**Title:** Alzheimer disease genetic risk factor *APOE* e4, and cognitive abilities in 111,739 UK Biobank participants.

**Authors**: Lyall, D.M.[1], Ward, J.[1], Ritchie, S.J.[2], Davies, G.[2], Cullen, B. [1], Celis, C.[3], Bailey, M.E.S.[4], Anderson, J.[1], Evans, J. [1], Mackay, D.F. [1], McIntosh, A.M.[2], Sattar, N.[3], Smith, D.J.[1], Deary, I.J.[2]. & Pell, J.P.[1]

[1] Institute of Health & Wellbeing, University of Glasgow, Scotland, UK

[2] Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, Scotland, UK

[3] Institute of Cardiometabolic and Medical Sciences, University of Glasgow, Scotland, UK.

[4] School of Life Sciences, CMVLS, University of Glasgow, Scotland, UK.

**Key words:** UK Biobank; Epidemiology; Cognitive ability; *APOE*; Alzheimer disease.

**Conflict of interest disclosures:** IJD was a participant in UK Biobank; this does not influence any motivations, analyses, or interpretations with regard to this study. The authors have no other competing interests to report.

**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, Scotland UK.

| | |
|---|---|
| Abstract word count | 178 |
| Main text word count | 3,172 |
| References | 25 |
| Tables | 2 |
| Supplementary Tables | 2 |

**Author statement**

Conflict of interest disclosures

One co-author is a UK Biobank participant. One other co-author is a member of the UK Biobank scientific advisory committee. Neither of these conflicts of interest affected the current study with regards motivation, analysis or interpretation. There are no other actual or potential conflicts of interest including financial, personal or otherwise. No author's institution has contracts relating to this research through which it or any other organization may stand to gain financially now or in the future.

Ethical approval

This study was conducted under generic approval from the NHS National Research Ethics Service (approval letter dated 17th June 2011, Ref 11/NW/0382).

Financial Support

This research has been conducted using the UK Biobank resource; we are grateful to UK Biobank participants. UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council (MRC), 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 work of the Edinburgh-based authors was undertaken within the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (http://www.ccace.ed.ac.uk), part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and MRC is gratefully acknowledged.

Data verification

The data contained in the manuscript being submitted have not previously been published, have not been submitted elsewhere and will not be submitted elsewhere while under consideration at Age and Ageing.

**Abstract**

Background: The apolipoprotein (*APOE*) e4 locus is a genetic risk factor for dementia. Carriers of the e4 allele may be more vulnerable to conditions that are independent risk factors for cognitive decline, such as cardiometabolic diseases.

Objective: We tested whether any association with *APOE* e4 status on cognitive ability was larger in older ages, or in those with cardiometabolic diseases.

Subjects: UK Biobank includes over 500,000 middle- and older-aged adults who have undergone detailed medical and cognitive phenotypic assessment. Around 150,000 currently have genetic data. We examined 111,739 participants with complete genetic and cognitive data.

Methods: Baseline cognitive data relating to information processing speed, memory and reasoning were used. We tested for interactions with age and with presence vs. absence of type-2 diabetes (T2D), coronary artery disease (CAD) and hypertension.

Results: In several instances *APOE* e4 dosage interacted with older age and disease presence to affect cognitive scores. When adjusted for potentially confounding variables there was no *APOE* e4 effect on the outcome variables.

Conclusions: Future research in large independent cohorts should continue to investigate this important question, which has potential implications for aetiology related to dementia and cognitive impairment.

## Highlights

- It is of great public health interest to determine if genetic and environmental risk factors for dementia interact.
- Effects of *APOE* e4 genotype on cognitive ability may be stronger in older ages, or in people with cardiometabolic diseases.
- Large sample size, homogenous ancestry and cognitive data in UK Biobank make the cohort very valuable.
- We tested our hypothesis in around 112,000 UK Biobank participants with genetic, cognitive and sociodemographic data.
- We found no evidence that older age or cardiometabolic diseases modified association between *APOE* e4 and cognitive ability.

**Key words:** UK Biobank; Epidemiology; Cognitive ability; *APOE*; Alzheimer disease.

| | |
|---|---|
| Abstract word count | 178 |
| Main text word count | 3,172 |
| References | 25 |
| Tables | 2 |
| Supplementary Tables | 2 |

## Introduction

Alzheimer disease (AD) is a type of progressive dementia characterized by cognitive impairment, and genetic variation in the apolipoprotein e (*APOE*) region is a risk factor. Previous studies have reported worse cognitive abilities in non-demented people with the *APOE* e4 allele (e3/e4, e4/e4 and occasionally e2/e4 genotypes), compared with those with the e2/e2, e2/e3 or e3/e3 genotype. Wisdom et al. reported a meta-analysis including 40,942 participants with a mean age of 63.14 years (standard deviation [SD] = 13.10). Significant negative associations were found between presence of the e4 allele and cognitive abilities relating to memory, global cognitive function, executive function and information processing speed, but not verbal ability, primary memory, attention or visuospatial functions[1].

A recent paper using the CHARGE consortium data[2], which includes 543,949 European participants from 31 cohorts, all aged over 45 years and without dementia, identified 13 single nucleotide polymorphisms (SNPs) that were significantly associated with fluid cognitive function as derived from several diverse cognitive tests. The magnitude of the association between rs10119 in the *APOE* region, and worse cognitive ability, increased with the mean age of each cohort ($r$ = -0.42, P = 0.022). The effect was close to zero in younger cohorts, aged around 55-60 years, and most pronounced in the oldest cohorts, aged around 80 years[2]. A report using the Generation Scotland cohort (N = 18,337) reported effects of *APOE* e4 dosage (0 vs. 1 vs. 2) on Logical Memory (standardised beta = -0.095P = 0.003), Verbal Fluency (standardised beta = 0.075, P = 0.023), and Digit Symbol tests (standardised beta = -0.087, P = 0.004), only in participants aged over 60 years[3]. In summary, there is evidence from relatively large cross-sectional studies that the effect of e4 genotype may become stronger in older ages[4,5].

The *APOE* locus has been referred to as the 'frailty gene'[6,7], because it is associated with increased risk of diseases which lead to disability[8], and e4 carriers are perhaps less likely to recover well[9]. Recent studies have investigated possible interactions between the e4 allele and cardiometabolic diseases. Perna et al. reported that in the ESTHER and KAROLA cohorts (total N = 1800), the negative association between *APOE* e4 presence and performance on the Cognitive Telephone Screening Instrument (COGTEL) was significantly stronger in participants with hypercholesterolemia (interaction P-values <0.05). There is a lack of high quality data on the possible interactive effect of *APOE* e4 genotype and cardiometabolic risk factors on cognitive abilities[10].

UK Biobank is a very large, general population cohort study of 502,649 people from the UK, which includes sociodemographic, medical, cognitive and genetic data[11].The genetic data are to be made available in batches. Genetic data for the first 150,000 participants have been released so far. UK Biobank has cognitive data pertinent to a number of domains previously associated with *APOE* e4: episodic memory, executive function and processing speed[1]. This study aimed to ascertain whether *APOE* e4 allele dosage was associated with cognitive function in a large, single-cohort population sample, represented by the currently released UK Biobank genotype dataset; whether genotype effects are stronger in older ages, or interact with cardiometabolic disease.

## Methods

### Materials and procedure

The UK Biobank baseline assessment took place between 2006 and 2010 in one of 22 assessment centres. In total 502,649 participants were recruited, aged 40-70 years, and from the general population. Invitation letters were sent to eligible adults registered with the NHS and living within 25 miles of a study assessment centre. Participants completed a touchscreen questionnaire related to various topics including sociodemographic, physical and mental health, early life factors, and a relatively brief battery of cognitive tests.

This study focuses on three cognitive tests that were included in UK Biobank, each administered via computerised touchscreen interface. The first of these was a task with thirteen logic/reasoning-type questions and a two-minute time limit, labelled as 'fluid intelligence' in the UK Biobank protocol but hereafter referred to as 'verbal-numerical reasoning'; http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=20016). The maximum score is 13. The next task was a visual memory test labelled 'pairs-matching' (http://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=100030), where participants were asked to memorize the positions of six card pairs, and then match them from memory while making as few errors as possible. Scores on the pairs-matching test are for the number of errors that each participant made; therefore, higher scores reflect poorer cognitive function. We refer to this test as the memory task from here on. Finally, participants completed a timed test of symbol matching, similar to the common card game 'Snap' hereafter referred to as reaction time (RT) (http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=20023). The reasoning task was only added to the participant assessment part-way through the baseline assessment phase. Therefore, the sample sizes for the different tasks vary. UK Biobank cognitive data have been described previously[12].

Participants were asked during the baseline assessment about any previous or current cardiometabolic conditions that had been diagnosed by their doctor. Specifically, participants were asked whether their doctor had diagnosed myocardial infarction, angina, stroke, hypertension or type II diabetes (T2D). We defined coronary artery disease (CAD) as either myocardial infarction or angina. We excluded participants who stated only 'prefer not to answer'. Townsend deprivation indices were derived from postcode of residence[13]. They provide 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.

Genotyping

UK Biobank genotyping was conducted by Affymetrix using a bespoke BiLEVE Axiom array for ~50,000 participants, and the remaining ~450,000 (for the purposes of this study 100,000) on the Affymetrix UK Biobank Axiom array. The two are extremely similar. All genetic data were quality controlled and imputed by UK Biobank. The *APOE* e genotype is directly genotyped. Further information on the genotyping process is available (http://www.ukbiobank.ac.uk/scientists-3/genetic-data), including detailed technical documentation (http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UKBiobank_genotyping_QC_documentation-web.pdf). UK Biobank provides recommendations, which we followed, for which participants to exclude from analysis, generally because the sample failed QC, had significant missing data or heterozygosity. UK Biobank also provides 15 separate principal components (PCs) to be included as possible covariates in order to limit confounding by population stratification.

Statistical analyses

We included only participants of Caucasian ethnicity, who accounted for the vast majority (~95%) of UK Biobank participants. In common with previous studies on *APOE* e[1], we excluded participants with the e2/e4 genotype as they have a combination of potentially protective and risk alleles.

Because the RT scores were significantly positively distributed we transformed the variable with a natural log transform ('LN'). The memory error scores were significantly zero-value inflated and therefore transformed with an LN +1 equation. We ran the models coding e2/e2, e2/e3 and e3/e3, e3/e4, and e4/e4 as 0, 1 and 2 respectively, i.e. an e4 dosage model. We found no evidence that the results were meaningfully different when we analysed the data as a binary *APOE* e4 present vs. absent model, so simply report the dose effect.

Participants with zero, one or two *APOE* e4 alleles were compared on descriptive variables with $\chi^2$ tests for trend for ordinal data, using the "nptrend" function in STATA [14]. The association between *APOE* e4 dosage and the cognitive function test results was tested using linear regression. For log RT and log memory scores we report exponentiated betas where effects are multiplicative so that an effect size of 1.00 represents no change, and e.g. beta = 1.01 equates to a 1% increase in raw RT scores. For reasoning scores we report unstandardized betas. We included 10 genetic PCs, as provided by UK Biobank. We tested whether there were significant interactions between *APOE*, age, and diseases of CAD, T2D and hypertension, and also age in years.

We first ran a base model adjusted for stratification with 10 PCs, and then adjusted for the covariates of age and sex ('partially adjusted model'), and finally self-reported depression,

Townsend deprivation index, T2D, CAD and hypertension ('fully adjusted model'). We tested each interaction term in a separate model. Genotype measurement batch, assessment centre and type of array made no difference to the final results when included as a covariate and we therefore present results without this in the model. As an additional analysis we split the sample into those over vs. under 60 years of age, and tested for an effect of *APOE* e4 at first unadjusted, and then adjusted for stratification and sex. This method of analysis is similar to that reported by Marioni et al. Because of the large sample size and partly to offset the risk of multiple comparisons we elected to use P <0.001 as nominal significance. Age was centred on 60 years. All analyses were conducted using IBM SPSS v.22 unless otherwise stated.

**Results**

**[Insert Table 1 here]**

Of 152,248 UK Biobank participants who had *APOE* genotyping, we first excluded participants that either reported a brain disease that may affect cognitive ability directly (see Appendix 1 in supplementary data) or did not provide that information: this left 145,278 participants. We excluded non-Caucasian participants, leaving n = 130,381. We removed participants who had a relatedness coefficient above 0.0442 (n = 15,625), or mismatch between reported and genetic sex (n = 152), leaving N = 114,604. We then removed participants who failed UK Biobank or BiLeve quality controlling (QC; n = 281), leaving N = 114,323 One participant was removed because they were aged over 70 years (above the official upper baseline age limit). This left a total of 114,322 participants with *APOE* data. Of those, 1,087 (1.0%) had *APOE* genotype e2/e2; 13,352 (11.7%) had e2/e3; 2,583 (2.3%) had e2/e4; 72,001 (63.0%) had e3/e3; 22,029 (19.3%) had e3/e4

and 3,270 (2.9%) had e4/e4. The allele frequencies were: e2 = 7.92%, e3 = 78.46% and e4 = 13.62%. The *APOE* e genotype was in Hardy-Weinberg equilibrium (P = 0.08). Because we excluded the e2/e4 carriers from analysis, the final sample was 111,739. The mean age was 56.79 years (SD = 7.94), and 52,679 (47.1%) were male. Frequencies of diseases and demographics by genotype are shown in Table 1. Of the 111,739 participants, 36,135 had reasoning data, 111,651 pairs-matching data, and 111,137 RT data.

**[Insert Table 2 here]**

Reasoning scores

In the base model, partially adjusted and fully adjusted models there was no association between *APOE* e4 dosage and reasoning scores (Table 2). In the base model there was a significant interaction between *APOE* e4 and CAD, and between *APOE* e4 and age, but no associations remained significant in the fully adjusted model.

Reaction time

In the base model there were several significant interactions (i.e. P value <0.001) between *APOE* e4 and cardiometabolic diseases, and age (Table 2). Most of these interactions attenuated in the partially adjusted model however, and none were significant in the fully adjusted model.

Pairs matching (memory) error scores

In the base model there were significant interactions between *APOE* e4 dosage and hypertension, and age, on log memory errors. However there were no significant interactions in the partially or fully adjusted models (Table 2).

Age-stratified analyses

When we split the sample into participants aged over vs. under 60 years, there was no evidence of a significant *APOE* e4 effect (at P = 0.001) in either group whether unadjusted, or corrected for stratification (with 10 PCs) and sex (Appendix 2).

## Discussion

Interpretation

In our study of around 112,000 participants of UK Biobank we found some statistically significant interactions between *APOE* e4 dosage and age/disease variables: for reasoning scores with CAD and age; for RT scores with CAD, T2D, hypertension and age, and for pairs matching memory errors with hypertension and age. However none of these interactions were significant in the fully adjusted model.

Our findings partly contrast with the recent finding by Davies et al. who reported that the e4 effect on fluid mental ability was much more pronounced in older age[2]. Specifically, Davies et al. noted that the *APOE* e4 effect size was mostly apparent at ages above 60 years, increasing relatively linearly until around age 80 (the mean age of the oldest cohort). Marioni et al. found generally weaker associations in participants aged under 60 years; in contrast we found no significant associations with *APOE* e4 dosage when we stratified participants in this way. We

found no interactions with self-reported T2D, hypertension or CAD for any of the cognitive tests in fully adjusted analyses. This is in contrast to a recent studies which showed significant interaction between *APOE* e4 and presence of cardiometabolic disease on cognitive function[10,15].

Wisdom et al. reported on 77 studies including 40,942 participants with *APOE* and cognitive data, and reported generally small effects (*r* coefficients ranged from -0.07 for episodic memory, to -0.002 for verbal ability), which are mostly larger than seen here. (We recommend the Wisdom et al. report for collating a large number of individual studies, detailed therein, although several reports of *APOE* e4 genotype and non-pathological cognitive ability in older adults been published since, generally showing significant associations with small-to-medium effect sizes [3,10,16–21]). Our study is based on a single very large cohort where participants completed the same battery of cognitive function tests. In contrast, the report by Wisdom et al. pooled effect sizes from different studies using different measures (although the authors were careful to include only validated tests). Our results also statistically control for a number of factors – e.g. deprivation – that Wisdom et al. did not. Therefore, it is possible that our largely negative findings are a more accurate reflection of the impact of *APOE* e4 on the cognitive function of people in middle age or early old age who are not affected by dementia. However, these unexpected findings require significant validation and replication in future studies, particularly when the UK Biobank genetic data are available on the full cohort of 502,649 participants. It is common for effect sizes in single-locus association studies to weaken when more participants are added [22,23]. It should also be noted that UK Biobank did not assess any tasks related to verbal episodic memory, which is the domain which Wisdom et al. reported had the strongest association with *APOE* e4 genotype.

Our study was cross-sectional and did not consider the *effects* of *APOE* e4 on longitudinal cognitive change[24]. Future research will address this. In addition, since the baseline assessment,

UK Biobank has assessed participants on cognitive tests based on well-validated tasks such as Trail Making Test, and Tower of Hanoi, and we will investigate these also.

Limitations

UK Biobank are unlikely to be perfectly representative of the general population in terms of age, sex, ethnicity and socioeconomic status within the age-range recruited [25]. However, in common with other general population cohorts, it is likely to be unrepresentative in other aspects. People with very low cognitive ability may have been less likely to participate due to various reasons including but not limited to inability to provide informed consent, or difficulties in participating. Also, since recruitment occurred in middle and old age, survival bias may have been introduced. If the range of cognitive ability among participants was reduced as a result, this may have masked a true association between genotype and cognitive ability.

It is also possible that the UK Biobank cognitive tests are insensitive, which may mask or partly offset the deleterious effects of *APOE* e4 genotype on mental function. Due to the very large scale of UK Biobank, the cognitive battery had to be delivered within a very short time-window. The tests also needed to be self-completed via a touch-screen computer. Therefore, they were relatively novel although based in principal on existing cognitive tasks. It is possible that some participants, uncomfortable with what could be considered a 'high tech' computerized approach, may have declined to participant, and this could theoretically have introduced some bias although it should be noted that only a very small minority of participants refuse to complete the cognitive assessment. The tests used here have shown a main effect of age[12], implying they are at least relatively sensitive metrics of cognitive change.

## Conclusions

The *APOE* e4 locus has been significantly associated with cognitive ability in non-demented adults, and there is prior evidence that the effect may increase with age, and perhaps modify the negative association with cardiometabolic diseases. We tested for these effects in 111,739 UK Biobank participants. We found significant interactions between *APOE* e4 genotype with presence of cardiometabolic diseases, and older age on cognitive abilities, although the effects were small and did not survive correction for confounders. Given the large sample sizes reported here, it is also possible that the unadjusted significant findings reflect type-1 error. Future studies should continue to investigate whether genetic risk factors for lower mental ability have a greater effect into older age, including when more genetic and cognitive data is available in UK Biobank.

## Table contents



## References

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

2 Davies G, Armstrong N, Bis JC, Bressler J, Chouraki V, Giddaluru S, *et al.* Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide

association studies in the CHARGE consortium (N=53949). *Mol Psychiatry* 2015; **20**: 183–92.

3 Marioni RE, Campbell A, Scotland G, Hayward C, Porteous DJ, Deary IJ. Differential effects of the APOE e4 allele on different domains of cognitive ability across the life-course. *Eur J Hum Genet* 2015. doi:10.1038/ejhg.2015.210.

4 Papenberg G, Lindenberger U, Bäckman L. Aging-related magnification of genetic effects on cognitive and brain integrity. *Trends Cogn Sci* 2015. doi:10.1016/j.tics.2015.06.008.

5 Laukka EJ, Lövdén M, Herlitz A, Karlsson S, Ferencz B, Pantzar A, *et al.* Genetic effects on old-age cognitive functioning: a population-based study. *Psychol Aging* 2013; **28**: 262–74.

6 Smith JD. Apolipoproteins and aging: emerging mechanisms. *Ageing Res Rev* 2002; **1**: 345–65.

7 Deary IJ, Whiteman MC, Pattie A, Starr JM, Hayward C, Wright AF, *et al.* Apolipoprotein e gene variability and cognitive functions at age 79: a follow-up of the Scottish mental survey of 1932. *Psychol Aging* 2004; **19**: 367–71.

8 McCarron MO, Delong D, Alberts MJ. APOE genotype as a risk factor for ischemic cerebrovascular disease: A meta-analysis. *Neurology* 1999; **53**: 1308–1308.

9 Zhou W, Xu D, Peng X, Zhang Q, Jia J, Crutcher KA. Meta-Analysis of APOE 4 Allele and Outcome after Traumatic Brain Injury. *J Neurotrauma* 2008; **25**: 279–90.

10 Bangen KJ, Beiser A, Delano-Wood L, Nation DA, Lamar M, Libon DJ, *et al.* APOE Genotype Modifies the Relationship between Midlife Vascular Risk Factors and Later Cognitive Decline. *J Stroke Cerebrovasc Dis* 2013; **22**: 1361–9.

11 Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med* 2015; **12**: e1001779.

12 Cullen B, Nicholl BI, Mackay DF, Martin D, Ul-Haq Z, McIntosh A, *et al.* Cognitive function and lifetime features of depression and bipolar disorder in a large population sample: Cross-sectional study of 143,828 UK Biobank participants. *Eur Psychiatry* 2015; **30**: 950–8.

13 Townsend P. Townsend deprivation index. Natl. database Prim. care groups Trust. 1998.

14 StataCorp. Stata Statistical Software: Release 11. Econ. J. 2009; **102**. doi:10.2307/2234838.

15 Perna L, Mons U, Rujescu D, Kliegel M, Brenner H. Apolipoprotein E e4 and Cognitive Function: A Modifiable Association? Results from Two Independent Cohort Studies. *Dement Geriatr Cogn Disord* 2015; **41**: 35–45.

16 Davies G, Harris SE, Reynolds C a, Payton a, Knight HM, Liewald DC, *et al.* A genome-wide association study implicates the APOE locus in nonpathological cognitive ageing. *Mol Psychiatry* 2014; **19**: 76–87.

17 Gerritsen L, Comijs HC, Deeg DJH, Penninx BWJH, Geerlings MI. Salivary cortisol, APOE-ε4 allele and cognitive decline in a prospective study of older persons. *Neurobiol Aging* 2011; **32**: 1615–25.

18 Izaks GJ, Gansevoort RT, van der Knaap AM, Navis G, Dullaart RPF, Slaets JPJ. The association of APOE genotype with cognitive function in persons aged 35 years or older. *PLoS One* 2011; **6**: e27415.

19 Woodfield R, Grant I, Sudlow CLM. Accuracy of Electronic Health Record Data for Identifying Stroke Cases in Large-Scale Epidemiological Studies: A Systematic Review from the UK Biobank Stroke Outcomes Group. *PLoS One* 2015; **10**: e0140533.

20 Caselli RJ, Dueck AC, Locke DEC, Hoffman-Snyder CR, Woodruff BK, Rapcsak SZ, *et al.* Longitudinal modeling of frontal cognition in APOE μ4 homozygotes, heterozygotes, and noncarriers. *Neurology* 2011; **76**: 1383–8.

21 Lyall DM, Harris SE, Bastin ME, Muñoz Maniega S, Murray C, Lutz MW, *et al.* Are APOE [ genotype and TOMM40 poly-T repeat length associations with cognitive ageing mediated by brain white matter tract integrity? *Transl Psychiatry* 2014; **4**: e449.

22 Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, *et al.* Confidence and precision increase with high statistical power. *Nat Rev Neurosci* 2013; **14**: 585–6.

23 Lyall DM, Lopez LM, Bastin ME, Maniega SM, Penke L, Valdés Hernández M del C, *et al.* ADRB2, brain white matter integrity and cognitive ageing in the Lothian Birth Cohort 1936. *Behav Genet* 2013; **43**: 13–23.

24 Schiepers OJG, Harris SE, Gow AJ, Pattie A, Brett CE, Starr JM, *et al.* APOE E4 status predicts age-related cognitive decline in the ninth decade: longitudinal follow-up of the Lothian Birth Cohort 1921. *Mol Psychiatry* 2011; **17**: 315–24.

25 Smith DJ, Nicholl BI, Cullen B, Martin D, Ul-Haq Z, Evans J, *et al.* Prevalence and characteristics of probable major depression and bipolar disorder within UK Biobank: Cross-sectional study of 172,751 participants. *PLoS One* 2013; **8**. doi:10.1371/journal.pone.0075362.

**Table 1: Clinical and demographic characteristics**

| | ***APOE* e2/e2, e2/e3 and e3/e3 (N = 86,440)** | ***APOE* e3/e4 (N = 22,029)** | ***APOE* e4/e4 (N = 3,270)** | **P value** |
|---|---|---|---|---|
| Age, mean (SD) | 56.85 (7.94) | 56.59 (7.95) | 56.53 (7.89) | <0.001 |
| Sex, male N (%) | 40,773 (47.2) | 10,307 (46.8) | 1,599 (48.9) | 0.663 |
| Townsend deprivation score, mean (SD) | -1.48 (2.99) | -1.57 (2.94) | -1.47 (3.07) | 0.002 |
| Reasoning scores, mean (SD) | 6.16 (2.10) | 6.23 (2.10) | 6.13 (2.14) | 0.130 |
| Log transformed RT score, mean (SD) | 6.30 (0.18) | 6.30 (0.18) | 6.29 (0.18) | 0.262 |
| Untransformed RT score, median (IQR) | 535 (477-605) | 535 (477-605) | 531 (476-602) | 0.262 |
| Log-transformed pairs-matching errors, mean (SD) | 1.43 (0.64) | 1.43 (0.64) | 1.44 (0.65) | 0.273 |
| Untransformed pairs-matching errors, median (IQR) | 3 (2-5) | 3 (2-5) | 3 (2-6) | 0.273 |
| Depression, N (%) | 4,926 (5.7) | 1,345 (6.1) | 177 (5.4) | 0.195 |
| Hypertension, N (%) | 23,333 (27.0) | 5,905 (26.8) | 849 (26.0) | 0.226 |
| Type-2 diabetes, N (%) | 4,527 (5.2) | 990 (4.5) | 139 (4.3) | <0.001 |
| Coronary artery disease, N (%) | 3,838 (4.4) | 1,074 (4.9) | 157 (4.8) | 0.008 |

P-values (two-tailed) for frequency statistics (e.g. depression) are based on $Ç^2$ trend tests for continuous/categorical data. RT = reaction time, IQR = interquartile range, SD = standard deviation.

**Table 2:** Associations between *APOE* e4 dosage, cardiometabolic diseases and cognitive scores

| | **Base model[1]** | | | | **Partially adjusted[2]** | | | | **Fully adjusted[3]** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Beta coefficient | 95% CI | | P-value | Beta coefficient | 95% CI | | P-value | Beta coefficient | 95% CI | | P-value |
| | | | | | Fluid reasoning | | | | | | | |
| *APOE* e4 (dosage) | 0.031 | -0.013 | 0.076 | 0.161 | 0.029 | -0.015 | 0.072 | 0.202 | 0.031 | -0.013 | 0.074 | 0.164 |
| *APOE* e4 x CAD | -0.378 | -0.554 | -0.202 | <0.001 | -0.335 | -0.511 | -0.159 | <0.001 | 0.080 | -0.117 | 0.277 | 0.425 |
| *APOE* e4 x type-2 diabetes | -0.129 | -0.308 | 0.049 | 0.155 | -0.107 | -0.285 | 0.071 | 0.238 | 0.046 | -0.149 | 0.241 | 0.643 |
| *APOE* e4 x hypertension | -0.085 | -0.163 | -0.007 | 0.034 | -0.045 | -0.123 | 0.033 | 0.259 | 0.077 | -0.007 | 0.161 | 0.073 |
| *APOE* e4 x age | -0.018 | -0.023 | -0.013 | <0.001 | -0.002 | -0.008 | 0.003 | 0.367 | -0.002 | -0.008 | 0.003 | 0.356 |
| | | | | | Log reaction time* | | | | | | | |
| *APOE* e4 (dosage) | 0.999 | 0.996 | 1.001 | 0.197 | 1.000 | 0.998 | 1.002 | 0.806 | 1.001 | 0.998 | 1.003 | 0.605 |
| *APOE* e4 x CAD | 1.033 | 1.024 | 1.042 | <0.001 | 1.008 | 1.000 | 1.017 | 0.048 | 0.994 | 0.985 | 1.004 | 0.232 |
| *APOE* e4 x type-2 diabetes | 1.030 | 1.021 | 1.040 | <0.001 | 1.016 | 1.007 | 1.025 | <0.001 | 1.004 | 0.994 | 1.013 | 0.452 |
| *APOE* e4 x hypertension | 1.019 | 1.015 | 1.023 | <0.001 | 1.004 | 1.000 | 1.008 | 0.032 | 0.999 | 0.995 | 1.003 | 0.739 |
| *APOE* e4 x age | 1.005 | 1.005 | 1.005 | <0.001 | 1.000 | 1.000 | 1.000 | 0.138 | 1.000 | 1.000 | 1.000 | 0.092 |
| | | | | | Log memory errors* | | | | | | | |
| *APOE* e4 (dosage) | 1.004 | 0.997 | 1.012 | 0.278 | 1.007 | 0.999 | 1.014 | 0.074 | 1.007 | 0.999 | 1.014 | 0.078 |
| *APOE* e4 x CAD | 1.034 | 1.003 | 1.067 | 0.030 | 0.989 | 0.959 | 1.020 | 0.478 | 0.993 | 0.960 | 1.028 | 0.701 |
| *APOE* e4 x type-2 diabetes | 0.990 | 0.959 | 1.023 | 0.555 | 0.965 | 0.935 | 0.996 | 0.029 | 0.980 | 0.947 | 1.015 | 0.269 |
| *APOE* e4 x hypertension | 1.035 | 1.021 | 1.049 | <0.001 | 1.010 | 0.997 | 1.024 | 0.144 | 1.005 | 0.991 | 1.020 | 0.494 |
| *APOE* e4 x age | 1.008 | 1.007 | 1.009 | <0.001 | 1.000 | 0.999 | 1.000 | 0.398 | 1.000 | 0.999 | 1.000 | 0.403 |

CI confidence interval; CAD coronary artery disease; T2D type II diabetes
[1]adjusted for stratification (with ten principal components).
[2]adjusted for stratification, age and sex.
[3]additionally stratification, age, sex, depression, Townsend scores, T2D, hypertension and CAD.
*exponentiated beta coefficients.

**Appendix 1:** Excluded (self-reported) diseases

Brain cancer/primary malignant tumour

Brain haemorrhage

Brain/intracranial abscess

Cerebral aneurysm

Cerebral palsy

Chronic/degenerative neurological problem

Dementia/Alzheimer disease/cognitive impairment

Encephalitis

Epilepsy

Head injury

Infection of nervous system

Ischaemic stroke

Meningeal cancer/malignant meningioma

Meningioma (benign)

Meningitis

Motor neurone disease

Multiple sclerosis

Neurological injury/trauma

Neuroma (benign)

Other demyelinating condition

Other neurological problem

Parkinson disease

Spina bifida

Stroke

Subarachnoid haemorrhage

Subdural haematoma

Transient ischaemic attack

**Appendix 2:** *APOE* e4 and cognitive abilities in people stratified by under vs. over 60 years of age

Under 60 years

| | **Fluid reasoning (N = 19,252)** | | | | **Log reaction time (N = 61,969)** | | | | **Log memory errors (N = 62,145)** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Beta coefficient | 95% CI | | P-value | Beta coefficient* | 95% CI | | P-value | Beta coefficient* | 95% CI | | P-value |
| *APOE* e4 (dosage) | | | | | | | | | | | | |
| Unadjusted | 0.061 | 0.001 | 0.122 | 0.047 | 1.002 | 0.999 | 1.005 | 0.121 | 1.007 | 0.997 | 1.017 | 0.153 |
| Adjusted for sex and stratification | 0.058 | -0.003 | 0.118 | 0.061 | 1.002 | 0.999 | 1.005 | 0.142 | 1.007 | 0.998 | 1.017 | 0.140 |

Over 60 years

| | **Fluid reasoning (N = 16,883)** | | | | **Log reaction time (N = 49,168)** | | | | **Log memory errors (N = 49,506)** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *APOE* e4 (dosage) | | | | | | | | | | | | |
| Unadjusted | -0.011 | -0.075 | 0.053 | 0.733 | 0.997 | 0.993 | 1.000 | 0.043 | 1.004 | 0.992 | 1.016 | 0.498 |
| Adjusted for sex and stratification | -0.006 | -0.070 | 0.057 | 0.884 | 0.997 | 0.994 | 1.000 | 0.056 | 1.004 | 0.993 | 1.016 | 0.469 |

CI confidence interval; CAD coronary artery disease; T2D type II diabetes.
*exponentiated.