Article Title: Alzheimer's disease susceptibility gene apolipoprotein e (*APOE*) and blood biomarkers in UK Biobank (N=395,769).

Authors: Amy C. Ferguson[1] PhD, Rachana Tank[1] MSc, Laura M. Lyall[1] PhD, Joey Ward[1] BSc, Carlos Celis-Morales[2,3] PhD, Rona Strawbridge[1,4,5] PhD, Frederick Ho[1] PhD, Christopher D. Whelan[6] PhD, Jason Gill[2] PhD, Paul Welsh[2] PhD, Jana J. Anderson[1] PhD, Patrick B. Mark[2] PhD, Daniel F. Mackay[1] PhD, Daniel J. Smith[1] MD, Jill P.Pell[1] PhD, Jonathan Cavanagh[1] PhD, Naveed Sattar[1] FRSE, Donald M. Lyall[1] PhD.

[1]Institute of Health & Wellbeing, University of Glasgow, Scotland, UK.
[2]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, Scotland, UK.
[3]GEEAFyS, Universidad Católica del Maule, Talca, Chile
[4]Health Data Research UK.
[5]Cardiovascular Medicine, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.
[6]Research and Early Development (RED), Biogen Inc., Cambridge MA, 02139, USA.

Corresponding author:
Dr. Donald M. Lyall (Donald.Lyall@Glasgow.ac.uk)
1 Lilybank Gardens,
Institute of Health and Wellbeing,
University of Glasgow,
G12 8RZ,
Glasgow, UK.
Tel: +44 (0) 141 330 8547
Fax: +44 (0) 141 330 1874

Title character count: 98
Text word count: 3627
Abstract word count: 217
Tables: 1
Figures: 1
Supplementary tables: 4
Supplementary figures: 0
References: 51

Search terms: Alzheimer disease; dementia; UK Biobank; cholesterol; APOE.

Abstract

Background

Alzheimer's disease (AD) is a neurodegenerative condition where the underlying aetiology is still unclear. Investigating the potential influence of apolipoprotein e (*APOE*), a major genetic risk factor, on common blood biomarkers could provide a greater understanding of the mechanisms of AD and dementia risk.

Objective

Our objective was to conduct the largest (to date) single-protocol investigation of blood biomarkers in the context of *APOE* genotype, in UK Biobank.

Methods

After quality control and exclusions, data on 395,769 participants of White European ancestry were available for analysis. Linear regressions were used to test potential associations between *APOE* genotypes and biomarkers.

Results

Several biomarkers significantly associated with *APOE* e4 'risk' and e2 'protective' genotypes (vs. neutral e3/e3). Most associations supported previous data: for example, e4 genotype was associated with elevated low-density lipoprotein cholesterol (LDL) (standardized beta [b] = 0.150 standard deviations [SDs] per allele, $p<0.001$) and e2 with lower LDL (b = -0.456 SDs, $p<0.001$). There were however instances of associations found in unexpected directions: e.g. e4 and increased insulin-like growth factor (IGF-1) (standardized beta = 0.017, $p<0.001$) where lower levels have been previously suggested as an AD risk factor.

Conclusion

These findings highlight biomarker differences in non-demented people at genetic risk for dementia. The evidence here in supports previous hypotheses of involvement from cardiometabolic and neuroinflammatory pathways.

Introduction

Alzheimer disease (AD) is the most common form of dementia and an important public health issue [1], hypothesized to be the result of interactions between genetic and environmental risk factors [2]. *APOE* e4 genotype is the largest common single genetic risk factor for AD and cognitive decline behind increasing age [3], with the e2 allele potentially protective [1]. The exact mechanisms by which *APOE* genotype influences brain ageing are unclear but probably due to pleiotropic pathways stemming from its core role in lipid metabolism [4].

Several studies have investigated serum biomarker differences between AD patients vs. healthy individuals in order to identify potential risk factors, including low density lipoproteins (LDL) and insulin-like growth factor 1 (IGF-1) with sometimes conflicting results potentially due to methodological heterogeneity [5,6]. Many studies investigating serum levels in AD have focussed on specific biomarkers involved in β-amyloid precursor protein metabolism and phosphorylation [7]. There have been relatively few studies investigating a wide range of biomarkers in a "hypothesis-free" approach; and those studies which have done this appear to be limited by small sample size or have been cross-sectional in individuals with an extant diagnosis of dementia [8–10]. Relatively few biomarkers have been investigated in the context of AD-susceptibility gene *APOE*. Gaining a greater understanding of how *APOE* influences biomarker serum levels could be extremely beneficial: highlighting factors significantly associated with AD genetic risk could elucidate potential pathways involved in its development, pathophysiology and ultimately treatment [11].

In this study *APOE* genotype status was tested vs. a range of circulating serum blood biomarkers available for approximately 396,000 participants in UK Biobank. Two separate analyses were undertaken to investigate the influence of genotypic status on biomarker levels: differences per i) risk e4, or ii) protective e2 allele; each vs. neutral e3/e3 genotype. Further analyses were undertaken to investigate the associations in males and females separately due to *a priori* evidence for *APOE*-sex differences in AD pathophysiology [12]. To our knowledge, this is the first large-scale investigation of the relationship between *APOE* genotype status and a wide range of biomarkers in a population cohort.

Methods

Subjects

Over 502,000 UK residents aged 37-73 years were recruited to UK Biobank from 2006-2010. At one of 22 assessment centres across the UK, participants completed a range of phenotypic assessments and questionnaires, including genetic, urine and blood samples [13]. We focussed on participants with White British ancestry because there is evidence of different e4 frequencies across ethnicities[14].

Ethical approval

This secondary-data analysis study was conducted under generic approval from the NHS National Research Ethics Service (approval letter dated 17th June 2011, ref 11/NW/0382). Written informed consent was obtained from all participants in the study (consent for research, by UK Biobank).

Genotyping

UK Biobank participants were genotyped using Applied Biosystems UK BiLEVE Axiom array by Affymetrix and Applied Biosystems UK Biobank Axiom Array which share 95% marker content [13]. *APOE* e status was based on two single nucleotide polymorphisms (SNPs): rs7412 and rs429358. Stringent quality control and processing were applied to the data, detailed at http://www.ukbiobank.ac.uk/scientists-3/genetic-data and http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UKBiobank_genotyping_QC_documentation-web.pdf.

Biomarker collection and processing

Biomarker levels were analysed in UK Biobank from serum and packed red blood cell samples obtained from all UK Biobank participants at baseline [15]. Stringent quality controls were applied to the assays used measure biomarker levels, details of biomarker quality control, instrumentation and analysis methods are available at: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/biomarker_issues.pdf, https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_biochemistry.pdf, http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/haematology.pdf, and http://www.ukbiobank.ac.uk/wp-content/uploads/2018/11/BCM023_ukb_biomarker_panel_website_v1.0-Aug-2015-edit-2018.pdf. Processing of very low levels of oestradiol and rheumatoid factor (RF) were recorded as "missing" (in the original data); these missing values were recoded conservatively as the square root of the minimum stated detectable value if individuals had data for a remaining biomarker [16] and were not coded as 'no data returned' or having unrecoverable aliquot problems. Oestradiol and RF require attention in this context because they were the biomarkers highlighted by UK Biobank as the variables with by far the highest frequency of values below reportable levels (75-90% of results); no assays had >0.1% of results above their reportable range.

Dementia outcomes

We validated *APOE* genotype's association with dementia/AD outcomes (in UK Biobank) as a check, having previously shown associations in expected directions with brain structure [3] and

cognitive abilities [17]. Dementia and AD outcomes were generated using self-report, hospital admission and death record data, with hospital and death record data utilising International Classification of Diseases version 10 (ICD-10 codes). Individuals were designated as cases ("all-cause dementia" or "Alzheimer disease") if they had indicated dementia or AD in either self-report, or through hospital or death records – derived by UK Biobank based on self-report, hospital admission and death reports [18]. Those coded as missing were designated as controls (i.e. did not self-report dementia/AD, and these diagnoses were not present in hospital/death records).

Covariates

Participants self-reported their smoking history, and we collated past and current smokers into 'ever' (vs. never). Participants self-reported medication use for cholesterol, high blood pressure or insulin. We excluded those who did not know or preferred not to answer for these various items (<5%). Townsend deprivation indices were derived from postcode of residence [19]. This provides an area-based measure of socioeconomic deprivation derived from aggregated data on car ownership, household overcrowding, owner occupation and unemployment. Higher Townsend scores equate to higher levels of area-based socioeconomic deprivation. We additionally controlled for potential population stratification using UK Biobank-derived principal components (PCs) 1-5, and genotypic array[13]. Height was measured (Seca 202 stadiometer) and weight was measured to the nearest 0.1 kg (BC-418 MA body composition analyser; Tanita Corp). Body mass index (BMI) was derived from weight in kilograms divided by height in meters squared. All-cause cancer was derived based on self-report at baseline. Month of assessment was recorded by UK Biobank. We have previously described and derived an 'any self-reported inflammatory condition' variable[20], including 58 conditions described in an open-access report. Participant assessment centre was recorded by UK Biobank, and educational attainment was self-reported.

Statistical analyses

The linear regressions reported reflect average SD-changes per e4 or e2 allele vs. neutral e3/e3 genotype, i.e. e3/e3 vs. e3/e4 vs. e4/e4 (dose), and e2/e2 vs. e2/e3 vs. e3/e3 (dose). Associations with *APOE* genotype vs. dementia/AD were tested using binary logistic regressions reporting odds ratios (OR) and their 95% confidence intervals. We corrected for multiple testing using False Discovery Rate (FDR)[21], conservatively collating all tests. Biomarkers which were not normally distributed were log transformed prior to Z-score transformation and reanalysed: the resulting effect sizes were unchanged and therefore we report the original estimates.

Three linear regression models were used to investigate potential associations with each biomarker and adjusted for potential confounders. Model 1 ('minimally-adjusted') adjusted for age, sex, baseline assessment centre, principal components 1-5 for population stratification, and genotyping array. Model 2 ('partially-adjusted') also included self-reported diabetes, high blood pressure and coronary heart disease (comprised of angina plus myocardial infarction [22]). Model 3 ('fully adjusted') also controlled for self-reported cholesterol, hormone replacement therapy, oral contraceptive, insulin or hypertension medication, Townsend deprivation scores, and ever vs. never smoking. The cross-sectional association between *APOE* vs. dementia outcomes were also analysed using these models.

As additional sensitivity analyses we adjusted for dummy-variable 'any self-reported chronic inflammatory condition' (n=64,996; 17%); underweight (BMI<18.5; n=1,962 or 0.5%) or obesity (BMI≥30, n = 297,738 or 75.4%) vs. normal to overweight (18.5 to 30 BMI; n=95,001 or 24.1%), month of assessment, university/college degree vs. not, and finally we additionally corrected any significant associations with IGF-1 for self-reported baseline cancer history (n=33,406; 8%) [23]. SNP data was collated and quality controlled with PLINK V1.90, and analysed with Stata V16.

Results

We excluded participants with non-white British ancestry (N=78,672; 16%), sex mismatch (self-report versus genetic), chromosomal aneuploidy, excessive heterozygosity and genotype missing rate >10%. We excluded the minority of participants with e2/e4 (n=2,556; 0.7%) genotype because this included potentially protective and risk alleles [24]. We removed outliers >5SDs from the mean for each biomarker. This left overall N=395,769 which varied slightly by biomarker: Table 1 shows sample size and key values per biomarker.

[Table 1 here]

<u>*APOE* associations with dementia phenotypes</u>

As a form of replication and to support the utility of investigating *APOE* genotype in the UK Biobank cohort, we tested for e4 and e2 allele count (vs. e3/e3) against all-cause dementia (n=1,852; 0.5%), and specific AD diagnosis (n=722; 0.2%). We found that e4 was associated with increased AD (fully-adjusted OR = 3.51 per allele, 95% CI = 3.14 to 3.92, P<0.001) and all-cause dementia (OR = 2.59, 95% CI = 2.40 to 2.79, P<0.001), whereas, e2 was correspondingly associated with decreased AD (OR = 0.59, 95% CI = 0.40 to 0.85, P = 0.005) and all-cause dementia (OR = 0.78, 95% CI = 0.65 to 0.93, P = 0.007). An unadjusted chi-square test showed 64%% of people with AD had at least one e4 allele vs. 36% in the non-AD group.

APOE associations with biomarker values

Several significant associations (at nominal p<0.05) were identified between *APOE* genotype status and biomarker values, as shown in Figure 1, which shows fully-adjusted estimates. Increasing e4 allele count associated with significant differences in several biomarker values (Supplementary Table S1). There were associations between e4 allele count vs. higher LDL, IGF-1, sex hormone binding globulin (SHBG), total bilirubin, triphosphate levels, ApoB and total cholesterol. Negative associations were found between e4 genotype and lower high-density lipoprotein (HDL), haemoglobin A1c (HbA1c), lipoprotein A, phosphate, C-reactive protein (CRP), gamma glutamyl transferase (GGT), vitamin D, creatinine, urate, and urea. The largest effect sizes – based on fully-adjusted model results- were seen for total cholesterol (0.13 SDs per allele in the fully adjusted model), ApoB (0.20), CRP (-0.12) and LDL (0.15), with the rest <0.1SDs per e4 allele. There was no evidence of a significant fully adjusted association between e4 and oestradiol, aspartate transaminase (AST), albumin, testosterone and RF levels.

[Figure 1 here]

Significant associations were seen between e2 allele count vs. lower LDL, HDL, IGF-1, total bilirubin, direct bilirubin, vitamin D, CRP, cystatin C (CysC), ApoA, ApoB, creatinine and alkaline phosphatase (Figure 1). These are in the opposite directions of effect reported for the associations to the e4 allele (vs. e3/e3) as expected. Associations between *APOE* e2 and HbA1c, lipoprotein A, SHBG and triphosphate levels were also identified, however, these effects were in the same direction as e4. The largest effect sizes were for LDL (-0.46 SDs per allele in the fully-adjusted model), triphosphate (0.13), ApoB (-0.61) and total cholesterol (-0.35), with the rest <0.1SDs per e2 allele.

There were instances of significant association for e2 but not e4 vs. e3/e3. These were: negative associations between e2 vs. protein and aspartate aminotransferase levels, and positive association between e2 and testosterone. There was no statistically significant association between e2 count and oestradiol, phosphate, GGT, urate, urea and alanine aminotransferase levels (Supplementary Table S1). Both e4 and e2 were also found to associate with increased SHBG, decreased protein and increased triphosphate (vs. e3/e3).

Sex-specific analyses

There were instances of male/female sex vs. genotype interactions for several biomarkers (Supplementary Table S2). Out of 33 biomarkers, 16 showed some interaction (P<0.05): LDL, HDL, HbA1c, Oestradiol, RF, SHBG, Testosterone, protein, triphosphate, urate, ApoB, ApoA, total

cholesterol, AST and alanine transaminase (ALT). We provide complete sex-specific Z-score associations in Supplementary Table S3. Most of the significant associations in the collated analyses remained so in individual sexes. Certain associations were only significant in males: e4 vs. oestradiol and calcium, and e2 vs. IGF-1, SHBG, testosterone and cystatin c.

Sensitivity analyses

All nominally significant associations survived correction for FDR. Results were unchanged in terms of effect size and P-value when individuals with incident dementia/AD were removed from analyses. When we added presence of any self-reported inflammatory condition, month of assessment, degree vs. not, underweight or obesity (vs. normal to overweight) to the final model, no results were meaningfully changed (Supplementary Table S4). Results were unchanged when we re-analysed all outcome variables with inverse-rank normalisation to avoid potential false positives caused by outlying values (prior to removing values >5SD from the mean), and when we used the maximum 40 PCs (vs. 5 reported). IGF-1 results were unchanged when we additionally controlled for all-cause cancer. As a check, we re-ran all final-model tests as 0 vs. 1 allele, and 0 vs. 2 allele contrasts (rather than a 0/1/2 dose effect); results were consistently indicative of dose effects in the same direction.

Discussion

The *APOE* e4 allele is known to associate in UK Biobank with worse non-demented cognitive abilities [17], cerebrovascular health [3] and here clinically-ascertained AD/dementia risk (mostly; a small minority of cases were baseline self-report). In this study we found several significant associations per e4 allele on circulating biomarker levels (vs. neutral e3/e3). In many instances these were supported by corresponding associations between the putatively protective e2 genotype in the opposite direction as would be expected. Gaining a greater understanding of the biomarker profiles of individuals at genetic risk of AD may be useful in the future for earlier detection of AD and could potentially highlight pathways as therapeutic targets [11]. In some instances, the directions of effect for e4 ('risk') and e2 ('protective') alleles conflicted with what would be expected based on levels reported elsewhere in people with prevalent AD. For example: lower levels of IGF-1 have previously associated with increased risk of AD and cognitive decline [6,25]. By contrast here we showed association between e4 and increased IGF-1, and between e2 genotype and lower IGF-1. This is surprising as IGF-1 stimulates neurogenesis and promotes cell survival [26]. Previous reports could reflect some degree of reverse causality or bias in cross-sectional AD patient sample studies showing lower IGF-1 levels i.e. where disease onset affects biomarker health.

Interpretation

Vitamin D

Previous studies have reported *APOE* e4 association with higher vitamin D serum concentration [27] however here e4 associated with decreased, and e2 with increased vitamin D. In e4 homozygotes, a higher vitamin D concentration has been associated with higher memory function, suggesting higher vitamin D levels could be protective for people at risk for AD [28].

Lipids

Investigations into the effect of differing cholesterol levels on cognitive decline and AD have produced conflicting results with both low and high cholesterol being associated [8,29–31]. The lack of consensus could be partially due to the smaller sample sizes previously used. The three *APOE* alleles encode for different ApoE protein isoforms with altered lipid interactions in serum; the e3 encoded protein isoform is associated with "normal" plasma lipid levels [32]. *APOE* e4 and its associated higher LDL have been previously associated with early onset of AD [33] while e2 has protective effects on the concentrations of cholesterol, lipids and phospholipids [34]. Here we reinforce those observations, particularly in the context of respective dose e2/e4 protective vs. deleterious associations [5,8,34,35].

It has been hypothesised that lipid metabolism is important in the pathophysiology of AD [11]. The significant associations between *APOE* genotype status and ApoA/ApoB support this. ApoA and ApoB proteins are major surface proteins of HDL and LDL, respectively [36]; previously ApoA has been associated with lower risk of cardiovascular disease [37], whereas, ApoB is reported to be proatherogenic [36]. We identified lower ApoA levels in e4 carriers and higher levels in e2 carriers; ApoA has reportedly neuroprotective effects by inhibiting β-amyloid plaque aggregation [36,38]. Our findings are supported by previous associations reported between decreased serum ApoA and increased AD risk [38]. With ApoB, we found e4 was associated with elevated levels of ApoB and e2 with lower levels of ApoB compared to e3. This is consistent with the direction of effect reported in most [10,33] but not all studies [36].

It is hypothesised *APOE* allele-encoded protein isoforms have different affinities for lipoproteins [39]. It has been difficult to define the exact effects of *APOE* genotype on lipoprotein A from previous studies' many small sample sizes. Our data showed that both *APOE* e2 and e4 were significantly associated with decreased levels of lipoprotein A (vs. e3)[39]. Elevated levels of lipoprotein A are known to have a causal relationship with myocardial infarction, which has previously been associated with increased dementia risk [39]. The influence of lower lipoprotein A on the pathophysiology of AD remains unclear.

APOE genotype status may modify the effect of sex hormones on dementia symptoms [40]. We did not find any evidence for a significant association between *APOE* genotype and oestradiol in the whole sample, although in males e4 associated with lower levels. Oestradiol

has suggested overall neuroprotective effects; however, evidence is conflicting, and this relationship is not fully understood [40]. We found an association between *APOE* e2 and higher testosterone levels, specific to males. There is conflicting evidence regarding the effect of testosterone on AD risk [41,42]; some cross-sectional studies report lower levels of testosterone in AD patients vs. healthy controls [42],consistent with our findings as e2 is associated with decreased AD risk [1].

Inflammation

The underlying pathophysiology of AD has been suggested to be at least partially influenced by neuroinflammation [43]. Serum CRP levels are a marker for inflammation but the evidence for association with AD risk is conflicting. Interactions between *APOE* e4 and elevated CRP have been reported to associate with early onset of AD [44]. However, consistent with our findings e4 has been associated with lower CRP levels, and e2 with higher CRP levels [10,45]. Further research is required to elucidate the effects of CRP levels on the pathophysiology of AD, particularly given it is a marker of acute rather than chronic inflammation [43]. Another inflammatory marker associated with increased risk of AD is raised GGT [46]. We identified lower levels of GGT in e4 carriers: this may reflect bias in cross-sectional studies of GGT and AD.

Alkaline phosphatase may have some involvement in the inflammatory/AD process [47]. We identified lower levels of alkaline phosphatase in e4 carriers and elevated levels in e2 carriers. It has been suggested that alkaline phosphatase could potentially be used as a therapy to reduce neuroinflammation in AD [47]; our findings may support this but more in depth investigations are required.

CysC, typically a marker of kidney dysfunction, is also involved in modulation of inflammatory responses and reported to have neuroprotective effects in AD as it co-localizes with β-amyloid and inhibits oligomerization [48]. We found that e4 associated with lower CysC and e2 with higher CysC levels compared to e3. Lower baseline CysC has been reported to precede AD onset in an 11-year longitudinal study of non-demented elderly men at baseline; in one small study (N=82) [49] which suggested the finding may be due to attrition bias because higher CysC is a risk factor for cardiovascular disease and earlier mortality. Our findings potentially lend support to low CysC serum levels as an AD risk factor.

Other biomarkers

We report suggestive association between e4 and lower phosphate levels, potentially contradicting prior research showing association between (age-dependent) higher serum phosphorous and incident dementia [50]. We found significant associations between e4 and higher total bilirubin and urea, and lower direct bilirubin, urate, creatinine, calcium, and alanine

aminotransferase. In the whole sample analysis, we found evidence of associations between e2 and higher direct bilirubin, creatinine and aspartate aminotransferase, and lower total bilirubin and albumin. Both e4 and e2 associated in same direction with HbA1c, SHBG, protein and triphosphate. This is unexpected: e4 and e2 tend to show opposing effects regarding AD. We found proportionally moderate evidence of interaction between sex, genotype and biomarker; this is a significant area of research and warrants further study along with investigation age- and multimorbidity-related interactions [12].

Limitations and future research

A limitation of UK Biobank is that participants are overall likely to have fewer health conditions, be better educated, of older age, female, and living in less socio-economically deprived areas than the general population [13]. This study was conducted in individuals of White European ancestry only and so these results may not be generalizable to a more mixed population. Findings may not be truly representative of the effects of *APOE* genotype in the wider UK population [13]. There may be conflicting biases at play in at least some of the results. For example, confounding effects may exist where e4 carriers are of poorer health, affecting their lifestyle and in turn influencing biomarker values, or by contrast selection bias, where the e4 carriers here are of particularly good health relative to the general population.

Although this study reports significant associations between serum biomarkers and *APOE* genotype, which could influence the risk of AD, these alleles are unlikely to be entirely responsible [1]. Associations may to some extent reflect early prodromal AD; future study may investigate biomarkers in AD-by-proxy (i.e. family history) cases. Some biomarker levels may only be pathogenic in combination with other biomarkers [8]. Some of the biomarkers may be influenced by environmental factors such as seasonality, although this should to some extent average out; sensitivity analyses showed no evidence of confounding. The effect sizes reported here are in many cases small and the results require replication; the clinical applicability of these findings remains unclear. It is not necessarily possible to identify the exact biological pathways involved in the pathophysiology of AD from these analyses. Further work is required to investigate the underlying pathways to identify processes which could be modified or targeted to decrease the risk of AD, including longitudinal biomarker data [12,35]. Future research e.g. using Mendelian randomization may investigate whether pharmaceutically altering serum biomarker levels or implementing lifestyle changes to manage these biomarkers may be beneficial to individuals at greater risk of developing AD.

Conclusion

The exact influence of *APOE* on AD and dementia pathophysiology is unclear. Through this study we have identified associations between the high-risk AD gene locus *APOE* and a range of serum blood biomarkers in UK Biobank. These associations highlight potential pathways involved in the development of cognitive impairment and have potential to lead to earlier detection of AD risk through the analysis of biomarkers and *APOE* genotype.

Funding: UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government and the Northwest Regional Development Agency. It has also had funding from the Welsh Assembly Government and the British Heart Foundation. The funders had no role in study design, data collection or management, analyses or interpretation of the data, nor preparation, review or approval of the manuscript. DML is supported by *The Neurosciences Foundation*, and *American Psychological Foundation*. Smith is partially funded by the Lister institute. Cavanagh is funded by the Sackler Trust, Wellcome Trust, Medical Research Council, and holds a Wellcome Trust strategic award, an industrial-academic collaboration with Janssen & Jannsen, GlaxoSmithKline, and Lundbeck. Pell has received funding from the Medical Research Council and Chief Scientist Office

Declaration of interest: Sattar has consulted for Amgen, Inc., Sanofi, AstraZeneca, Eli Lilly, and has sat on the Medical UK Biobank Scientific Advisory Board. Pell has sat on the Medical Research Council Strategy Board and UK Biobank Scientific Advisory Board. None of these disclosures are directly related to the study, nor its conception, analysis or interpretation.

Acknowledgements: This research has been conducted using the UK Biobank resource; we are grateful to UK Biobank participants. This project was completed using UK Biobank application 17689 (PI: DML).

Author contributions

Study concept and design: DML.

Statistical analysis: ACF; DML.

Drafted the manuscript: ACF, DML.

Critically revised content: all authors.

Obtained principal study funding: DML.

References

[1] Van Cauwenberghe C, Van Broeckhoven C, Sleegers K (2016) The genetic landscape of Alzheimer disease: clinical implications and perspectives. *Genet. Med.* **18**, 421–30.

[2] Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL (2006) Role of Genes and Environments for Explaining Alzheimer Disease. *Arch. Gen. Psychiatry* **63**, 168.

[3] Lyall DM, Cox SR, Lyall LM, Celis-Morales C, Cullen B, Mackay DF, Ward J, Strawbridge RJ, McIntosh AM, Sattar N, Smith DJ, Cavanagh J, Deary IJ, Pell JP Association between APOE e4 and white matter hyperintensity volume, but not total brain volume or white matter integrity. *Brain Imaging Behav.* 1–9.

[4] Holtzman DM, Herz J, Bu G (2012) Apolipoprotein E and apolipoprotein E receptors: normal biology and roles in Alzheimer disease. *Cold Spring Harb. Perspect. Med.* **2**, a006312.

[5] Moriarty PM, Varvel SA, Gordts PLSM, McConnell JP, Tsimikas S (2017) Lipoprotein(a) Mass Levels Increase Significantly According to APOE Genotype: An Analysis of 431 239 Patients. *Arterioscler. Thromb. Vasc. Biol.* **37**, 580–588.

[6] Kimoto A, Kasanuki K, Kumagai R, Shibata N, Ichimiya Y, Arai H (2016) Serum insulin-like growth factor-I and amyloid beta protein in Alzheimer's disease: relationship with cognitive function. *Psychogeriatrics* **16**, 247–254.

[7] Olsson B, Lautner R, Andreasson U, Öhrfelt A, Portelius E, Bjerke M, Hölttä M, Rosén C, Olsson C, Strobel G, Wu E, Dakin K, Petzold M, Blennow K, Zetterberg H (2016) CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol.* **15**, 673–684.

[8] Szalkai B, Grolmusz VK, Grolmusz VI (2017) Identifying combinatorial biomarkers by association rule mining in the CAMD Alzheimer's database. *Arch. Gerontol. Geriatr.* **73**, 300–307.

[9] Tynkkynen J, Chouraki V, van der Lee SJ, Hernesniemi J, Yang Q, Li S, Beiser A, Larson MG, Sääksjärvi K, Shipley MJ, Singh-Manoux A, Gerszten RE, Wang TJ, Havulinna AS, Würtz P, Fischer K, Demirkan A, Ikram MA, Amin N, Lehtimäki T, Kähönen M, Perola M, Metspalu A, Kangas AJ, Soininen P, Ala-Korpela M, Vasan RS, Kivimäki M, van Duijn CM, Seshadri S, Salomaa V (2018) Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: A prospective study in eight cohorts. *Alzheimers. Dement.* **14**, 723–733.

[10] Soares HD, Potter WZ, Pickering E, Kuhn M, Immermann FW, Shera DM, Ferm M, Dean RA, Simon AJ, Swenson F, Siuciak JA, Kaplow J, Thambisetty M, Zagouras P, Koroshetz WJ, Wan HI, Trojanowski JQ, Shaw LM, Biomarkers Consortium Alzheimer's Disease Plasma

Proteomics Project (2012) Plasma biomarkers associated with the apolipoprotein E genotype and Alzheimer disease. *Arch. Neurol.* **69**, 1310–7.

[11] Yamazaki Y, Painter MM, Bu G, Kanekiyo T (2016) Apolipoprotein E as a Therapeutic Target in Alzheimer's Disease: A Review of Basic Research and Clinical Evidence. *CNS Drugs* **30**, 773–89.

[12] Riedel BC, Thompson PM, Brinton RD (2016) Age, APOE and sex: Triad of risk of Alzheimer's disease. *J. Steroid Biochem. Mol. Biol.* **160**, 134–47.

[13] Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J (2018) The UK Biobank resource with deep phenotyping and genomic data. *Nature*.

[14] Eisenberg DTA, Kuzawa CW, Hayes MG (2010) Worldwide allele frequencies of the human apolipoprotein E gene: climate, local adaptations, and evolutionary history. *Am. J. Phys. Anthropol.* **143**, 100–11.

[15] Welsh C, Celis-Morales CA, Brown R, Mackay DF, Lewsey J, Mark PB, Gray SR, Ferguson LD, Anderson JJ, Lyall DM, Cleland JG, Jhund PS, Gill JMR, Pell JP, Sattar N, Welsh P (2019) Comparison of Conventional Lipoprotein Tests and Apolipoproteins in the Prediction of Cardiovascular Disease. *Circulation*.

[16] Sinnott-Armstrong N, Tanigawa Y, Amar D, Mars NJ, Aguirre M, Venkataraman GR, Wainberg M, Ollila HM, Pirruccello JP, Qian J, Shcherbina A, FinnGen, Rodriguez F, Assimes TL, Agarwala V, Tibshirani R, Hastie T, Ripatti S, Pritchard JK, Daly MJ, Rivas MA (2019) Genetics of 38 blood and urine biomarkers in the UK Biobank. *bioRxiv* 660506.

[17] Lyall DM, Celis-Morales C, Lyall LM, Graham C, Graham N, Mackay DF, Strawbridge RJ, Ward J, Gill JMR, Sattar N, Cavanagh J, Smith DJ, Pell JP (2019) Assessing for interaction between *APOE* ε4, sex, and lifestyle on cognitive abilities. *Neurology*.

[18] Wilkinson T, Schnier C, Bush K, Rannikmäe K, Henshall DE, Lerpiniere C, Allen NE, Flaig R, Russ TC, Bathgate D, Pal S, O'Brien JT, Sudlow CLM, Dementias Platform UK and UK Biobank on behalf of DPU and U (2019) Identifying dementia outcomes in UK Biobank: a validation study of primary care, hospital admissions and mortality data. *Eur. J. Epidemiol.* **34**, 557–565.

[19] Townsend P (1998) Townsend deprivation index. *Natl. database Prim. care groups Trust.*

[20] Lyall LM, Cullen B, Lyall DM, Leighton SP, Siebert S, Smith DJ, Cavanagh J (2019) The

associations between self-reported depression, self-reported chronic inflammatory conditions and cognitive abilities in UK Biobank. *Eur. Psychiatry* **60**, 63–70.

[21] Benjamini Y, Hochberg Y (1995) Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300.

[22] Lyall DM, Celis-Morales CA, Anderson J, Gill JMR, Mackay DF, McIntosh AM, Smith DJ, Deary IJ, Sattar N, Pell JP (2017) Associations between single andmultiple cardiometabolic diseases and cognitive abilities in 474 129 UK Biobank participants. *Eur. Heart J.* **38**, 577–583.

[23] Weroha SJ, Haluska P (2012) The Insulin-Like Growth Factor System in Cancer. *Endocrinol. Metab. Clin. North Am.* **41**, 335–350.

[24] Wisdom NM, Callahan JL, Hawkins KA (2011) The effects of apolipoprotein E on non-impaired cognitive functioning: a meta-analysis. *Neurobiol. Aging* **32**, 63–74.

[25] Vidal J-S, Hanon O, Funalot B, Brunel N, Viollet C, Rigaud A-S, Seux M-L, le-Bouc Y, Epelbaum J, Duron E (2016) Low Serum Insulin-Like Growth Factor-I Predicts Cognitive Decline in Alzheimer's Disease. *J. Alzheimer's Dis.* **52**, 641–649.

[26] Lee E, Son H (2009) Adult hippocampal neurogenesis and related neurotrophic factors. *BMB Rep.* **42**, 239–44.

[27] Egert S, Rimbach G, Huebbe P (2012) ApoE genotype: from geographic distribution to function and responsiveness to dietary factors. *Proc. Nutr. Soc.* **71**, 410–424.

[28] Maddock J, Cavadino A, Power C, Hyppönen E (2015) 25-Hydroxyvitamin D, APOE ε4 genotype and cognitive function: findings from the 1958 British birth cohort. *Eur. J. Clin. Nutr.* **69**, 505–508.

[29] Wagner M, Helmer C, Tzourio C, Berr C, Proust-Lima C, Samieri C (2018) Evaluation of the Concurrent Trajectories of Cardiometabolic Risk Factors in the 14 Years Before Dementia. *JAMA Psychiatry* **75**, 1033.

[30] Li G, Shofer JB, Kukull WA, Peskind ER, Tsuang DW, Breitner JCS, McCormick W, Bowen JD, Teri L, Schellenberg GD, Larson EB (2005) Serum cholesterol and risk of Alzheimer disease. *Neurology* **65**, 1045–1050.

[31] Reitz C, Tang M-X, Schupf N, Manly JJ, Mayeux R, Luchsinger JA (2010) Association of higher levels of high-density lipoprotein cholesterol in elderly individuals and lower risk of late-onset Alzheimer disease. *Arch. Neurol.* **67**, 1491–7.

[32] Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Barker-Collo S, Bartels DH, Bell ML, Benjamin EJ, Bennett D, Bhalla K, Bikbov B, Bin Abdulhak A, Birbeck G, Blyth F, Bolliger I, Boufous S, Bucello C, Burch M, Burney P, Carapetis J, Chen H, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahodwala N, De Leo D, Degenhardt L, Delossantos A, Denenberg J, Des Jarlais DC, Dharmaratne SD, Dorsey ER, Driscoll T, Duber H, Ebel B, Erwin PJ, Espindola P, Ezzati M, Feigin V, Flaxman AD, Forouzanfar MH, Fowkes FGR, Franklin R, Fransen M, Freeman MK, Gabriel SE, Gakidou E, Gaspari F, Gillum RF, Gonzalez-Medina D, Halasa YA, Haring D, Harrison JE, Havmoeller R, Hay RJ, Hoen B, Hotez PJ, Hoy D, Jacobsen KH, James SL, Jasrasaria R, Jayaraman S, Johns N, Karthikeyan G, Kassebaum N, Keren A, Khoo J-P, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lipnick M, Lipshultz SE, Ohno SL, Mabweijano J, MacIntyre MF, Mallinger L, March L, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGrath J, Mensah GA, Merriman TR, Michaud C, Miller M, Miller TR, Mock C, Mocumbi AO, Mokdad AA, Moran A, Mulholland K, Nair MN, Naldi L, Narayan KMV, Nasseri K, Norman P, O'Donnell M, Omer SB, Ortblad K, Osborne R, Ozgediz D, Pahari B, Pandian JD, Rivero AP, Padilla RP, Perez-Ruiz F, Perico N, Phillips D, Pierce K, Pope CA, Porrini E, Pourmalek F, Raju M, Ranganathan D, Rehm JT, Rein DB, Remuzzi G, Rivara FP, Roberts T, De León FR, Rosenfeld LC, Rushton L, Sacco RL, Salomon JA, Sampson U, Sanman E, Schwebel DC, Segui-Gomez M, Shepard DS, Singh D, Singleton J, Sliwa K, Smith E, Steer A, Taylor JA, Thomas B, Tleyjeh IM, Towbin JA, Truelsen T, Undurraga EA, Venketasubramanian N, Vijayakumar L, Vos T, Wagner GR, Wang M, Wang W, Watt K, Weinstock MA, Weintraub R, Wilkinson JD, Woolf AD, Wulf S, Yeh P-H, Yip P, Zabetian A, Zheng Z-J, Lopez AD, Murray CJL, AlMazroa MA, Memish ZA, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Barker-Collo S, Bartels DH, Bell ML, Benjamin EJ, Bennett D, Bhalla K, Bikbov B, Abdulhak A Bin, Birbeck G, Blyth F, Bolliger I, Boufous S, Bucello C, Burch M, Burney P, Carapetis J, Chen H, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahodwala N, De Leo D, Degenhardt L, Delossantos A, Denenberg J, Des Jarlais DC, Dharmaratne SD, Dorsey ER, Driscoll T, Duber H, Ebel B, Erwin PJ, Espindola P, Ezzati M, Feigin V, Flaxman AD, Forouzanfar MH, Fowkes FGR, Franklin R, Fransen M, Freeman MK, Gabriel SE, Gakidou E, Gaspari F, Gillum RF, Gonzalez-Medina D, Halasa YA, Haring D, Harrison JE, Havmoeller R, Hay RJ, Hoen B, Hotez PJ, Hoy D, Jacobsen

KH, James SL, Jasrasaria R, Jayaraman S, Johns N, Karthikeyan G, Kassebaum N, Keren A, Khoo J-P, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lipnick M, Lipshultz SE, Ohno SL, Mabweijano J, MacIntyre MF, Mallinger L, March L, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGrath J, Memish ZA, Mensah GA, Merriman TR, Michaud C, Miller M, Miller TR, Mock C, Mocumbi AO, Mokdad AA, Moran A, Mulholland K, Nair MN, Naldi L, Narayan KMV, Nasseri K, Norman P, O'Donnell M, Omer SB, Ortblad K, Osborne R, Ozgediz D, Pahari B, Pandian JD, Rivero AP, Padilla RP, Perez-Ruiz F, Perico N, Phillips D, Pierce K, Pope CA, Porrini E, Pourmalek F, Raju M, Ranganathan D, Rehm JT, Rein DB, Remuzzi G, Rivara FP, Roberts T, De León FR, Rosenfeld LC, Rushton L, Sacco RL, Salomon JA, Sampson U, Sanman E, Schwebel DC, Segui-Gomez M, Shepard DS, Singh D, Singleton J, Sliwa K, Smith E, Steer A, Taylor JA, Thomas B, Tleyjeh IM, Towbin JA, Truelsen T, Undurraga EA, Venketasubramanian N, Vijayakumar L, Vos T, Wagner GR, Wang M, Wang W, Watt K, Weinstock MA, Weintraub R, Wilkinson JD, Woolf AD, Wulf S, Yeh P-H, Yip P, Zabetian A, Zheng Z-J, Lopez AD, Murray CJL (2012) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2095–128.

[33] Wingo TS, Cutler DJ, Wingo AP, Le N-A, Rabinovici GD, Miller BL, Lah JJ, Levey AI (2019) Association of Early-Onset Alzheimer Disease With Elevated Low-Density Lipoprotein Cholesterol Levels and Rare Genetic Coding Variants of *APOB*. *JAMA Neurol.* **76**, 809.

[34] Karjalainen J-P, Mononen N, Hutri-Kähönen N, Lehtimäki M, Juonala M, Ala-Korpela M, Kähönen M, Raitakari O, Lehtimäki T (2019) The effect of apolipoprotein E polymorphism on serum metabolome - a population-based 10-year follow-up study. *Sci. Rep.* **9**, 458.

[35] Whelan CD, Mattsson N, Nagle MW, Vijayaraghavan S, Hyde C, Janelidze S, Stomrud E, Lee J, Fitz L, Samad TA, Ramaswamy G, Margolin RA, Malarstig A, Hansson O (2019) Multiplex proteomics identifies novel CSF and plasma biomarkers of early Alzheimer's disease. *Acta Neuropathol. Commun.* **7**, 169.

[36] Tynkkynen J, Hernesniemi JA, Laatikainen T, Havulinna AS, Sundvall J, Leiviskä J, Salo P, Salomaa V (2016) Apolipoproteins and HDL cholesterol do not associate with the risk of future dementia and Alzheimer's disease: the National Finnish population study (FINRISK). *Age (Dordr).* **38**, 465–473.

[37] Emerging Risk Factors Collaboration TERF, Di Angelantonio E, Gao P, Pennells L, Kaptoge S, Caslake M, Thompson A, Butterworth AS, Sarwar N, Wormser D, Saleheen D, Ballantyne

CM, Psaty BM, Sundström J, Ridker PM, Nagel D, Gillum RF, Ford I, Ducimetiere P, Kiechl S, Koenig W, Dullaart RPF, Assmann G, D'Agostino RB, Dagenais GR, Cooper JA, Kromhout D, Onat A, Tipping RW, Gómez-de-la-Cámara A, Rosengren A, Sutherland SE, Gallacher J, Fowkes FGR, Casiglia E, Hofman A, Salomaa V, Barrett-Connor E, Clarke R, Brunner E, Jukema JW, Simons LA, Sandhu M, Wareham NJ, Khaw K-T, Kauhanen J, Salonen JT, Howard WJ, Nordestgaard BG, Wood AM, Thompson SG, Boekholdt SM, Sattar N, Packard C, Gudnason V, Danesh J (2012) Lipid-related markers and cardiovascular disease prediction. *JAMA* **307**, 2499–506.

[38] Slot RER, Van Harten AC, Kester MI, Jongbloed W, Bouwman FH, Teunissen CE, Scheltens P, Veerhuis R, van der Flier WM (2017) Apolipoprotein A1 in Cerebrospinal Fluid and Plasma and Progression to Alzheimer's Disease in Non-Demented Elderly. *J. Alzheimer's Dis.* **56**, 687–697.

[39] Kritharides L, Nordestgaard BG, Tybjærg-Hansen A, Kamstrup PR, Afzal S (2017) Effect of APOE ε Genotype on Lipoprotein(a) and the Associated Risk of Myocardial Infarction and Aortic Valve Stenosis. *J. Clin. Endocrinol. Metab.* **102**, 3390–3399.

[40] Xing Y, Jia JP, Ji XJ, Tian T (2013) Estrogen associated gene polymorphisms and their interactions in the progress of Alzheimer's disease. *Prog. Neurobiol.* **111**, 53–74.

[41] Chu L-W, Tam S, Wong RLC, Yik P-Y, Song Y, Cheung BMY, Morley JE, Lam KSL (2010) Bioavailable testosterone predicts a lower risk of Alzheimer's disease in older men. *J. Alzheimers. Dis.* **21**, 1335–45.

[42] Panizzon MS, Hauger R, Xian H, Vuoksimaa E, Spoon KM, Mendoza SP, Jacobson KC, Vasilopoulos T, Rana BK, McKenzie R, McCaffery JM, Lyons MJ, Kremen WS, Franz CE (2014) Interaction of APOE genotype and testosterone on episodic memory in middle-aged men. *Neurobiol. Aging* **35**, 1778.e1–8.

[43] Heneka MT, Carson MJ, El Khoury J, Landreth GE, Brosseron F, Feinstein DL, Jacobs AH, Wyss-Coray T, Vitorica J, Ransohoff RM, Herrup K, Frautschy SA, Finsen B, Brown GC, Verkhratsky A, Yamanaka K, Koistinaho J, Latz E, Halle A, Petzold GC, Town T, Morgan D, Shinohara ML, Perry VH, Holmes C, Bazan NG, Brooks DJ, Hunot S, Joseph B, Deigendesch N, Garaschuk O, Boddeke E, Dinarello CA, Breitner JC, Cole GM, Golenbock DT, Kummer MP (2015) Neuroinflammation in Alzheimer's disease. *Lancet. Neurol.* **14**, 388–405.

[44] Tao Q, Ang TFA, DeCarli C, Auerbach SH, Devine S, Stein TD, Zhang X, Massaro J, Au R, Qiu WQ (2018) Association of Chronic Low-grade Inflammation With Risk of Alzheimer Disease in ApoE4 Carriers. *JAMA Netw. open* **1**, e183597.

620 [45] Martiskainen H, Takalo M, Solomon A, Stančáková A, Marttinen M, Natunen T, Haapasalo
621 A, Herukka S-K, Kuusisto J, Soininen H, Kivipelto M, Laakso M, Hiltunen M (2018)
622 Decreased plasma C-reactive protein levels in APOE ε4 allele carriers. *Ann. Clin. Transl.*
623 *Neurol.* **5**, 1229–1240.

624 [46] Yavuz BB, Yavuz B, Halil M, Cankurtaran M, Ulger Z, Cankurtaran ES, Aytemir K, Ariogul S
625 (2008) Serum elevated gamma glutamyltransferase levels may be a marker for oxidative
626 stress in Alzheimer's disease. *Int. Psychogeriatrics* **20**, 815–823.

627 [47] Pike AF, Kramer NI, Blaauboer BJ, Seinen W, Brands R (2015) An alkaline phosphatase
628 transport mechanism in the pathogenesis of Alzheimer's disease and neurodegeneration.
629 *Chem. Biol. Interact.* **226**, 30–39.

630 [48] Kaur G, Levy E (2012) Cystatin C in Alzheimer's disease. *Front. Mol. Neurosci.* **5**, 79.

631 [49] Sundelöf J, Ärnlöv J, Ingelsson E, Sundström J, Basu S, Zethelius B, Larsson A, Irizarry MC,
632 Giedraitis V, Rönnemaa E, Degerman-Gunnarsson M, Hyman BT, Basun H, Kilander L,
633 Lannfelt L (2008) Serum cystatin C and the risk of Alzheimer disease in elderly men.
634 *Neurology* **71**, 1072.

635 [50] Li T, Xie Y, Bowe B, Xian H, Al-Aly Z (2017) Serum phosphorus levels and risk of incident
636 Dementia. *PLoS One* **12**,.

637

638 **Table 1**: Baseline descriptive values.

	Mean (SD)	**Median**	**Min**	**Max**
Age	56.9 (8.0)	58	39	73
Townsend deprivation score	-1.57 (2.9)	2.35	-6.26	10.8
All-cause dementia	**Cases N (%)**		**Controls N (%)**	
	1,852 (0.47%)		393,917 (99.53%)	
Alzheimer's disease	**Cases N (%)**		**Controls N (%)**	
	722 (0.18%)		395,047 (99.82%)	
	Male		**Female**	
Sex	181,000 (45.7%)		214,771 (54.3%)	

	e3/e3	e3/e4	e4/e4	e3/e2	e2e2
***APOE* genotype**	230,094 (59.7%)	94,395 (24.5%)	9,518 (2.4%)	49,082 (12.7%)	2,557 (0.6%)

	Non-smoker	Previously smoked	Current smoker
Smoking status	216,094 (54.8%)	138,744 (35.2%)	39,580 (10%)

	No	**Yes**
CHD	466,843 (96.5%)	16,920 (3.5%)
High BP	381,072 (78.8%)	102,691 (21.2%)
Type 2 diabetes	376,233 (95.3%)	18,702 (4.7%)
Medication	329,322 (83.4%)	65,634 (16.6%)

Biomarkers (not standardised)	**Mean**	**Standard deviation**	**Min**	**Max**	**N**
Low-density lipoprotein	3.6	0.9	0.5	7.9	376,668
High-density lipoprotein	1.5	0.4	0.2	3.4	345,315
Insulin-growth like factor-1	21.4	5.5	1.4	49.6	375,135

Hemoglobin A1C	35.6	5.1	15.0	68.1	374,702
Lipoprotein A	44.1	49.5	3.8	189.0	300,249
Oestradiol	71.6	173.8	8.5	1313.3	349,998
Phosphate	1.2	0.2	0.4	2.0	344,894
Rheumatoid factor	4.4	4.9	3.2	47.4	374,322
Sex hormone-binding globulin	51.4	26.4	0.4	189.3	341,418
Total Bilirubin	9.0	3.9	1.1	31.2	374,181
Direct bilirubin	1.8	0.7	1.0	6.0	320,111
Testosterone	6.5	6.0	0.4	36.7	341,975
Protein	72.4	4.0	36.3	92.5	345,085
Triphosphate	1.7	1.0	0.2	6.9	375,888
Urate	309.1	79.9	89.1	708.8	376,905
Vitamin D	49.9	20.8	10.0	154.0	360,598
C-reactive protein	2.3	3.0	0.1	24.2	373,817
Gamma-glutamyltransferase	34.9	28.1	5.0	241.7	374,958
Cystatin C	0.9	0.1	0.3	1.8	376,382
ApoB	1.0	0.2	0.4	2.0	375,576
ApoA	1.5	0.3	0.4	2.5	343,506
Creatinine	72.0	14.5	10.8	161.4	376,640
Total cholesterol	5.7	1.1	1.4	11.4	377,349
Calcium	2.4	0.1	1.2	2.8	345,223
Urea	5.4	1.3	0.8	12.3	376,453
Aspartate transaminase (AST)	25.8	7.5	3.3	78.7	374,496
Alanine transaminase (ALT)	23.0	11.6	3.0	93.7	375,656
Alkaline phosphatase	83.1	22.9	8.0	215.0	376,658
Albumin	45.2	2.6	18.9	58.2	345,558

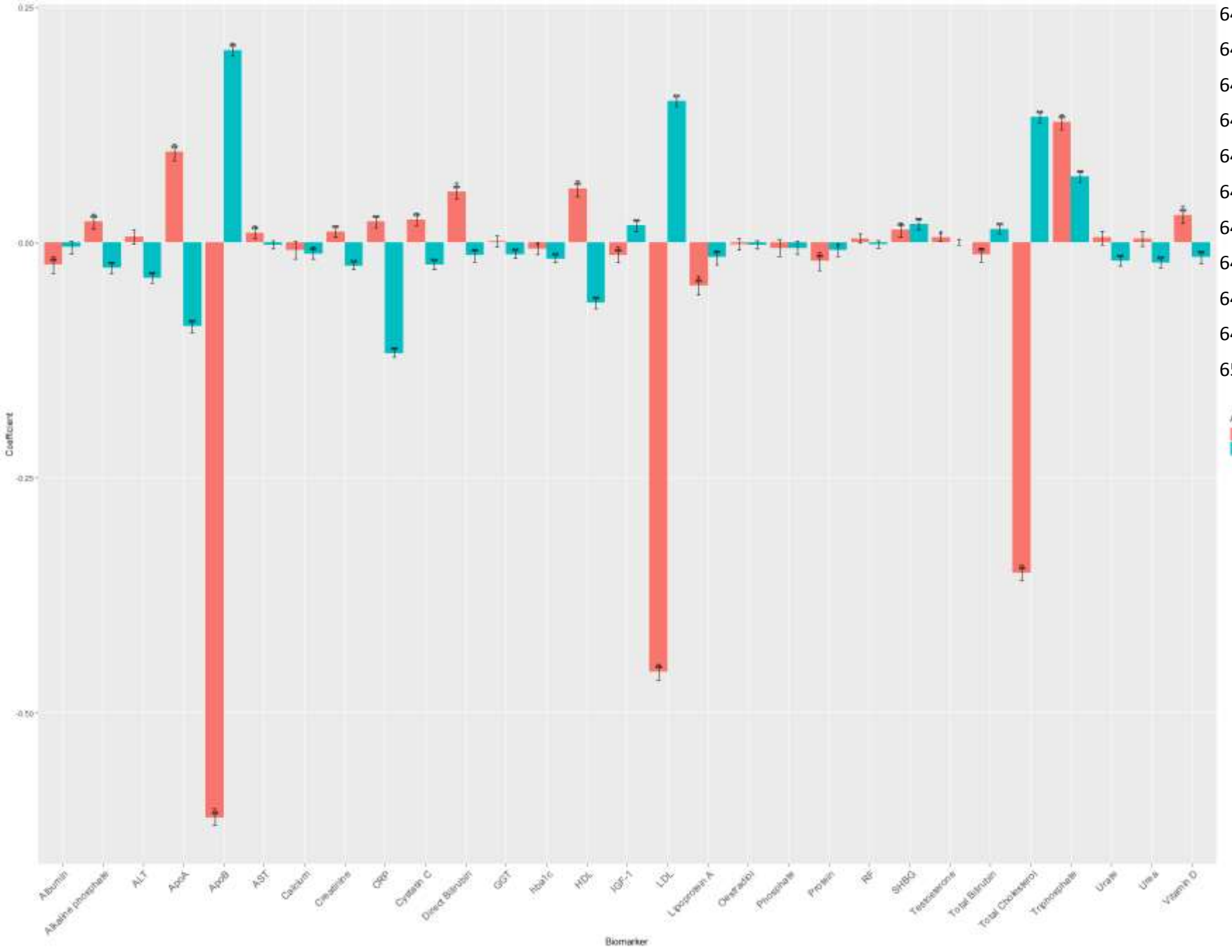

Figure 1: linear regression fully-adjusted standardized betas comparing *APOE* e4 and e2 genotypes (per allele) vs. neutral e3/e3, on biomarker values. P<0.05, **P<0.001