0.01 |
| Neuropilin-2/cystatin-C | -0.21 | -0.41, -0.01 | -2.02 | 0.044 | 0.01 |
| Alpha-1-microglobulin/bikunin precursor | -0.19 | -0.373, -0.004 | -2.01 | 0.045 | 0.01 |

The protein concentrations were transformed into a natural log, and the analyses were adjusted for age and *APOE* ε4 carrier status

^a An eta-squared value of < 0.02 is considered a small effect size

Table 3 Results from the interaction analyses between plasma protein levels and genetic variants for predicting overall cognitive function

| Exposure | B_{coeff} | 95% CI | t-value | p-value | ^a Partial eta-squared |
|--|--------------------|----------------|---------|---------|----------------------------------|
| Interaction between apolipoprotein C-I and rs3764650 (<i>ABCA7</i>) | | | | | |
| Apolipoprotein C-I | 0.17 | 0.03, 0.32 | 2.42 | 0.016 | 0.01 |
| <i>ABCA7</i> _carrier | 0.10 | -0.06, 0.26 | 1.27 | 0.204 | < 0.01 |
| Apolipoprotein C-I# <i>ABCA7</i> _carrier | -0.44 | -0.83, -0.05 | -2.24 | 0.026 | 0.01 |
| Interaction between apolipoprotein C-I and rs11136000 (<i>CLU</i>) | | | | | |
| Apolipoprotein C-I | 0.17 | 0.03, 0.31 | 2.37 | 0.018 | 0.01 |
| <i>CLU</i> _carrier | -0.001 | -0.12, 0.12 | -0.02 | 0.986 | < 0.01 |
| Apolipoprotein C-I# <i>CLU</i> _carrier | -0.29 | -0.58, -0.01 | -2.00 | 0.046 | 0.01 |
| Interaction between complement C1q subcomponent subunit C and rs6265 (<i>BDNF</i>) | | | | | |
| Complement C1q subcomponent subunit C | -0.22 | -0.438, -0.005 | -2.01 | 0.045 | 0.01 |
| <i>BDNF</i> _carrier | -0.01 | -0.13, 0.12 | -0.10 | 0.919 | < 0.01 |
| Complement C1q subcomponent subunit C# <i>BDNF</i> _carrier | -0.58 | -1.06, -0.11 | -2.43 | 0.015 | 0.01 |
| Interaction between glutathione peroxidase 3 and rs610932 (<i>MS4A6A</i>) | | | | | |
| Glutathione peroxidase 3 | 0.28 | 0.04, 0.51 | 2.29 | 0.023 | 0.01 |
| <i>MS4A6A</i> _carrier | -0.02 | -0.15, 0.10 | -0.38 | 0.707 | < 0.01 |
| Glutathione peroxidase 3# <i>MS4A6A</i> _carrier | -0.64 | -1.12, -0.15 | -2.56 | 0.011 | 0.01 |
| Interaction between lysozyme C and rs6265 (<i>BDNF</i>) | | | | | |
| Lysozyme C | -0.19 | -0.36, -0.01 | -2.06 | 0.040 | 0.01 |
| <i>BDNF</i> _carrier | -0.01 | -0.14, 0.11 | -0.23 | 0.815 | < 0.01 |
| Lysozyme C# <i>BDNF</i> _carrier | -0.36 | -0.71, -0.01 | -2.00 | 0.047 | 0.01 |
| Interaction between lysozyme C and rs610932 (<i>MS4A6A</i>) | | | | | |
| Lysozyme C | -0.18 | -0.36, -0.01 | -2.05 | 0.041 | 0.01 |
| <i>MS4A6A</i> _carrier | -0.02 | -0.15, 0.10 | -0.39 | 0.696 | < 0.01 |
| Lysozyme C# <i>MS4A6A</i> _carrier | 0.45 | 0.12, 0.79 | 2.65 | 0.008 | 0.02 |

The protein concentrations were transformed into a natural log, and the analyses were adjusted for age and *APOE* $\epsilon 4$ carrier status. The individual main effects are from a model without the interaction term

^a An eta-squared value of < 0.02 is considered a small effect size

bone health, the presence of a mood disorder impacted the biomarker findings. The interaction between protein markers and mood disorder was examined for only 58 participants. Among individuals with a current diagnosis of a mood disorder, increasing apolipoprotein C-I levels correlated with higher cognitive function, while an inverse pattern was detected for people without a mood disorder (Fig. 3G).

Variable selection and multimarker model building using LASSO

An 11-feature multimarker panel was identified through LASSO in the following order of importance: apolipoprotein A-I, neuropilin-2/cystatin-C, apolipoprotein C-I, alpha-1-microglobulin/bikunin precursor, complement C1q subcomponent subunit B, insulin-like growth factor-binding protein complex acid labile subunit, glutathione peroxidase 3, age, alpha-1-acid glycoprotein 1, lysozyme C and carboxypeptidase N catalytic chain (Additional

file 1: Fig. S2). The cross-validated LASSO model had a mean squared error of 0.38, which was smaller than the total mean squared error (0.56) of the regression model that included the same 11 features identified through LASSO. There was a good overlap between the results obtained from LASSO and individual linear regression analyses as out of the 10 proteins revealed through LASSO, eight were identified through regression analyses (refer to Table 1).

Pathway mapping and interaction network of proteins identified in the multianalyte panel

The proteins identified through LASSO were mapped to different molecular pathways such as amyloid fibre formation, lipoprotein assembly and clearance, innate immune system and transport of insulin-like growth factor. Further bioinformatic analyses revealed an interaction network among these proteins as shown in Additional file 1: Fig. S3. Similar protein-protein

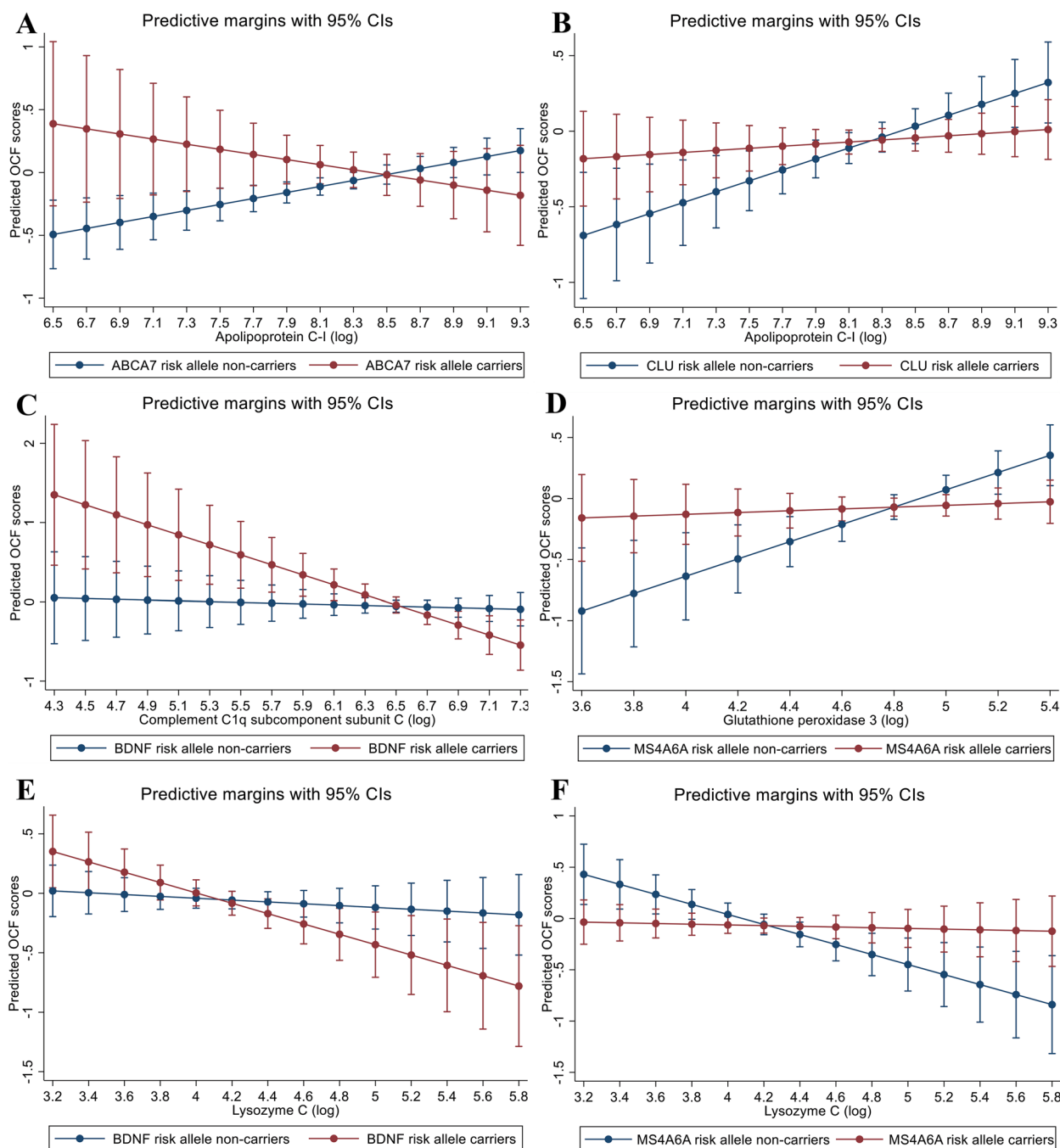


Fig. 2 Predictive margins plots displaying interactions of protein markers with AD-related genetic variants for predicting overall cognitive function. **A** Interaction of apolipoprotein C-I with *ABCA7* risk allele. **B** Interaction of apolipoprotein C-I with *CLU* risk allele. **C** Interaction of complement C1q subcomponent subunit C with *BDNF* risk allele. **D** Interaction of glutathione peroxidase 3 with *MS4A6A* risk allele. **E** Interaction of lysozyme C with *BDNF* risk allele. **F** Interaction of lysozyme C with *MS4A6A* risk allele

interactions were observed with the two tools. Of these 11 proteins (given that neuropilin-2 and cystatin-C are distinct proteins), eight were found to be interacting with one another. Despite having distinct biological

functions, these proteins interacted through more than one pathway, suggesting an underlying physiological link.

Table 4 Results from the interaction analyses between plasma protein levels and variables related to bone health and mood disorder for predicting overall cognitive function

| Exposure | B_{coeff} | 95% CI | t-value | p-value | ^a Partial eta-squared |
|---|--------------------|----------------|---------|---------|----------------------------------|
| Interaction between carboxypeptidase N catalytic chain and femoral neck BMD | | | | | |
| Carboxypeptidase N catalytic chain | 0.20 | −0.02, 0.42 | 1.81 | 0.072 | 0.01 |
| Femoral neck BMD | 0.19 | −0.29, 0.66 | 0.78 | 0.434 | < 0.01 |
| Carboxypeptidase N catalytic chain#Femoral neck BMD | −2.21 | −3.86, −0.57 | −2.65 | 0.008 | 0.02 |
| Interaction between glutathione peroxidase 3 and total hip BMD | | | | | |
| Glutathione peroxidase 3 | 0.36 | 0.12, 0.60 | 2.93 | 0.004 | 0.02 |
| Total hip BMD | 0.23 | −0.19, 0.66 | 1.08 | 0.282 | < 0.01 |
| Glutathione peroxidase 3#Total hip BMD | −1.86 | −3.51, −0.21 | −2.21 | 0.027 | 0.01 |
| Interaction between apolipoprotein A-I and spine BMD | | | | | |
| Apolipoprotein A-I | 0.38 | 0.11, 0.65 | 2.73 | 0.007 | 0.02 |
| Spine BMD | 0.03 | −0.28, 0.35 | 0.20 | 0.840 | < 0.01 |
| Apolipoprotein A-I#Spine BMD | 1.49 | 0.08, 2.89 | 2.08 | 0.038 | 0.01 |
| Interaction between apolipoprotein C-I and major osteoporotic fractures in the last 20 years | | | | | |
| Apolipoprotein C-I | 0.18 | 0.04, 0.32 | 2.47 | 0.014 | 0.01 |
| MOF_20yrs | −0.08 | −0.30, 0.15 | −0.66 | 0.510 | < 0.01 |
| Apolipoprotein C-I#MOF_20yrs | 0.65 | 0.25, 1.06 | 3.19 | 0.002 | 0.02 |
| Interaction between complement C1q subcomponent subunit B and major osteoporotic fractures in the last 20 years | | | | | |
| Complement C1q subcomponent subunit B | −0.24 | −0.470, −0.005 | −2.00 | 0.046 | 0.01 |
| MOF_20yrs | −0.05 | −0.28, 0.18 | −0.45 | 0.654 | < 0.01 |
| Complement C1q subcomponent subunit B#MOF_20yrs | −1.52 | −2.45, −0.59 | −3.21 | 0.001 | 0.02 |
| Interaction between complement C1q subcomponent subunit C and major osteoporotic fractures in the last 20 years | | | | | |
| Complement C1q subcomponent subunit C | −0.22 | −0.434, 0.001 | −1.95 | 0.051 | 0.01 |
| MOF_20yrs | −0.05 | −0.28, 0.18 | −0.42 | 0.674 | < 0.01 |
| Complement C1q subcomponent subunit C#MOF_20yrs | −1.20 | −2.28, −0.12 | −2.19 | 0.029 | 0.01 |
| Interaction between apolipoprotein C-I and a current mood disorder | | | | | |
| Apolipoprotein C-I | −0.02 | −0.42, 0.38 | −0.09 | 0.925 | < 0.01 |
| Current mood disorder | 0.03 | −0.30, 0.36 | 0.19 | 0.852 | < 0.01 |
| Apolipoprotein C-I#Current mood disorder | 0.81 | 0.05, 1.56 | 2.15 | 0.036 | 0.08 |

The protein concentrations were transformed to a natural log, and the analyses were adjusted for age and *APOE* $\epsilon 4$ carrier status. The individual main effects are from a model without the interaction term. MOF_20yr refers to the occurrence of a major osteoporotic fracture in the last 20 years

^a An eta-squared value of < 0.02 is considered a small effect size

Discussion

The present study identified plasma proteins associated with overall cognitive function among ageing men without severe cognitive impairment. In age- and *APOE* $\epsilon 4$ carrier status-adjusted univariate analysis, 10 proteins, which are known to be associated with physiological processes such as lipid transport, immune responses, protection against oxidative damage and transmembrane transport, displayed an association with cognitive function. We suspected that the observed relationship between protein markers and cognition might be impacted by the presence of genetic variants linked to AD risk. Allelic variants belonging to four genes: *ABCA7*, *CLU*, *BDNF* and *MS4A6A*, showed an interaction with

some of the protein markers for predicting cognitive function, suggesting that individuals carrying the risk alleles may show an altered biomarker profile. These interactions may further aid our understanding of the underlying biology of the disease. For instance, the protein apolipoprotein C-I, which plays a major role in lipid metabolism, displayed interaction with risk variants of *ABCA7* and *CLU* genes that also have a role in lipid metabolism. The gene *ABCA7* encodes for a transmembrane protein involved in the packaging of lipids into lipoprotein particles, while *CLU* encodes for an apolipoprotein that transports cholesterol in the brain [29, 30]. These interactions suggest a possible role for lipid transport and metabolism in modulating cognitive function.

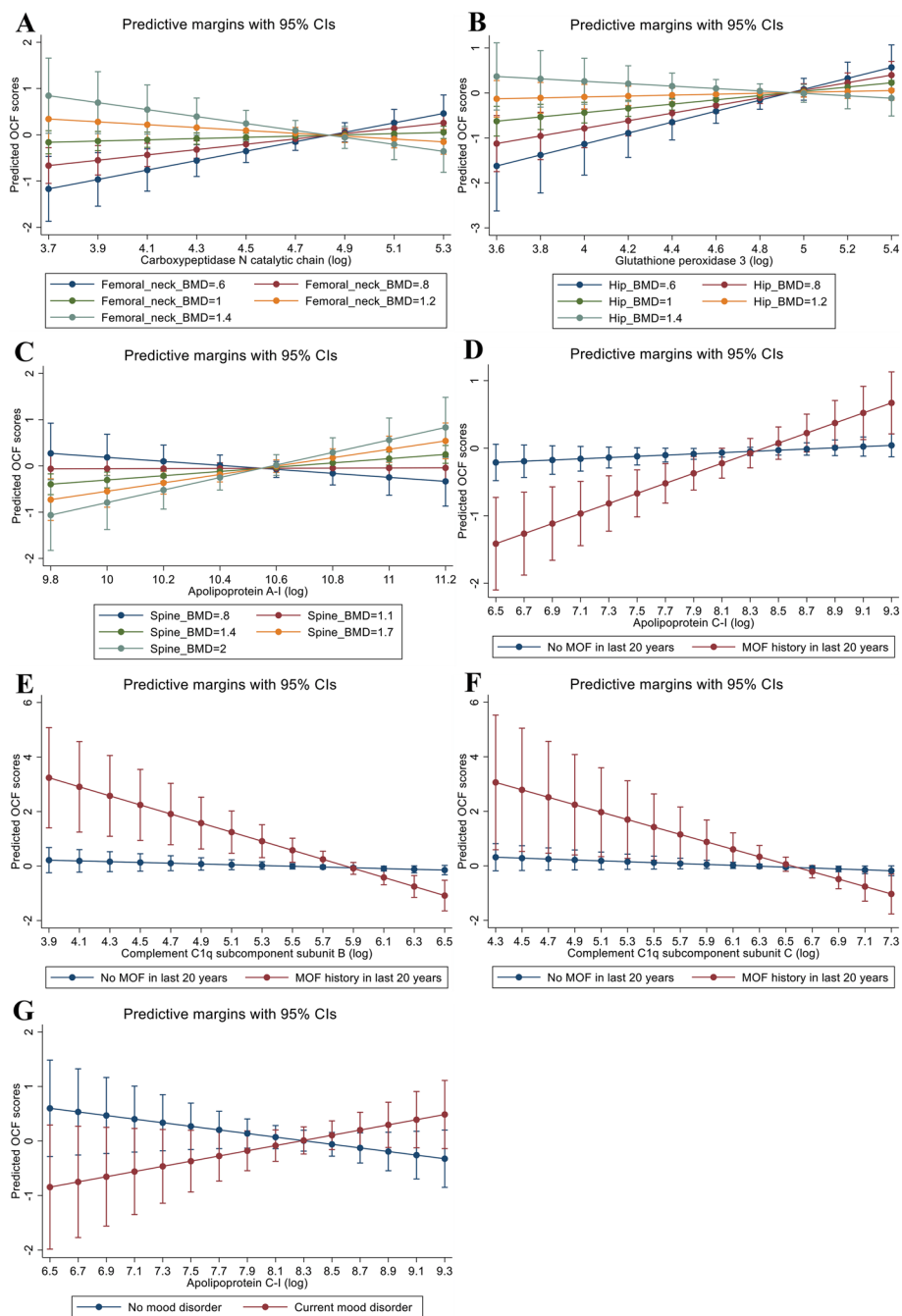


Fig. 3 Predictive margin plots displaying the interactions of protein markers with AD-related health conditions for predicting overall cognitive function. **A** Interaction of carboxypeptidase N catalytic chain with femoral neck BMD. **B** Interaction of glutathione peroxidase 3 with total hip BMD. **C** Interaction of apolipoprotein A-I with spine BMD. **D** Interaction of apolipoprotein C-I with major osteoporotic fractures in the last 20 years. **E** Interaction of complement C1q subcomponent subunit B with major osteoporotic fractures in the last 20 years. **F** Interaction of complement C1q subcomponent subunit C with major osteoporotic fractures in the last 20 years. **G** Interaction of apolipoprotein C-I with a current mood disorder

The effect of genotype on blood protein profiles for AD has been previously demonstrated by Soares et al. who found the *APOE* genotype to be associated with a unique biochemical plasma profile [3]. However, not much is

known about the impact of other risk-conferring genetic variants on plasma biomarkers as *APOE* remains the most studied genetic risk factor to date even though AD

can develop among individuals without the *APOE* $\epsilon 4$ risk allele.

Next, interactions were explored between protein markers and variables related to bone health and mood disorders. Poor bone health and mood disorders, especially depression, are some of the known modifiable risk factors associated with AD, and it is worthwhile to investigate whether their presence influences the relationship between biomarker levels and cognitive function. Bone mineral density at the femoral neck, total hip and spine along with a history of major osteoporotic fracture and diagnosis of a mood disorder, modified the relationship between plasma protein markers and cognition. Overall, the interaction results highlight the need to include genetic variants and other risk factors in biomarker studies, which face the inherent challenge of reproducibility. This is more important for health conditions such as AD, which is not driven by a single causative factor but arises due to a complex interplay between genetic and environmental risk factors. As the analyses above focused on a single protein analyte, a multianalyte panel associated with cognitive function was identified that comprised 10 protein markers and age. Eight of these plasma proteins were also found to be associated with cognitive function through individual regression analyses. Further analysis revealed dysregulation of pathways such as amyloid fibre formation and other biological mechanisms that may not directly relate to $A\beta$ pathology such as lipoprotein assembly and immune responses. Encouragingly, proteins such as apolipoprotein A-I, apolipoprotein C-I and cystatin-C, which were identified through both univariate regression analyses and LASSO, are some of the most widely investigated plasma biomarkers for cognitive decline [31–34]. Among these proteins, the highest effect size was observed for apolipoprotein A-I through regression analyses wherein increasing levels of this protein were associated with better cognitive function. This is consistent with previous studies that have demonstrated that reduced blood apolipoprotein A-I levels are associated with an increased risk of cognitive decline [31, 35, 36].

Although these biomarker findings have the potential to be used for wide-scale population screening, they need to be first replicated across independent population cohorts to ensure high accuracy and reproducibility. Future prospective studies are required to see whether these protein markers can also predict any long-term cognitive change. Despite these challenges, blood-based screening to detect early cognitive changes and identify high-risk individuals offers several advantages over expensive PET imaging with limited availability or CSF measurements that involve an invasive

lumbar puncture. Furthermore, high-throughput mass spectrometric platforms allow simultaneous quantitation of multiple proteins using sample volume as low as 30 μ L, facilitating efficient blood-based screening tests [5].

This was a one-of-a-kind study that highlights the need to investigate the role of genetic risk factors other than *APOE* and health conditions in biomarker analysis. Our study was strengthened by the use of a population-based cohort where participants were drawn at random from the general population and did not comprise individuals with severe cognitive impairment or dementia. Also, an ethnically homogeneous population may yield more precise results. However, our findings may not be generalisable to other populations, and thus, future studies are required to extend these findings to other ethnically diverse cohorts. The present study included only male participants as cognitive function using CBB was evaluated for the first time for the GOS male cohort in their 15-year follow-up phase. We are collecting data for the female cohort in their current follow-up and plan to conduct a similar study for women in the future when their follow-up assessment phase is completed. Also, as this was an exploratory study, in order to mitigate the possibility of a false discovery rate, future confirmatory studies with adjustments for multiple comparisons are expected. Another limitation of our study was the lack of brain imaging data to correlate with biochemical findings.

Overall, this study supports the hypothesis that a relationship between plasma protein levels and cognitive function exists, and a blood-based proteomic signature can be exploited to enrich participants for more comprehensive AD testing. Early screening of high-risk groups may provide an opportunity to make lifestyle or pharmacological interventions before the onset of clinical symptoms, which become apparent only after irreversible neurological damage has occurred [2]. The study also underscores the importance of including information on genetic risk factors and other health conditions associated with AD in order to establish reproducible biomarkers.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13195-023-01294-7>.

Additional file 1: Fig. S1. GOS participant selection flow diagram. **Fig. S2.** Features identified using LASSO regression for predicting cognitive function. **Fig. S3.** Interaction networks of proteins identified in the multi-analyte panel. A Interaction networks generated using NetworkAnalyst. B Interaction networks generated using STRINGS. *APOA1 = Apolipoprotein A-I, CPN1 = Carboxypeptidase N catalytic chain, CST3 = Cystatin-C, ORM1 = Alpha-1-acid glycoprotein 1, AMBP = Alpha-1-microglobulin/bikunin precursor, APOC1 = Apolipoprotein C-I, LYZ = Lysozyme C and C1QB =

Complement C1q subcomponent subunit B. **Table S1.** List of 269 proteins measured in blood plasma samples. **Table S2.** The distribution of risk alleles among the study participants.

Acknowledgements

We acknowledge the men who participated in the study and the staff who contributed to the data collection. We would also like to thank the University of Victoria – Genome British Columbia Proteomics Centre, Canada, and the Australian Genome Research Facility for conducting the proteomic and genotyping analysis, respectively.

Authors' contributions

K.M., J.A.P., L.J.W. and V.B.G. conceived the study. K.M. was responsible for the experiments, shipment of samples and conducted the statistical analyses in consultation with M.M. The funding was acquired by J.A.P., L.J.W. and V.B.G. Recruitment of study cohort and assessment procedures were overseen by J.A.P., L.J.W. and S.X.S. The original draft was written by K.M. and the manuscript was revised by M.M., J.A.P., L.J.W., K.W. and B.L.N.

Funding

The Geelong Osteoporosis Study was funded by the National Health and Medical Research Council (NHMRC) Australia [Grant Nos. 299831, 628582]. KM was supported by the Deakin University Postgraduate Research Scholarship (DUPRS). LJW is supported by an NHMRC Emerging Leadership Fellowship (1174060). VBG is supported by NHMRC (RM34909). The funding bodies had no involvement in the study design, data collection, data analysis and preparation and submission of the manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the senior authors upon reasonable request.

Declarations

Ethics approval and consent to participate

All participants provided written informed consent to participate in the study, which was approved by the Human Research Ethics Committee at Barwon Health (00/56-E7) and Deakin University (2013-320). All procedures performed were in accordance with the ethical standards of the institutional and national research committees and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Deakin University, IMPACT – The Institute for Mental and Physical Health and Clinical Translation, School of Medicine, Geelong, VIC 3216, Australia. ²Baker Heart and Diabetes Institute, Melbourne, VIC, Australia. ³Biostatistics Unit, Faculty of Health, Deakin University, Burwood, VIC, Australia. ⁴Department of Medicine-Western Health, The University of Melbourne, St Albans, VIC, Australia. ⁵Department of Epidemiology and Preventive Medicine, Monash University, Prahran, VIC, Australia. ⁶Barwon Health, Geelong, VIC, Australia. ⁷Department of Geriatric Medicine, Barwon Health, Geelong, VIC, Australia.

Received: 5 May 2023 Accepted: 21 August 2023

Published online: 01 September 2023

References

- 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:eaa07220.
- Burnham SC, Faux NG, Wilson W, Laws SM, Ames D, Bedo J, et al. A blood-based predictor for neocortical A β burden in Alzheimer's disease: results from the AIBL study. *Mol Psychiatry.* 2014;19:519–26.
- Soares HD, Potter WZ, Pickering E, Kuhn M, Immermann FW, Shera DM, et al. Plasma biomarkers associated with the apolipoprotein E genotype and Alzheimer disease. *Arch Neurol.* 2012;69:1310–7.
- Zetterberg H, Schott JM. Blood biomarkers for Alzheimer's disease and related disorders. *Acta Neurol Scand.* 2022;146:51–5.
- Bhardwaj M, Weigl K, Tikk K, Holland-Letz T, Schrotz-King P, Borchers CH, et al. Multiplex quantitation of 270 plasma protein markers to identify a signature for early detection of colorectal cancer. *Eur J Cancer.* 2020;127:30–40.
- Bazenet C, Lovestone S. Plasma biomarkers for Alzheimer's disease: much needed but tough to find. *Biomark Med.* 2012;6:441–54.
- Reitz C, Mayeux R. Alzheimer disease: epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochem Pharmacol.* 2014;88:640–51.
- Mehta K, Thandavan SP, Mohebbi M, Pasco JA, Williams LJ, Walder K, et al. Depression and bone loss as risk factors for cognitive decline: a systematic review and meta-analysis. *Ageing Res Rev.* 2022;76:101575.
- Gallagher D, Kiss A, Lanctot K, Herrmann N. Depressive symptoms and cognitive decline: a longitudinal analysis of potentially modifiable risk factors in community dwelling older adults. *J Affect Disord.* 2016;190:235–40.
- John A, Patel U, Rusted J, Richards M, Gaysina D. Affective problems and decline in cognitive state in older adults: a systematic review and meta-analysis. *Psychol Med.* 2019;49:353–65.
- Laudisio A, Fontana DO, Rivera C, Ruggiero C, Bandinelli S, Gemma A, et al. Bone mineral density and cognitive decline in elderly women: results from the InCHIANTI study. *Calcif Tissue Int.* 2016;98:479–88.
- Lui LY, Stone K, Cauley JA, Hillier T, Yaffe K. Bone loss predicts subsequent cognitive decline in older women: the study of osteoporotic fractures. *J Am Geriatr Soc.* 2003;51:38–43.
- Pasco JA, Nicholson GC, Kotowicz MA. Cohort profile: Geelong Osteoporosis Study. *Int J Epidemiol.* 2012;41:1565–75.
- Sui SX, Holloway-Kew KL, Hyde NK, Williams LJ, Leach S, Pasco JA. Muscle strength and gait speed rather than lean mass are better indicators for poor cognitive function in older men. *Sci Rep.* 2020;10:10367.
- Sui SX, Williams LJ, Holloway-Kew KL, Hyde NK, Anderson KB, Tembo MC, et al. Skeletal muscle density and cognitive function: a cross-sectional study in men. *Calcif Tissue Int.* 2021;108:165–75.
- Sui SX, Williams LJ, Holloway-Kew KL, Hyde NK, Leach S, Pasco JA. Associations between muscle quality and cognitive function in older men: cross-sectional data from the Geelong Osteoporosis Study. *J Clin Densitom.* 2022;25:133–40.
- Pasco JA, Stuart AL, Sui SX, Holloway-Kew KL, Hyde NK, Tembo MC, et al. Dynapenia and low cognition: a cross-sectional association in postmenopausal women. *J Clin Med.* 2021;10:173.
- Devenney E, Hodges JR. The Mini-Mental State Examination: pitfalls and limitations. *Pract Neurol.* 2017;17:79–80.
- Williams LJ, Jacka FN, Pasco JA, Coulson CE, Quirk SE, Stuart AL, et al. The prevalence and age of onset of psychiatric disorders in Australian men. *Aust N Z J Psychiatry.* 2016;50:678–84.
- Rufus-Membere P, Holloway-Kew KL, Diez-Perez A, Kotowicz MA, Pasco JA. Associations between bone material strength index, calcaneal quantitative ultrasound, and bone mineral density in men. *J Endocr Soc.* 2021;5:bvaa179.
- Chandrasekaran V, Pasco JA, Stuart AL, Brennan-Olsen SL, Berk M, Hodge JM, et al. Anticonvulsant use and bone health in a population-based study of men and women: cross-sectional data from the Geelong Osteoporosis Study. *BMC Musculoskelet Disord.* 2021;22:172.
- Pasco JA, Lane SE, Brennan-Olsen SL, Holloway KL, Timmeny EN, Bucki-Smith G, et al. The epidemiology of incident fracture from cradle to senescence. *Calcif Tissue Int.* 2015;97:568–76.
- Gaither C, Popp R, Mohammed Y, Borchers CH. Determination of the concentration range for 267 proteins from 21 lots of commercial human plasma using highly multiplexed multiple reaction monitoring mass spectrometry. *Analyst.* 2020;145:3634–44.
- Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet.* 2007;39:17–23.

25. Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam CP, et al. Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Arch Neurol*. 2012;69:1318–25.
26. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res*. 2019;47:W234–W41.
27. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49:D605–12.
28. Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res*. 2022;50:D687–92.
29. Ramirez LM, Goukasian N, Porat S, Hwang KS, Eastman JA, Hurtz S, et al. Common variants in *ABCA7* and *MS4A6A* are associated with cortical and hippocampal atrophy. *Neurobiol Aging*. 2016;39:82–9.
30. Almeida JFF, Dos Santos LR, Trancozo M, de Paula F. Updated meta-analysis of *BIN1*, *CR1*, *MS4A6A*, *CLU*, and *ABCA7* variants in Alzheimer's disease. *J Mol Neurosci*. 2018;64:471–7.
31. Song F, Poljak A, Crawford J, Kochan NA, Wen W, Cameron B, et al. Plasma apolipoprotein levels are associated with cognitive status and decline in a community cohort of older individuals. *PLoS One*. 2012;7:e34078.
32. Hu WT, Holtzman DM, Fagan AM, Shaw LM, Perrin R, Arnold SE, et al. Plasma multianalyte profiling in mild cognitive impairment and Alzheimer disease. *Neurology*. 2012;79:897–905.
33. Hye A, Riddoch-Contreras J, Baird AL, Ashton NJ, Bazenet C, Leung R, et al. Plasma proteins predict conversion to dementia from prodromal disease. *Alzheimer's Dement*. 2014;10:799–807.e2.
34. Johnstone D, Milward EA, Berretta R, Moscato P, Alzheimer's Disease Neuroimaging I. Multivariate protein signatures of pre-clinical Alzheimer's disease in the Alzheimer's Disease Neuroimaging Initiative (ADNI) plasma proteome dataset. *PLoS One*. 2012;7:e34341.
35. Kawano M, Kawakami M, Otsuka M, Yashima H, Yaginuma T, Ueki A. Marked decrease of plasma apolipoprotein AI and AII in Japanese patients with late-onset non-familial Alzheimer's disease. *Clin Chim Acta*. 1995;239:209–11.
36. Liu HC, Hu CJ, Chang JG, Sung SM, Lee LS, Yuan RY, et al. Proteomic identification of lower apolipoprotein A-I in Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2006;21:155–61.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

