

Blood-based biomarker candidates of cerebral amyloid using PiB PET in non-demented elderly

Sarah Westwood^{a,b}, Emanuela Leoni^{b,c,d}, Abdul Hye^b, Steven Lynham^e, Mizanur R Khondoker^{b,f}, Nicholas J. Ashton^b, Steven J. Kiddle^g, Alison L. Baird^a, Ricardo Sainz-Fuertes^{b,h}, Rufina Leung^b, John Grafⁱ, Cristina Tan Hehirⁱ, David Baker^j, Cristina Cereda^d, Chantal Bazenet^b, Malcolm Ward^e, Madhav Thambisetty^k, Simon Lovestone^{a,*}

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

^bKing's College London, Institute of Psychiatry, Psychology and Neuroscience, and NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK

^cDepartment of Brain and Behavioural Science, University of Pavia, Pavia, Italy

^dLaboratory of Experimental Neurobiology, "C. Mondino" National Institute of Neurology Foundation, Pavia, Italy

^eProteomics Core Facility, Centre of Excellence for Mass Spectrometry, Institute of Psychiatry, Kings College London, London, UK

^fDepartment of Applied Health Research, University College London, London, UK

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

^hPartnerships in Care, North London Clinic, London, UK

ⁱDiagnostics, Imaging and Biomedical Technologies, GE Global Research, Niskayuna, NY, USA

^jJanssen R&D, Neurosciences, Titusville, NJ, USA

^kUnit of Clinical and Translational Neuroscience, Laboratory of Behavioral Neuroscience, National Institute on Ageing, Baltimore, MD, USA

*Corresponding author. Correspondence to: Simon Lovestone, Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, OX3 7JX. Tel.: +44 (0)1865 223910; Fax: +44 (0)1865 793101; simon.lovestone@psych.ox.ac.uk

Running Title: Blood biomarkers of PiB PET cerebral amyloid

Abstract

Increasingly, clinical trials for Alzheimer's disease are being conducted earlier in the disease phase and with biomarker confirmation using *in vivo* amyloid PET imaging or CSF tau and A β measures to quantify pathology. However, making such a pre-clinical AD diagnosis is relatively costly and the screening failure rate is likely to be high. Having a blood based marker that would reduce such costs and accelerate clinical trials through identifying potential participants with likely pre-clinical AD would be a substantial advance.

In order to seek such a candidate biomarker, discovery phase proteomic analyses using 2DGE and gel-free LC-MS/MS for high and low molecular weight analytes were conducted on longitudinal plasma samples collected over a twelve year period from non-demented older individuals who exhibited a range of ^{11}C -PiB PET measures of amyloid load. We then sought to extend our discovery findings by investigating whether our candidate biomarkers were also associated with brain amyloid burden in disease, in an independent cohort. Seven plasma proteins including A2M, Apo-A1 and multiple complement proteins, were identified as pre-clinical biomarkers of amyloid burden and were consistent across three time points ($p < 0.05$). Five of these proteins also correlated with brain amyloid measures at different stages of the disease ($q < 0.1$).

Here we show that it is possible to detect a plasma based biomarker signature indicative of AD pathology at a stage long before the onset of clinical disease manifestation. As in previous studies, acute phase reactants and inflammatory markers dominate this signature.

Keywords: Alzheimer disease; Biological Markers; Preclinical; Plasma; Blood; Amyloid; Proteomics

Introduction

Although there have been considerable advances in our understanding of Alzheimer's disease (AD) the results of clinical trials testing AD treatments have so far been disappointing [1, 2]. A recent study suggested about 99.6% of trials aimed at preventing or reversing the disease between 2002 and 2012 have failed [3]. Only a few pharmaceutical interventions have been approved to treat AD and these only temporarily treat symptoms and are not disease modifying. Potential explanations for the failure of disease-modifying therapies to date could include i) the relatively advanced stage of disease targeted in clinical trials ii) the high rate of absence of AD pathology in participants included in these trials. This suggests that trials will increasingly be targeting the pre-clinical AD population and therefore necessitating the use of biomarkers such as positron emission tomography (PET) imaging or cerebrospinal fluid (CSF) measures to confirm AD pathology. However, such measures are relatively invasive and costly and current screening failure rates are high. Having a blood marker that identifies participants with likely pre-clinical AD would provide a pre-screening filter for PET imaging or CSF measures, reducing screening failures and accelerating clinical trials. Moreover, given the ease of repeated sampling of blood, such a measure might find utility as a companion biomarker in clinical trials.

Research has suggested that there is a peripheral signal of disease in the blood of people with clinically diagnosed AD [4, 5]. Moreover, such a signal has also been shown to reflect AD brain pathology [6-8]. Therefore a peripheral measure may also be a valuable biomarker in the preclinical stages. In 2013, for example, we reported a protein signature in non-demented elderly people from the Baltimore Longitudinal Study of Ageing (BLSA) that could differentiate between individuals who would go on to show amyloid pathology using PET imaging ten years later [9]. In the present study we aimed to expand upon this work by

increasing the coverage of the proteome investigated and by identifying proteins that are stable in their ability across time to reflect amyloid as a continuous measure in cognitively healthy individuals.

The present study used two complementary proteomic techniques for the discovery of biologically relevant plasma biomarkers of pre-clinical AD. Two Dimensional Gel Electrophoresis (2DGE) was used to identify larger proteins (~20-250kDa), whereas gel-free liquid chromatography-tandem mass spectrometry (LC-MS/MS) methodology was tailored to focus on the identification of low molecular weight (LMW) proteins (<30kDa). Plasma proteins were assessed for their relationship to global brain amyloid load at three time points; 12 years before *in vivo* amyloid PET imaging, 6 years before amyloid PET imaging, and concurrently to amyloid PET imaging. Proteins that were consistently associated with amyloid PET at all three time points were investigated further as candidate biomarkers throughout the disease in an independent cohort, using pre-existing data generated by Ashton et al [7].

Material and Methods

Discovery cohort: Baltimore Longitudinal Study of Ageing neuroimaging sub-study (BLSA-NI)

The BLSA was initiated in 1958 as a prospective longitudinal study of healthy human ageing [10-14]. The primary objective of the study is to understand the physical and cognitive changes associated with normal ageing. The BLSA neuroimaging sub-study is currently in its 19th year and includes a subset of BLSA volunteers who return annually for imaging and clinical evaluations. The neuroimaging sub-study is described in detail by Resnick et al [15].

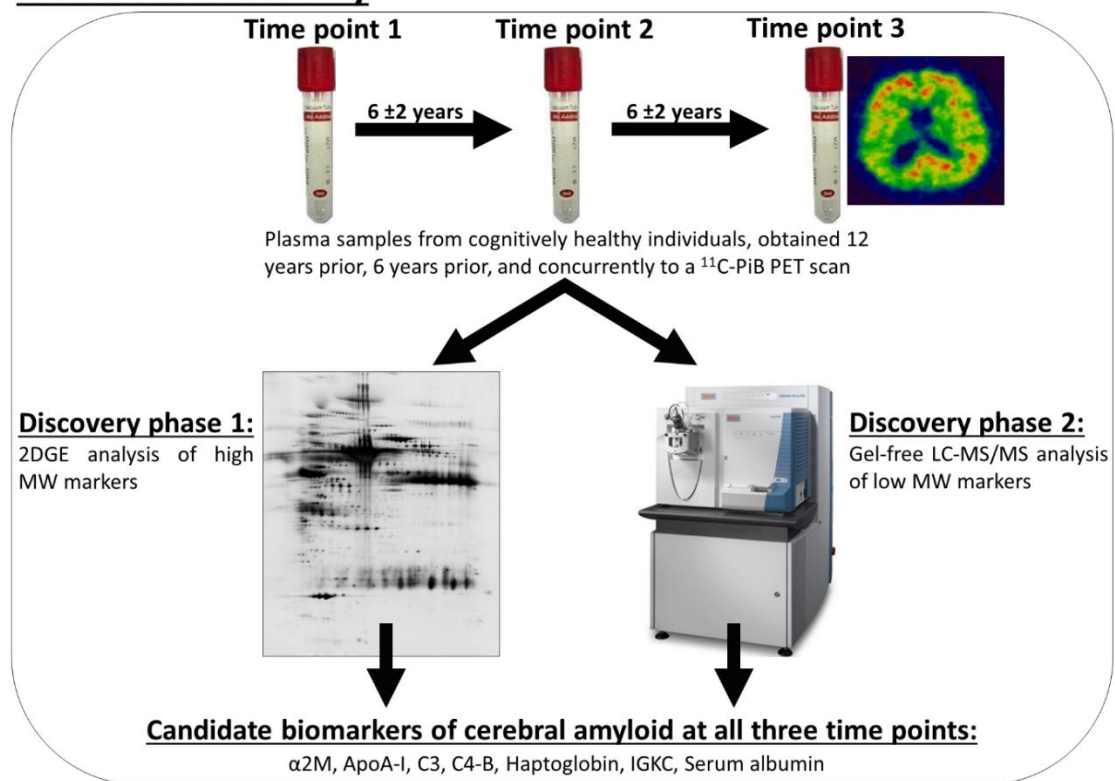
Discovery cohort: assessments, blood collection and processing

In the initial 2DGE experiments (discovery phase 1) we examined plasma samples from 54 subjects from the BLSA-NI, whose demographics are shown in Table 1. All subjects had serial blood sampling over a 12 year period prior to a ^{11}C -PiB PET scan. Blood samples were obtained in the morning after an overnight fast, centrifuged at 3000 rpm for 30 minutes at 4°C, and plasma collected and stored in 1.5ml aliquots at -80°C. This study was approved by the local institutional review board. All participants provided written informed consent.

Plasma samples from three time points approximately 6 years apart were selected for proteomic analysis; time point 1 = 12 \pm 2 years prior to the PET scan, time point 2 = 6 \pm 2 years prior to the PET scan, and time point 3 = concurrent to the PET scan (Figure 1). Plasma samples from a subset of 25 of these subjects, matched for age and gender, were further analysed for the discovery of low molecular weight proteins as biomarkers (discovery phase 2).

Extensive neuropsychological testing was completed at each time-point and apolipoprotein E genotyping was performed for each participant [16]. All subjects remained cognitively normal throughout the 12 year period of this study.

Biomarker Discovery



Extension study

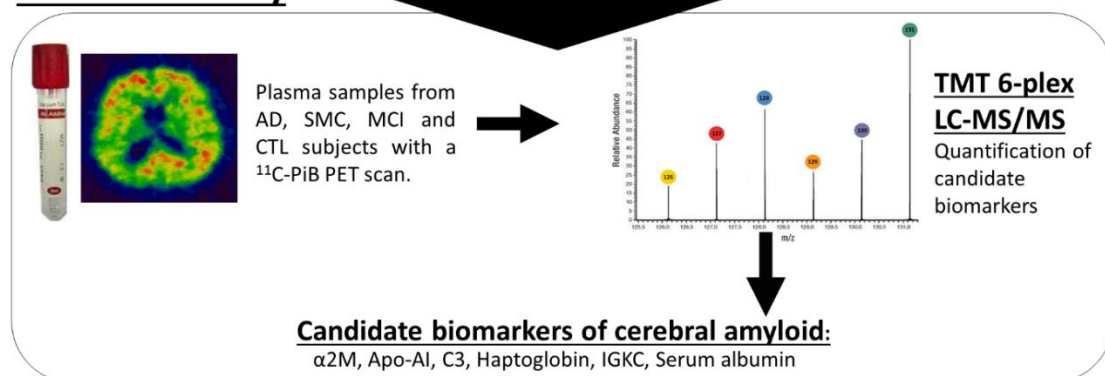


Figure 1: Schematic diagram illustrating the experimental design and work flow of the present study for discovery of plasma proteins associated with brain amyloid burden in non-demented and demented older individuals.

Abbreviations: 2DGE, Two Dimensional Gel Electrophoresis; α 2M, alpha-2-macroglobulin; AD, Alzheimer's disease; Apo-AI, Apolipoprotein A-I; C3, Complement C3; C4-B, Complement C4-B; CTL, Control; IGKC, Ig Kappa Chain C Region; LC-MS/MS, Liquid Chromatography Tandem Mass Spectrometry; MCI, Mild Cognitive Impairment; MW, Molecular Weight; SMC, Subjective Memory Complaint; TMT, Tandem Mass Tagging.

BLSA-NI ^{11}C -PiB PET

The PiB PET imaging methodology of the BLSA-NI study is detailed elsewhere [17]. The mean cortical distribution volume ratio (DVR) was calculated by averaging values from orbitofrontal, prefrontal, superior frontal, parietal, lateral temporal, occipital, and anterior and posterior cingulate regions. Each of the 54 subjects had one ^{11}C -PiB PET scan taken concurrently to their final plasma sample (time point 3).

Discovery study, phase 1: High molecular weight studies using two dimensional gel electrophoresis (2DGE)

2DGE analysis was performed as detailed in Hye et al (2006) [4]. Each sample was analysed in duplicate. Gel images were analysed using Progenesis SameSpots software (Non-linear Dynamics) and normalised volumes for each of the protein spots were exported. As each sample had been analysed in duplicate the coefficient of variation (CV) and average normalised volumes were calculated for every duplicate spot. Average volumes were selected for statistical analysis, unless the CV was >40% then the spot closest in value to the overall normalised spot volume mean was selected. A script was written in R (version 3.0.0) to complete the above pre-processing.

Protein identification by liquid chromatography – tandem mass spectrometry (LC-MS/MS)

2DGE spots that had statistically significant changes in volume associated with ^{11}C -PiB PET DVR were manually excised from 2DGE preparative gels. The proteins present in a given spot were trypsin digested and the resulting peptides were subjected to LC-MS/MS for protein identification. A Thermo EASY n-LC II system (Thermo Scientific) was coupled to an LTQ Orbitrap Velos Pro for LC-MS/MS protein identification. Peptides were resolved by reverse phase chromatography (35 minutes) using 20x collision-induced dissociation (CID) scans

following each Fourier transform mass spectrometry (FTMS) scan (2x μ Scans at 30,000 resolving power @ 400 m/z). CID was carried out on 20 of the most intense ions from each FTMS scan.

Pre-processing of LC-MS/MS data

Raw data files produced in Excalibur software (Thermo Scientific) were processed using Proteome Discoverer V1.3 (Thermo Scientific) and peptide spectral matches and subsequent protein identifications were determined using Mascot (v2.3; <http://www.matrixscience.com>).

Discovery study, phase 2: Low molecular weight (LMW) using gel-free LC-MS/MS

Gel-free LC-MS/MS was performed to provide a complementary evaluation of the proteome to that already conducted using 2DGE. Prior depletion of abundant and high MW plasma proteins allowed for the extraction and identification of proteins with a LMW (<30kDa). Samples were first passed through an albumin and IgG depletion column (Proteoprep Immunoaffinity Albumin and IgG Depletion Kit from Sigma), then remaining large proteins were removed using a 30kDa cut off filter (Amicon Ultrafree-MC centrifugal filter tubes from Merck Millipore). The LMW filtrate was then trypsin digested and LC-MS/MS analysed, as described previously, with triplicate injections.

LMW LC-MS/MS data pre-processing

Raw data was processed as previously described in discovery phase 1. Data was then loaded into Scaffold 4 (Proteome Software Inc., USA) allowing the merging of triplicate sample injections into one BioSample. Proteins were filtered using the following cut off values: protein threshold = 95%, minimum number of peptides = 2, peptide threshold = 80%. Proteins with more than 50% missing data, decoy proteins, and any contaminants (keratin and any trypsin)

were removed. To enable comparisons across samples the Normalised Total Spectra quantitative value was selected for statistical analysis.

Extension cohort: The Australian Imaging, Biomarkers and Lifestyle Flagship Study of Ageing (AIBL)

The AIBL study, launched in 2006, is a prospective longitudinal study aiming to discover biomarkers, cognitive characteristics, and health and lifestyle factors contributing to the subsequent development of Alzheimer's disease (<http://www.aibl.csiro.au/>). The study design including additional subject recruitment, diagnosis and PET imaging have been described previously [18].

Extension cohort: assessments, blood collection and processing

This study further examined pre-existing proteomic data from 78 subjects from the AIBL study, generated by Ashton et al [7], which included; 19 healthy controls, 31 subjective memory complaints (SMC), 22 mild cognitive impairment (MCI) and 6 AD patients, all of which had undergone ^{11}C -PiB PET scans. Detailed diagnostic criteria and further cohort information is available in Ellis et al 2009 [18]. Subjects had been selected to be enriched for clear cases of PiB negativity (PiB-) and positivity (PiB+) based upon a standardized uptake value ratio (SUVR) cut off of 1.3. Standardised clinical assessments included Mini-Mental State Examination (MMSE), and apolipoprotein E (APOE) genotypes were available. The details of blood collection and sample processing have been previously reported [8]. Demographics and PiB group distributions are shown in Table 2.

AIBL ^{11}C -PiB PET

The PiB PET imaging methodology of the AIBL study is detailed elsewhere [19]. The SUVR values were calculated using the cerebellar grey matter as the reference region and neocortical

amyloid burden (NAB) was determined using the average SUVR of the mean of frontal, superior parietal, lateral temporal, lateral occipital, and anterior and posterior cingulate regions.

Extension Study: Tandem Mass Tag LC-MS/MS

Tandem mass tag (TMT) labelling enables a high throughput multiplex approach for LC-MS/MS analysis. TMT 6Plex (Thermo Scientific) combined with LC-MS-MS therefore allows the use of the same methodology as in gel-free LC-MS/MS discovery phase experiments but now multiplexed with up to six samples including a reference sample measured simultaneously. Ashton et al [7] had previously performed TMT 6Plex LC-MS/MS on 78 samples from the AIBL study, and therefore this pre-existing dataset was here further analysed. Protein identifications observed in different gel fractions were considered as separate entities defined as protein MW isoforms.

Statistical analysis; discovery and extension studies

All statistical analyses were performed using R (version 3.0.0). Discovery data was analysed using partial correlation; proteomic data was correlated with ^{11}C -PiB PET DVR whilst controlling for six covariates (age, gender, education (years), body mass index (BMI), cholesterol, and APOE $\epsilon 4$ status). For 2DGE data a penalised regression analyses using elastic net penalty [20] with leave-one-out cross validation (LOOCV) was additionally used to select a smaller subset of important spots which could be picked for LC-MS/MS analysis. Collinearity is common in 2DGE datasets perhaps as different forms of the same protein may be present within multiple 2DGE spots. An elastic-net regression model was chosen as it deals with overfitting and accounts for multi-collinearity between predictors (spots) through encouraging a grouping effect where strongly correlated predictors tend to be in or out of the model together. This type of regression model is also preferable for high-dimensional data where the number of

explanatory variables is larger than the sample size ($p > n$). LOOCV was selected as the most appropriate cross validation approach for the sample size.

Extension study data was obtained from a pre-existing dataset generated by Ashton et al (2015) [7]. Briefly, protein MW isoforms had undergone the following statistical tests; Mann-Whitney U test and logistic regression (PiB+ and PiB- group comparison) and Spearman Rank Correlation (SRC) and linear regression (associations with PiB retention as a continuous measure). Each test had been completed for both the mean and median protein values, resulting in a total of eight statistical tests per protein MW isoform. For logistic and linear regression, age, gender and APOE $\epsilon 4$ status were used as covariates. For the purpose of this study, we targeted only the protein MW isoforms of interest for statistical analyses and re-calculated false discovery rate corrections for this subset of data.

Results: Discovery

We aimed in this study to identify blood based markers indicative of brain amyloid in cognitively healthy people using samples obtained from the BLSA. Furthermore we sought to identify proteins which remained stable in their biomarker ability over time. We then examined whether these proteins were also associated with PET amyloid in disease using samples and data obtained from the AIBL study. In both cases we predicated plasma studies not on clinical status but on PET imaging using in vivo ^{11}C -PiB PET derived measures of global brain amyloid burden. In order to explore the large dynamic range of proteins detected by MS/MS we utilised a combination of gel based and gel free methods of protein separation and also depleted the samples of high molecular weight proteins.

Discovery phase 1: 2DGE identification of high MW plasma protein markers of global ^{11}C -PiB PET DVR

We first performed 2DGE on 54 participants in the BLSA-NI, all without dementia at each assessment at three time points; 12 years before PET imaging, 6 years before PET imaging and concurrently with PET PiB imaging. In total, 1386 discrete silver stained protein spots, each representing proteins with similar molecular weight and charge characteristics, were identified as being well defined and present on each gel.

Partial correlation was performed to identify 2DGE protein spots that were significantly related to cerebral amyloid load whilst controlling for age, gender, education, BMI, cholesterol, and APOE $\epsilon 4$ status. Normalised spot volumes correlating with ^{11}C -PiB PET DVR at each time point, whilst controlling for covariates were identified (119, 47, and 73 protein spots nominally significant at baseline, time point 2, and time point 3 respectively, $p < 0.05$). 2DGE protein spots showing significant correlation were visually re-assessed to verify suitability for identification by mass spectrometry. Spots which were too small or faint to be accurately located by eye were excluded from further analysis. After these visual inspections the number of spots remaining were: 53 spots for time point 1, 6 spots for time point 2, and 33 spots for time point 3. These proteins were taken forward for further regression analysis. The above steps are summarised in Figure 2.

2DGE protein spots consistently identified on every gel

= 1386



Partial correlation with ^{11}C -PiB-PET DVR

	Time point 1	Time point 2	Time point 3
Significant spots:	119	47	73



Visual inspection

	Time point 1	Time point 2	Time point 3
Visible spots:	53	6	33



Penalised elastic-net regression model

Figure 2: Flow chart illustrating the steps taken for two-dimensional gel electrophoresis protein spot analysis.

Abbreviations: DVR, Distribution Volume Ratio; PiB, Pittsburgh compound B; 2DGE, Two Dimensional Gel Electrophoresis.

For each time point an elastic-net penalised regression model with LOOCV was performed on the remaining spots with seven covariates included; age, gender, BMI, cholesterol, education (years), APOE ϵ 4 status, and prediction time (years between plasma sample and ^{11}C -PiB PET scan). The elastic-net model with the minimum cross-validated Mean Squared Error (MSE) was selected as the model of interest. To obtain a measure of how well these elastic-net protein models were predicting PET amyloid load, multiple linear regressions were performed. The resulting regression models could explain 83%, 53% and 90% of the variance in ^{11}C -PiB PET DVR levels 12 years later, 6 years later, and concurrently (Table 3).

The protein spots on 2DGE gels that were included in each regression model and hence had some value in predicting PiB amyloid load were selected for identification with mass spectrometry. Manual spot picking of significant spots was performed as described in Hye et al (2006) [4]. The proteins found to be abundantly present in the 2DGE spots significantly related to ^{11}C -PiB PET DVR at each of the three time points are listed in the supplementary tables Supplementary Table 1, Supplementary Table 2, and

Supplementary Table 3.

Discovery phase 2: LC-MS/MS identification of LMW plasma protein markers of global PiB PET DVR

We then performed a gel-free tandem mass spectrometry experiment using a subset of the same sample set in order to supplement the 2DGE/MS data that preferentially examined abundant high molecular weight proteins. The analysis of the resulting data followed broadly the same workflow first using partial correlation with ^{11}C -PiB PET DVR whilst controlling for covariates (age, gender, education (years), BMI, cholesterol, and APOE e4 status) to identify proteins which have a relationship with amyloid load. Partial correlations at baseline, time point 2, and time point 3 revealed 5, 2, and 2 proteins, respectively, which nominally significantly correlated with ^{11}C -PiB PET DVR 12 years later, 6 years later and concurrently ($p < 0.05$) (data shown in Supplementary tables Supplementary Table 4, Supplementary Table 5, and Supplementary Table 6).

Combining data from the two discovery approaches

Table 4 combines together all of the proteins identified as candidate biomarkers in the two discovery phases. Seven proteins were found as candidate biomarkers for ^{11}C -PiB PET DVR consistently across all three time points; alpha-2-macroglobulin, apolipoprotein A-I, complement C3, complement C4B, haptoglobin, Ig kappa chain C region, and serum albumin.

Results: Extension study

TMT LC-MS/MS performed on AIBL subjects

Proteins identified as candidate biomarkers across all three time points were taken forward for further investigation. We then wanted to measure each of these discovery phase proteins simultaneously in multiple samples from an entirely independent cohort to allow for relative

quantitation in replication phase study. TMT combined with LC-MS/MS is an approach that allows the use of the same methodology to analyse proteins but now multiplexed with up to six samples including a reference sample simultaneously. This approach had previously been performed by Ashton et al [7] on 78 samples from the AIBL study; TMT LC-MS/MS had been performed on all samples and after a stringent data clean up procedure (detailed in Ashton et al, 2015, [7]) 381 protein MW isoforms were confidently measured, representing 116 unique protein groups. Of our 7 candidate biomarkers from the discovery phase analyses all but one (Complement C4b) were detected in this dataset. As protein identifications observed in different gel MW fractions were considered as separate entities these 7 proteins consisted of 45 protein MW isoforms which were then targeted for statistical analyses.

A total of 15 protein MW isoforms passed at least one statistical test (uncorrected $p < 0.05$). After false discovery rate (FDR) correction 8 protein MW isoforms passed at least one statistical test ($q < 0.1$) (shown in Table 5). These 8 protein MW isoforms represented $\alpha 2M$, serum albumin, APO-A1, C3, and haptoglobin. This result adds support to the discovery findings that these 5 proteins perform as biomarkers of PET amyloid load, and importantly indicates that their ability remains stable throughout different stages of Alzheimer's disease.

Discussion

It would be of considerable benefit to AD clinical trials if markers in peripheral fluids could be found that correlated with pathology specific markers derived from PET or CSF measurements. Such markers might be less specific than direct measures of pathological proteins but might make a substantial contribution to reducing the costs of screen failure and accelerate identification of potential participants to interventional trials. As part of efforts to search for such markers, plasma proteins were assessed for their relationship to amyloid load 12 years

prior, 6 years prior, and concurrently to estimation of cerebral amyloid pathology using a ^{11}C -PiB PET scan.

This study used two complementary proteomic techniques for the discovery of plasma biomarkers of pre-clinical AD pathology. Together these two techniques provide a relatively comprehensive proteome analysis of plasma; a tissue challenging for biomarker discovery both because of its complexity and the dynamic range of protein concentration and size. Each technique identified unique candidate biomarkers and three candidate proteins overlapped between the two techniques, providing a form of technical replication.

Seven proteins correlated with amyloid burden at all three time points and were taken forward to extension experiments in an independent cohort comprised of cognitively healthy individuals as well as those with subjective memory concerns, mild cognitive impairments, and Alzheimer's disease. Of these proteins, complement C4b was the only protein not detected using the multiplexing analysis, possibly because of a failure of TMT labelling or because C4b has high peptide sequence homology with another protein detected, such as complement C4a, and consequently peptides may have been misidentified. Five of the six proteins detected in these extension studies showed a significant association with PET amyloid load even when FDR corrections were applied. These five proteins are: α 2M, serum albumin, Apo-A1, C3, and haptoglobin. Together these results provide strong support for the ability of these proteins to reflect amyloid burden throughout the disease course, from a pre-clinical phase through to established clinical syndromes.

All five of these proteins have been previously identified in plasma as candidate biomarkers associated with AD disease status and/or pathology [4, 9, 21-26]. In fact α 2M and C3 are two of the most reproducible plasma protein markers of AD, associating with AD-related phenotypes in five independent cohorts [5]. It is encouraging to see the same proteins

consistently reported as plasma AD biomarkers across studies. The findings here add support for their utility as AD biomarkers in the preclinical stage, and also demonstrates that their biomarker abilities remain stable over time.

In the discovery phase experiments a number of proteins were identified as biomarkers of PET amyloid load at 1 or 2 time points. Due to the statistical approach used (penalised regression model) it is not possible to rule out that these proteins may have also performed as biomarkers at all three time points but were omitted from the regression model due to collinearity with another protein. Of particular interest are fibrinogen chains alpha, beta and gamma, all of which were significantly associated with PET amyloid load at two time points, and have been identified in previous studies as biomarkers of PET amyloid load in AD cohorts [26, 27]. Furthermore fibrinogen has been shown to cross the blood-brain-barrier and co-localise with amyloid plaques in AD brain tissue and AD animal models [28-30]. The results of this study therefore replicate the potential biomarker ability of fibrinogen and furthermore indicate its possible utility during the preclinical stage of the disease.

The proteins found here to be implicated in pre-clinical AD are involved in a variety of systems, including; complement and inflammation, coagulation, haemostasis, fibrinolysis, and blood-brain-barrier integrity. Previous research has implicated these systems in AD pathogenesis [31-34]. Such systems may be causally related to AD onset, or be affected as a result of the disease. Additionally the involvement of these systems in AD pathology may be inter-linked or they may act independently, further work is needed to clarify these relationships. It may be possible to identify the resulting contribution of some of these pathways (e.g. chronic inflammation) as high risk factors for the development of AD in healthy individuals, comparable to that of more-established risk factors such as APOE ϵ 4 status. At the very least, the presence of these biomarkers in healthy elderly individuals indicates potential early intervention targets worthy

of further investigation. Further work examining how these plasma proteins may be associated with the longitudinal cognitive decline and brain atrophy data available within the BLSA-NI would also be beneficial to see whether similar biological systems are again highlighted.

We acknowledge the limitations of this study largely caused by the relatively small cohort of subjects used in this investigation. Though we believe the advantages of the longitudinal replication achieved in the discovery phases compensate for this, and enable the detection of proteins robust in their biomarker ability over time. The unbalanced discovery phase cohorts are also a limitation of this work to be considered, however this design was necessary due to sample availability and cost limitations. Validation within other healthy elderly cohorts with amyloid PET scans is also critically needed to assess the candidate markers ability to perform across cohorts and studies. We believe the proteomic markers identified in our study will serve as a useful starting point for future validation studies for such cohorts.

A further limitation of this study may be the lack of standardisation in the approaches used for the amyloid PET quantitative analyses between the discovery study (DVR) and the extension study (SUVR). In some cases differing DVR and SUVR quantitative approaches may impact upon the classification of amyloid positivity, however we believe the effects of using the two different approaches to be minimal in this study. In the discovery phase analyses subjects were not dichotomised into groups and the statistical analysis applied identified proteins which could predict continuous amyloid burden. As DVR and SUVR values are often reported to correlate the resulting markers should be minimally affected by the use of either method. The extension study did use an SUVR cut off (1.3) to classify subjects as amyloid negative/positive, however the cohort was enriched for clear cases of PiB negativity and positivity (mean negative SUVR = 1.1, mean positive SUVR = 2.3). Therefore the likelihood of these subjects being identically classified as amyloid negative/positive via an alternative analysis approach is high.

In conclusion, this study adds further evidence that plasma is a rich resource for AD biomarkers. Biological variation and the dynamic range of plasma complicates the discovery, but through the use of complementary techniques these issues can be minimised. Here we show it is also possible to detect these plasma based biomarkers at a stage long before the onset of clinical disease manifestations. We identified numerous candidate biomarkers, many of which have been previously been reported as AD biomarkers in the clinical stage of the disease, providing further evidence for such biomarkers and supporting their utility as biomarkers longitudinally throughout the disease. Quantitative assays are now needed to validate the proteins reported here as candidate biomarkers in the pre-clinical stage. If validated these proteins may, independently or in combination, provide a blood based screening measure to enrich populations for clinical trial recruitment, and allow repeated testing as companion biomarkers in such clinical trials.

Acknowledgements

We are grateful for grant funding from the Alzheimer's Society, Alzheimer's Research UK, and the NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at the South London and Maudsley NHS Foundation Trust and Kings College London. We are grateful to participants in the Baltimore Longitudinal Study of Aging for their invaluable contribution. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging. A portion of this work was funded by GE Healthcare and Janssen R&D. SJK is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1). The authors acknowledge data used in the preparation of this article obtained from The Australian, Imaging, Biomarkers and Lifestyle Study of Ageing (AIBL) which is a collaboration between CSIRO, Edith Cowan University (ECU), The Florey

Institute of Neuroscience and Mental Health (FINMH), National Ageing Research Institute (NARI) and Austin Health (www.AIBL.csiro.au).

References

- [1] Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B, Joffe S, Kieburtz K, Raman R, Sun X, Aisen PS, Siemers E, Liu-Seifert H, Mohs R (2014) Phase 3 Trials of Solanezumab for Mild-to-Moderate Alzheimer's Disease. *New England Journal of Medicine* **370**, 311-321.
- [2] Salloway S, Sperling R, Fox NC, Blennow K, Klunk W, Raskind M, Sabbagh M, Honig LS, Porsteinsson AP, Ferris S, Reichert M, Ketter N, Nejadnik B, Guenzler V, Miloslavsky M, Wang D, Lu Y, Lull J, Tudor IC, Liu E, Grundman M, Yuen E, Black R, Brashear HR (2014) Two Phase 3 Trials of Bapineuzumab in Mild-to-Moderate Alzheimer's Disease. *New England Journal of Medicine* **370**, 322-333.
- [3] Cummings JL, Morstorf T, Zhong K (2014) Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimer's Research & Therapy* **6**, 37-37.
- [4] Hye A, Lynham S, Thambisetty M, Causevic M, Campbell J, Byers HL, Hooper C, Rijdsdijk F, Tabrizi SJ, Banner S, Shaw CE, Foy C, Poppe M, Archer N, Hamilton G, Powell J, Brown RG, Sham P, Ward M, Lovestone S (2006) Proteome-based plasma biomarkers for Alzheimer's disease. *Brain* **129**, 3042-3050.
- [5] Kiddle SJ, Sattlecker M, Proitsi P, Simmons A, Westman E, Bazenet C, Nelson SK, Williams S, Hodges A, Johnston C, Soininen H, Kłoszewska I, Mecocci P, Tsolaki M, Vellas B, Newhouse S, Lovestone S, Dobson RJB (2014) Candidate Blood Proteome Markers of Alzheimer's Disease Onset and Progression: A Systematic Review and Replication Study. *Journal of Alzheimer's Disease* **38**, 515-531.
- [6] Voyle N, Baker D, Burnham SC, Covin A, Zhang Z, Sangurdekar DP, Tan Hehir CA, Bazenet C, Lovestone S, Kiddle S, Dobson RJ (2015) Blood Protein Markers of Neocortical Amyloid-beta Burden: A Candidate Study Using SOMAscan Technology. *J Alzheimers Dis* **46**, 947.
- [7] Ashton NJ, Kiddle SJ, Graf J, Ward M, Baird AL, Hye A, Westwood S, Wong KV, Dobson RJ, Rabinovici GD, Miller BL, Rosen HJ, Torres A, Zhang Z, Thurfjell L, Covin A, Hehir CT, Baker D, Bazenet C, Lovestone S (2015) Blood protein predictors of brain amyloid for enrichment in clinical trials? *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* **1**, 48-60.
- [8] Burnham SC, Faux NG, Wilson W, Laws SM, Ames D, Bedo J, Bush AI, Doecke JD, Ellis KA, Head R, Jones G, Kiiveri H, Martins RN, Rembach A, Rowe CC, Salvado O, Macaulay SL, Masters CL, Villemagne VL (2014) A blood-based predictor for neocortical A β burden in Alzheimer's disease: Results from the AIBL study. *Molecular Psychiatry* **19**, 519-526.
- [9] Thambisetty M, Tripaldi R, Riddoch-Contreras J, Hye A, An Y, Campbell J, Sojkova J, Kinsey A, Lynham S, Zhou Y, Ferrucci L, Wong DF, Lovestone S, Resnick SM (2010) Proteome-based plasma markers of brain amyloid-beta deposition in non-demented older individuals. *J Alzheimers Dis* **22**, 1099-1109.
- [10] Stone JL, Norris AH (1966) Activities and attitudes of participants in the Baltimore longitudinal study. *J Gerontol* **21**, 575-580.
- [11] Rowe JW, Andres R, Tobin JD, Norris AH, Shock NW (1976) The effect of age on creatinine clearance in men: a cross-sectional and longitudinal study. *J Gerontol* **31**, 155-163.
- [12] Schneider EL (1979) Cell replication and aging: in vitro and in vivo studies. *Fed Proc* **38**, 1857-1861.
- [13] Tzankoff SP, Norris AH (1979) Age-related differences in lactate distribution kinetics following maximal exercise. *Eur J Appl Physiol Occup Physiol* **42**, 35-40.

- [14] Shock NW, Greulich RC, Costa PT, Jr., Andres R, Lakatta EG, Arenberg D, Tobin JD (1984) *Normal Human Aging: The Baltimore Longitudinal Study on Aging*, NIH Publication, Washington, D.C., Government Printing Office.
- [15] Resnick SM, Pham DL, Kraut MA, Zonderman AB, Davatzikos C (2003) Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *J Neurosci* **23**, 3295-3301.
- [16] Thambisetty M, Beason-Held L, An Y, Kraut MA, Resnick SM (2010) APOE ϵ 4 genotype and longitudinal changes in cerebral blood flow in normal aging. *Archives of neurology* **67**, 93-98.
- [17] Sojkova J, Beason-Held L, Zhou Y, An Y, Kraut MA, Ye W, Ferrucci L, Mathis CA, Klunk WE, Wong DF, Resnick SM (2008) Longitudinal Cerebral Blood Flow and Amyloid Deposition: An Emerging Pattern? *Journal of Nuclear Medicine* **49**, 1465-1471.
- [18] Ellis KA, Bush AI, Darby D, De Fazio D, Foster J, Hudson P, Lautenschlager NT, Lenzo N, Martins RN, Maruff P, Masters C, Milner A, Pike K, Rowe C, Savage G, Szeoke C, Taddei K, Villemagne V, Woodward M, Ames D (2009) The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr* **21**, 672-687.
- [19] Rowe CC, Ellis KA, Rimajova M, Bourgeat P, Pike KE, Jones G, Fripp J, Tochon-Danguy H, Morandau L, O'Keefe G, Price R, Raniga P, Robins P, Acosta O, Lenzo N, Szeoke C, Salvado O, Head R, Martins R, Masters CL, Ames D, Villemagne VL (2010) Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol Aging* **31**, 1275-1283.
- [20] Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **67**, 301-320.
- [21] Thambisetty M, Hye A, Foy C, Daly E, Glover A, Cooper A, Simmons A, Murphy D, Lovestone S (2008) Proteome-based identification of plasma proteins associated with hippocampal metabolism in early Alzheimer's disease. *Journal of Neurology* **225**, 1712-1720.
- [22] Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam CP, Mondal A, Bedo J, Bush AI, Brown B, De Ruyck K, Ellis KA, Fowler C, Gupta VB, Head R, Macaulay SL, Pertile K, Rowe CC, Rembach A, Rodrigues M, Rumble R, Szeoke C, Taddei K, Taddei T, Trounson B, Ames D, Masters CL, Martins RN (2012) Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Archives of Neurology* **69**, 1318-1325.
- [23] O'Bryant SE, Xiao G, Barber R, Reisch J, Hall J, Cullum CM, Doody R, Fairchild T, Adams P, Wilhelmsen K, Diaz-Arrastia R, Texas Alzheimer's R, Care C (2011) A Blood-Based Algorithm for the Detection of Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders* **32**, 55-62.
- [24] Hye A, Riddoch-Contreras J, Baird AL, Ashton NJ, Bazenet C, Leung R, Westman E, Simmons A, Dobson R, Sattlecker M, Lupton M, Lunnon K, Keohane A, Ward M, Pike I, Zucht HD, Pepin D, Zheng W, Tunnicliffe A, Richardson J, Gauthier S, Soininen H, Kłoszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S (2014) Plasma proteins predict conversion to dementia from prodromal disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* **10**, 799-807.e792.
- [25] Thambisetty M, Simmons A, Hye A, Campbell J, Westman E, Zhang Y, Wahlund L-O, Kinsey A, Causevic M, Killick R, Kłoszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Spenger C, Lovestone S, for the AddNeuroMed c (2011) Plasma Biomarkers of Brain Atrophy in Alzheimer's Disease. *PLoS ONE* **6**, e28527.

- [26] Kiddle SJ, Thambisetty M, Simmons A, Riddoch-Contreras J, Hye A, Westman E, Pike I, Ward M, Johnston C, Lupton MK, Lunnon K, Soininen H, Kloszewska I, Tsolaki M, Vellas B, Mecocci P, Lovestone S, Newhouse S, Dobson R, for the Alzheimers Disease Neuroimaging I (2012) Plasma Based Markers of 11C PiB-PET Brain Amyloid Burden. *PLoS ONE* **7**, e44260.
- [27] Ashton N, Kiddle S, Graf J, Ward M, Baird A, Hye A, Westwood S, Wong k, Dobson R, Rabinovici G, Miller B, Rosen H, Torres A, Zhang Z, Thurfjell L, Covin A, Hehir C, Baker D, Bazenet C, Lovestone S, Ashton N, Kiddle S, Graf J, Ward M, Baird A, Hye A, Westwood S, Wong k, Dobson R, Rabinovici G, Miller B, Rosen H, Torres A, Zhang Z, Thurfjell L, Covin A, Hehir C, Baker D, Bazenet C, Lovestone S (2015) Blood protein predictors of brain amyloid for enrichment in clinical trials? *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring* **1**.
- [28] Cortes-Canteli M, Paul J, Norris EH, Bronstein R, Ahn HJ, Zmolodchikov D, Bhuvanendran S, Fenz KM, Strickland S (2010) Fibrinogen and β -Amyloid Association Alters Thrombosis and Fibrinolysis: A Possible Contributing Factor to Alzheimer's Disease. *Neuron* **66**, 695-709.
- [29] Klohs J, Baltes C, Prinz-Kranz F, Ratering D, Nitsch RM, Knuesel I, Rudin M (2012) Contrast-enhanced magnetic resonance microangiography reveals remodeling of the cerebral microvasculature in transgenic ArcA β mice. *Journal of Neuroscience* **32**, 1705-1713.
- [30] Ryu JK, McLarnon JG (2009) A leaky blood-brain barrier, fibrinogen infiltration and microglial reactivity in inflamed Alzheimer's disease brain. *Journal of Cellular and Molecular Medicine* **13**, 2911-2925.
- [31] Cortes-Canteli M, Zmolodchikov D, Ahn HJ, Strickland S, Norris EH (2012) Fibrinogen and altered hemostasis in Alzheimer's disease. *J Alzheimers Dis* **32**, 599-608.
- [32] Kitazawa M, Cheng D, Tsukamoto MR, Koike MA, Wes PD, Vasilevko V, Cribbs DH, LaFerla FM (2011) Blocking IL-1 signaling rescues cognition, attenuates tau pathology, and restores neuronal beta-catenin pathway function in an Alzheimer's disease model. *J Immunol* **187**, 6539-6549.
- [33] Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schurmann B, Heun R, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frolich L, Hampel H, Hull M, Rujescu D, Goate AM, Kauwe JSK, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Muhleisen TW, Nothen MM, Moebus S, Jockel K-H, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* **41**, 1088-1093.
- [34] Montagne A, Barnes SR, Sweeney MD, Halliday MR, Sagare AP, Zhao Z, Toga AW, Jacobs RE, Liu CY, Amezcua L, Harrington MG, Chui HC, Law M, Zlokovic BV (2015) Blood-brain barrier breakdown in the aging human hippocampus. *Neuron* **85**, 296-302.

Table 1: Discovery cohort; Demographics of selected subjects from the BLSA-NI

		N	Mean age at time point 1*	Mean MMSE at time point 1	APOE ε4 (carrier/non-carrier)	Mean ¹¹C-PiB PET cortical DVR	Mean no. years between first plasma sample and ¹¹C-PiB PET scan
Discovery phase 1	Female	29	67.62 +/- 6.55	29.25	8/21	1.17	11.39
	Male	25	65.64 +/- 5.54	29.13	7/18	1.13	11.45
	Total	54	66.70	29.18	15/39	1.15	11.43
Discovery phase 2	Female	13	66.31 +/- 6.25	29.00	4/9	1.08	11.71
	Male	12	66.67 +/- 4.98	28.92	2/10	1.10	11.55
	Total	25	66.48	28.96	6/19	1.09	11.63

Abbreviations: BLSA-NI, Baltimore Longitudinal Study of Ageing Neuroimaging Substudy; DVR, Distribution Volume Ratio; MMSE, Mini Mental State Examination.

* Data presented as mean ± standard deviation

Table 2: Extension study; Demographics of selected subjects from the AIBL cohort

	N	Gender (M/F)	Mean age*	CTL	SMC	MCI	AD	APOE ε4	(carrier/non-carrier)	Mean MMSE
Low neocortical	38	20/18	75.8 +/-	13	18	7	0	14/24		28.3
SUVR (PiB -)			6.53							
High neocortical	40	20/20	80.9 +/-	6	13	15	6	25/15		26.8
SUVR (PiB +)			8.22							
Total	78	40/38	78.35	19	31	22	6	39/39		27.55

Abbreviations: AD, Alzheimer's disease; APOE, Apolipoprotein E; CTL, control subject; M/F, male/female; MCI, mild cognitive impairment; MMSE, Mini Mental State Examination; PiB, Pittsburgh compound B; SUVR, standardized uptake value ratio; SMC, subjective memory complaint.

* Data presented as mean \pm standard deviation

Table 3: Elastic-net penalised regression model components per time point, and resulting linear regression analyses.

Time point	Elastic-net regression model components	Linear Regression		
		Rsquare	F(d.f.)	<i>p</i> -value
1	20 2DGE protein spots, age and APOE ε4	0.828	6.575(22,30)	<0.001
2	5 2DGE protein spots, age, gender, education, BMI, cholesterol, APOE ε4, and prediction years.	0.527	3.524(12,38)	<0.005
3	14 2DGE protein spots, age, APOE ε4, and prediction years.	0.898	18.728 (17,36)	<0.001

Abbreviations: 2DGE, two dimensional gel electrophoresis; APOE, Apolipoprotein E; BMI, body mass index.

Table 4: Combined list of proteins identified as biomarkers of cerebral amyloid from both discovery phase experiments, across all time points

Protein	Time point	Discovery phase: 1 (2DGE), 2 (Gel-free LC-MS/MS), or both.*
Alpha-2-macroglobulin	1, 2, 3	2DGE
Apolipoprotein A-I	1, 2, 3	2DGE
Complement C3	1, 2, 3	2DGE (1, 3) Gel-free LC-MS/MS (2)
Complement C4-B	1, 2, 3	Gel-free LC-MS/MS (1) 2DGE (2, 3)
Haptoglobin	1, 2, 3	2DGE
Ig kappa chain C region	1, 2, 3	Both (1) 2DGE (2, 3)
Serum albumin	1, 2, 3	2DGE
Ceruloplasmin	1, 2	2DGE
Alpha-1-antitrypsin	1, 3	2DGE
Complement component C6	1, 3	2DGE
Fibrinogen beta chain	1, 3	Both (1) 2DGE (3)
Fibrinogen gamma chain	1, 3	2DGE
Ig alpha-1 chain C region	1, 3	2DGE
Ig gamma-1 chain C region	1, 3	2DGE
Ig gamma-2 chain C region	1, 3	2DGE
Ig lambda-2 chain C regions	1, 3	2DGE
Serum amyloid P-component	1, 3	2DGE
Fibrinogen alpha chain	2, 3	Both (2) 2DGE (3)
Protein AMBP	2, 3	2DGE

Apolipoprotein A-IV	1	2DGE
Beta-2-glycoprotein 1	1	2DGE
Complement C1r subcomponent	1	2DGE
Immunoglobulin lambda-like polypeptide 5	1	2DGE
Inter-alpha-trypsin inhibitor heavy chain H4	1	2DGE
Leucine rich alpha 2 glycoprotein 1	1	Gel-free LC-MS/MS
Zinc alpha 2 glycoprotein	1	Gel-free LC-MS/MS
Complement factor H	2	2DGE
Alpha 1 acid glycoprotein 2	3	Gel-free LC-MS/MS
Alpha-amylase 1	3	2DGE
Annexin A1	3	2DGE
Apolipoprotein L1	3	2DGE
Beta 2 microglobulin	3	Gel-free LC-MS/MS
Complement C4-A	3	2DGE
Dermcidin	3	2DGE
Desmoplakin	3	2DGE
Serum paraoxonase/arylesterase 1	3	2DGE

Abbreviations: 2DGE, two dimensional gel electrophoresis; LC-MS/MS, liquid chromatography tandem mass spectrometry.

*Specific time points indicated in brackets

Table 5: Extension study; Protein MW isoforms which passed FDR correction ($q < 0.1$)

		Linear regression			Spearman's Rank		
					Correlation		
Gel10	Protein	Beta	<i>P</i>-value	q value	Rho	<i>P</i>-value	q value
fraction							
3	Alpha-2-macroglobulin				0.290	0.010	0.098
9	Alpha-2-macroglobulin	-0.196	0.037	0.079			
1	Serum Albumin	0.160	0.020	0.078			
10	Apolipoprotein A-I	-0.211	0.013	0.078			
6	Apolipoprotein A-I	-0.157	0.023	0.078			
6	Complement C3	-0.148	0.035	0.079			
8	Haptoglobin	-0.170	0.026	0.078			
5	Haptoglobin	-0.198	0.005	0.075	-0.280	0.013	0.098

Abbreviations: FDR, false discovery rate; MW, molecular weight.

Supplementary data:

Supplementary Table 1: Most abundant proteins identified in baseline 2DGE spots nominally significantly related to PiB PET DVR ($p < 0.05$)

Spot ID	Most abundant proteins identified in 2DGE spot
X1052	Alpha-1-antitrypsin
	Alpha-2-macroglobulin
	Beta-2-glycoprotein 1
	Fibrinogen beta chain
	Ig gamma-1 chain C region
X1128	Fibrinogen gamma chain
	Serum albumin
X1155	Fibrinogen gamma chain
	Serum albumin
X1273	Alpha-1-antitrypsin
	Apolipoprotein A-IV
	Haptoglobin

	Inter-alpha-trypsin inhibitor heavy chain H4
X1287	Alpha-2-macroglobulin
	Haptoglobin
	Serum albumin
X1400	Alpha-1-antitrypsin
	Haptoglobin
	Ig gamma-1 chain C region
	Ig gamma-2 chain C region
	Ig kappa chain C region
	Inter-alpha-trypsin inhibitor heavy chain H4
X1702	Ig kappa chain C region
	Serum albumin
X1730	Ig kappa chain C region
	Ig lambda-2 chain C regions
	Immunoglobulin lambda-like polypeptide 5

	Serum albumin
	Serum amyloid P-component
X1792	Apolipoprotein A-I
X1988	Alpha-2-macroglobulin
	Serum albumin
X2020	Alpha-2-macroglobulin
	Complement C3
	Ig gamma-1 chain C region
	Ig gamma-2 chain C region
	Serum albumin
X2026	Alpha-2-macroglobulin
	Complement C3
	Serum albumin
X2048	Alpha-1-antitrypsin
	Ceruloplasmin
	Inter-alpha-trypsin inhibitor heavy chain H4

X2157	Beta-2-glycoprotein 1
	Fibrinogen beta chain
	Ig gamma-1 chain C region
	Ig gamma-2 chain C region
	Serum albumin
X393	Alpha-2-macroglobulin
	Serum albumin
X529	Alpha-2-macroglobulin
	Ceruloplasmin
	Ig alpha-1 chain C region
	Serum albumin
X632	Alpha-2-macroglobulin
	Complement C3
	Serum albumin
X639	Alpha-2-macroglobulin
	Complement C3
	Complement component C6

	Serum albumin
X740	Alpha-2-macroglobulin
	Ceruloplasmin
	Complement C1r subcomponent
	Ig alpha-1 chain C region
	Serum albumin
X765	Alpha-2-macroglobulin
	Serum albumin

Supplementary Table 2: Most abundant proteins identified in time point 2 2DGE spots nominally significantly related to PiB PET DVR ($p<0.05$)

Spot ID	Most abundant proteins identified in 2DGE spot
X1309	Alpha-2-macroglobulin
	Complement C4-B
	Haptoglobin
	Serum albumin
X1660	Alpha-2-macroglobulin
	Protein AMBP
	Serum albumin
X1763	Alpha-2-macroglobulin
	Apolipoprotein A-I
	Ig kappa chain C region
	Serum albumin
X2271	Alpha-2-macroglobulin
	Ceruloplasmin
	Complement factor H

	Serum albumin
X949	Fibrinogen alpha chain
	Serum albumin

Supplementary Table 3: Most abundant proteins identified in time point 3 2DGE spots nominally significantly related to PiB PET DVR ($p<0.05$)

Spot ID	Most abundant proteins identified in 2DGE spot
X1110	Fibrinogen beta chain
	Fibrinogen gamma chain
	Ig gamma-1 chain C region
	Ig gamma-2 chain C region
	Serum albumin
X1170	Alpha-2-macroglobulin
	Complement C3
	Fibrinogen gamma chain
	Serum albumin
	Serum paraoxonase/arylesterase 1
X1172	Alpha-2-macroglobulin
	Fibrinogen gamma chain
	Serum albumin
X1303	Apolipoprotein L1

	Complement C4-A
	Dermcidin
	Haptoglobin
	Serum albumin
x1309	Alpha-2-macroglobulin
	Complement C4-B
	Haptoglobin
	Serum albumin
X1377	Alpha-1-antitrypsin
X1593	Alpha-1-antitrypsin
	Protein AMBP
X1597	Alpha-1-antitrypsin
	Alpha-amylase 1
	Annexin A1
	Complement C4-A
	Desmoplakin
	Fibrinogen alpha chain

	Ig gamma-1 chain C region
	Ig kappa chain C region
X1710	Alpha-1-antitrypsin
	Alpha-2-macroglobulin
	Complement C4-A
	Fibrinogen beta chain
	Ig gamma-1 chain C region
	Ig kappa chain C region
	Ig kappa chain C region
	Ig lambda-2 chain C regions
	Serum albumin
	Serum amyloid P-component
X1762	Alpha-1-antitrypsin
	Alpha-2-macroglobulin
	Apolipoprotein A-I
	Ig kappa chain C region
	Serum albumin

X381	Alpha-2-macroglobulin
	Complement C3
	Serum albumin
X386	Alpha-2-macroglobulin
	Serum albumin
X584	Alpha-1-antitrypsin
	Alpha-2-macroglobulin
	Ig alpha-1 chain C region
	Serum albumin
X642	Alpha-1-antitrypsin
	Complement C3
	Complement component C6
	Serum albumin

Supplementary Table 4: Baseline partial correlation results for proteins with PiB PET DVR.

Table lists nominally significant proteins ($p < 0.05$). Proteins ranked in order of significance based on p value. Full length protein MW and percentage of data available also listed.

Protein	Correlation coefficient	P -value	MW (kDa)	Complete data (%)
Complement C4 B	-0.498	0.022	193	96
Leucine rich alpha 2 glycoprotein 1	0.607	0.022	39	64
Zinc alpha 2 glycoprotein	0.478	0.025	34	100
Ig kappa chain C region	0.455	0.035	12	100
Fibrinogen beta chain	-0.471	0.038	56	92

Supplementary Table 5: Time point 2 partial correlation results for proteins with PiB PET DVR. Table lists nominally significant proteins (red, $p < 0.05$). Proteins ranked in order of significance based on p value. Full length protein MW and percentage of data available also listed.

Protein	Correlation coefficient	<i>P</i> -value	MW (kDa)	Complete data (%)
Complement C3	-0.567	0.017	187	91
Fibrinogen alpha chain	-0.477	0.042	95	100

Supplementary Table 6: Time point 3 partial correlation results for proteins with PiB PET DVR. Table lists nominally significant proteins (red, $p < 0.05$). Proteins ranked in order of significance based on p value. Full length protein MW and percentage of data available also listed.

Protein	Correlation coefficient	<i>P</i> -value	MW (kDa)	Complete data (%)
Alpha 1 acid glycoprotein 2	0.520	0.043	24	73
Beta 2 microglobulin	0.574	0.047	14	61