Article Title: Assessing for interaction between *APOE* e4, sex and lifestyle on cognitive abilities.

Authors: Donald M. Lyall* PhD[1], Carlos Celis-Morales PhD[2], Laura M. Lyall PhD[1], Christopher Graham[1], Nicholas Graham MBChB[1], Daniel F. Mackay PhD[1], Rona J. Strawbridge PhD[1,3], Joey Ward BSc[1], Jason M. Gill PhD[2], Naveed Sattar MD[2], Jonathan Cavanagh PhD[1], Daniel J. Smith PhD[1], Jill P. Pell MD[1]

* Conducted statistical analysis

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

[2]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Scotland UK.

[3]Department of Medicine Solna, Karolinska Institute, Stockholm, Sweden.

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

Search terms: *Alzheimer's disease; Cognitive aging; Risk factors in epidemiology; Association studies in genetics; All Neuropsychology/Behavior.*

Title character count: 85

Text word count: 3,501

Abstract word count: 158

Tables: 3

Figures: 0

Supplementary tables: 5

Supplementary figure: 1

References: 39.

Authors' contributions: Study concept or design: DML. Drafting/revising the manuscript for content: DML, CC, LML, NG, DFM, RJS, JW, JMG, NS, JC, DJS, JPP. Analysis or interpretation of data: DML, CC, LML, NG, DFM, RJS, JW, JMG, NS, JC, DJS, JPP.

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. DJS is supported by MRC Mental Health Data Pathfinder Award (reference MC_PC_17217). RJS is supported by a UKRI Innovation-HDR-UK Fellowship (MR/S003061/1),

Financial disclosure statement: Smith is partially funded by the Lister institute. Sattar has consulted for Amgen, Inc., Sanofi, AstraZeneca, Eli Lilly, and has sat on the Medical UK Biobank Scientific Advisory Board. 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 and 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. Doctors Lyall (DM), Celis-Morales, Lyall (LM), Graham (N), Mackay, Strawbridge, Prof. Gill, Mr. Graham (C) and Mr. Ward report no disclosures. This study is not industry sponsored.

Abstract

Objective: To test for interactions between apolipoprotein *(APOE)* e4 genotype, and lifestyle factors on worse cognitive abilities in UK Biobank.

Methods: Using UK Biobank cohort data, we tested for interactions between *APOE* e4 allele presence, lifestyle factors of alcohol intake, smoking, total physical activity and obesity, and sex, on cognitive tests of reasoning, information processing speed and executive function (n range=70,988-324,725 depending on the test). We statistically adjusted for potential confounders of age, sex, deprivation, cardiometabolic conditions, and educational attainment.

Results: There were significant associations between *APOE* e4 and worse cognitive abilities, independent of potential confounders, and between lifestyle risk factors and worse cognitive abilities, however there were no interactions at multiple correction-adjusted P<0.05, against our hypotheses.

Conclusions: Our results do not provide support for the idea that e4 genotype increases vulnerability to the negative effects of lifestyle risk factors on cognitive ability, but rather support a primarily outright association between *APOE* e4 genotype and worse cognitive ability.

Introduction

There is some evidence that associations between known lifestyle-based risk factors for worse cognitive abilities - e.g. diabetes[1], stress[2], traumatic brain injury[3], lower exercise[4], or air pollution[5], and female sex[6,7] – are larger in terms of effect size in people who possess an *APOE* e4 allele (vs. possessing non-risk e2 or e3 alleles). With regards dementia as an outcome, there are similar findings for physical activity, dietary fat, alcohol intake and smoking[8]. Essentially: people with the e4 allele may be more vulnerable to the effects of lifestyle risk factors on cognitive faculties. The potential biological rationale for this is that the *APOE* locus moderates lipid metabolism which influences brain-relevant factors like white matter myelination and neuronal repair; meaning e4 carriers may be more 'frail' and vulnerable to the negative effects of sub-optimal lifestyle risk factors[9,10]. There have been instances of null results, however[11]. It is also possible that there is a degree of 'file-drawer' where null results are less likely to be published[12]. There have been few large-scale systematic investigations into whether *APOE* e4 interacts with lifestyle risk factors associated with worse cognitive abilities, in a single cohort with a standard methodological procedure.

UK Biobank is a large general population cohort with approximately 502,000 participants[13]. All participants have baseline medical, cognitive and sociodemographic data, and genetic data. We hypothesised that there would be a significant statistical interaction where known lifestyle factors would have larger associations with cognitive abilities in people who possessed *APOE* e4 genotype (vs. non-e4).

Methodology

Study design and participants

The UK Biobank cohort is a large prospective general population cohort where baseline assessment took place between 2006 and 2010 in 22 assessment centres[13]. In total, 502,628 participants aged 40–70 years were recruited 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 comprehensive touch-screen questionnaire including sociodemographic characteristics, physical and mental health, and a brief battery of cognitive tests. Across 2014-2015, participants that had provided an email address were invited to complete a remote, web-based questionnaire including cognitive tests. The project was completed using application number 17689 (PI: Lyall).

Cognitive assessment

At baseline assessment participants completed five tests of cognitive ability, which were novel and computerised. We have described these in detail, in an open-access report[14]. For the current study, we focussed on the two tests that showed acceptable intra-participant stability across on average 4 years (intraclass *r* range = 0.54 to 0.65). In the first test, most participants completed a timed test of symbol matching, like the common card game 'Snap' hereafter referred to as reaction time (RT). The second test was a task with 13 logic/reasoning-type questions and a 2-min time limit, labelled as 'fluid intelligence' and referred to here simply as reasoning[15]. The maximum score is 13. The reasoning task was only added to the battery part way through the baseline assessment phase and so around n=~150k participants completed it.

We did not examine the baseline tests of pairs-matching, prospective memory or numeric memory. The pairs-matching task was markedly zero-inflated (indicating floor effect) and did not show good longitudinal stability in n=~20k with repeat data (r<0.2 across four years on average); prospective memory had around 94% overall success rate and thus had a degree of ceiling effect, and numeric memory was only completed by around n=48k overall and did not have longitudinal data to suggest good reliability. These considerations have been described previously[14].

After baseline assessment (2006-2010), between 2014 and 2015 participants were invited to complete a web-based questionnaire, where responders completed, amongst other things, web-based versions two well-known cognitive tasks called 'Trail making test A/B' (TMT-A and TMT-B; processing speed and speed/executive function respectively) and 'Digit symbol substitution' (executive function), each sensitive to the effects of cognitive ageing[16,17]. Independent studies have shown good correlation between computerized vs. paper-and-pen versions of the tests[18,19].

Sociodemographic and medical data

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 diabetes. We defined coronary heart disease (CHD) as either myocardial infarction or angina. We excluded participants who stated only 'prefer not to answer'.

Participants reported their highest educational attainment and this was recoded into a simpler college/university degree vs. no degree variable. Townsend deprivation indices were derived from postcode of residence[20]. 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.

Physical activity was self-reported and weighted for intensity: self-reported minutes of walking (×3.3), moderate exercise (×4.0) and vigorous exercise (×8.0; this is a common calculation[21]). These were then summated to create an overall physical activity score, which was then split into quintiles to simplify analysis.

Participants whose BMI was 40 or over were considered very severely obese as per World Health Organisation (WHO) guidelines; we chose a cut-off of 40 rather than say 30 ('moderately obese') because there is evidence of reverse causality where moderately high BMI can show a protective effect under some circumstances[22]. (Note that final results were virtually identical when we used a BMI of 30 as a cut-off).

In terms of smoking we compared 'never' vs. 'current' smokers. Frequency of alcohol intake was recorded as never, special occasions only, 1–3 times per month, 1–2 times per week, 3–4 times per week, daily/almost daily. Because our interest is in high vs. low alcohol intake we split this into a binary variable: participants who reported 'Daily or almost daily' (i.e. high) vs. 'One to three times a month'; 'Special occasions only' and 'Never' (i.e. low). Participants were asked if there was a reason they had stopped drinking, e.g. due to doctor's advice, health precaution etc.: participants who reported

this were removed from analysis, to help reduce confounding where low alcohol intake was due to poor health.

Genetic data

UK Biobank genotyping was conducted by Affymetrix using a bespoke BiLEVE Axiom array for ~50,000 participants and the remaining ~450,000 on the Affymetrix UK Biobank Axiom array. All genetic data were quality controlled by UK Biobank as described by the protocol paper[23]. 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 (https://biobank.ctsu.ox.ac.uk/crystal/docs/genotyping_sample_workflow.pdf). The two *APOE* e SNPs – rs7412 and rs429358 – were both in Hardy Weinberg equilibrium (P>0.05) assessed with PLINK V1.90[24].

Standard Protocol Approvals, Registrations, and Patient Consents

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).

Data availability statement

UK Biobank is an open access resource available to verified researchers upon application (http://www.ukbiobank.ac.uk/). Analysis syntax is available upon request.

Statistical analysis

We used two models: partially adjusted and fully adjusted. The partially adjusted model was statistically corrected for the potential confounders of: age, sex, genotypic array, assessment centre and eight genetic principal components (PCs; to correct for potential stratification). The fully adjusted model was additionally corrected for Townsend deprivation scores, self-reported diabetes, CHD, hypertension, and university/college degree ('yes' vs. 'no')[1]. We report descriptive statistics according to EQUATOR guidelines. The dependent variables in the linear regression were the cognitive scores for reasoning, log RT, log TMT A and B, and Digit symbol scores.

We first tested for associations between *APOE* e4 and lifestyle factors on cognitive abilities, using linear regression and reporting standardized betas (i.e. on a per-SD scale of effect). We then tested for two-way interactions between *APOE* e4 genotype with male vs. female sex, and e4 with lifestyle factors. Finally, we tested for additional three-way interactions (*APOE*; sex; lifestyle). TMT and reaction time scores were log-transformed due to a positive skew. We removed outliers above 3.30 SDs from the mean (<0.1%). We corrected for multiple testing using the False Discovery Rate (FDR)[25,26]. Power calculations were performed using G*Power 3[27]. Stata V.14 was used for statistical analyses. For additional comparison with previous meta-analyses, we have provided Cohen's *d* effect size estimates for unadjusted *APOE* e4/cognitive associations. All supplementary tables and figures are available from Dryad.

Results

Descriptives

There were 487,377 participants with *APOE* e genotype data. We excluded participants with non-white British ancestry, self-report vs. genetic sex mismatch, putative sex chromosomal aneuploidy, excess heterozygosity, and missingness rate >0.1. This left n=408,228. We removed participants who reported a neurological condition (~5%; see Lyall et al.[14]); the inclusion of which could drive type-1 errors due to skewed results (results were unchanged when we included these participants). This left 389,778 participants. Finally, we accounted for relatedness between participants by removing one random participant in cases where two individuals were 1st cousins or closer. This left 326,535 participants for whom genotype frequencies of *APOE* were e2/e2 n=2,133 (1%), e2/e3 n=40,460 (12%), e2/e4=8,348 (3%), e3/e3=189,728 (58.0%), e3/e4 n=77,963 (24%) and e4/e4 n=7,923 (2%). Descriptive statistics for cognitive scores and cardiometabolic conditions are shown in Tables 1 and 2, and demographic factors are show in Supplementary table e-1.

The mean age at baseline was 56.79 (standard deviation [SD] = 8.00), and 150,071 (46%) participants were male. The mean age at time of completing the internet tests was 61.8 years (SD=7.60). Using an *APOE* e4 present vs. absent model excluding e2/e4 (protective/risk alleles) genotype carriers, results in sample sizes per group of: e4+ n=85,886 (e3/e4; e4e4) vs. e4- n=232,301 (e2/e2; e2/e3; e3/e3), total n=318,187. In terms of cognitive data: reasoning data were available in n=105,913, reaction time in n=324,725, TMT A (processing speed) in n=70,988 and B (speed plus executive function) in n=71,055, with Digit symbol substitution (executive function) in n=79,840. All

significant phenotypic/genetic associations with cognitive abilities reported hereafter remained significant after correction for type-1 error.

A power calculation showed that based on a Cohen's D of 0.1 (a 'small' effect size being 0.2) and group difference ratio of 2:1 (based arbitrarily on never vs. current smoker ratio), 95% power to detect an effect would be achieved at n=4,872, suggesting the current analyses have generally good power.

[Table 1 here]

APOE e4 and lifestyle associations with cognitive abilities

Table 3 shows standardised beta associations between *APOE* e4 genotype, lifestyle factors, and cognitive abilities: there were significant associations between e4 genotype and worse log TMT-A times (fully adjusted model standardised beta = 0.032, 95% CI = 0.016 to 0.048, P<0.001), TMT-B times (fully adjusted standardised beta = 0.047, 95% CI = 0.032 to 0.062, P <0.001) and Digit symbol substitution scores (fully adjusted standardised beta = -0.054, 95% CI = -0.068 to -0.040, P<0.001).

Unadjusted *APOE* e4/cognitive score associations were of very small magnitude (i.e. under 0.2) for each of log RT (Cohen's d = 0.003), reasoning (d = -0.003), log TMT A (d = -0.014), log TMT B (-0.023), and Digit symbol coding (d = 0.035). Effect sizes were similar for untransformed RT and TMT A/B values.

In terms of lifestyle factors: there were significant associations for smoking with reasoning, TMT-A and -B times and Digit symbol substitution scores (all P<0.001; Table 1). There were significant associations for alcohol intake and obesity, but the sign of these

associations changed for alcohol and obesity where they appeared protective in the fully adjusted models for various tests. Physical activity did not significantly associate with any cognitive outcomes. When all analyses were corrected for type-1 error with FDR, all significant associations remained statistically significant (FDR-adjusted P-values all<0.05).

[Table 2 here]

Two-way interactions: APOE e4 and sex; APOE e4 and lifestyle.

We tested for *APOE* e4 by sex interactions, with the results shown in Supplementary Table e-2. There were two significant interactions: for log RT (fully-adjusted model P=0.045), and fluid reasoning (P=0.034). Stratifying by sex using the fully-adjusted models showed that the e4 effect was stronger in males vs. females for log RT (P = 0.068 vs. 0.375 respectively) although still non-significant; and not appreciably different for fluid reasoning scores (P = 0.155 vs. 0.136). For Digit symbol substitution there was a significant interaction between e4 and obesity (final model P value <0.001). Stratified, this appeared to be due to a significantly deleterious effect of e4 genotype in non-obese participants (fully-adjusted standardized beta = -0.058, 95% CI = -0.072 to -0.044, P <0.001), but protective in obese participants (fully-adjusted standardized beta = 0.176, 95% CI = 0.058 to 0.295, P = 0.004). All other tested two-way interactions were not significant (P>0.05).

[Table 3 here]

Three-way interactions: APOE e4, sex, and lifestyle.

We tested for significant *APOE* e4/sex/lifestyle interactions, with the results shown in Supplementary Table e-3. All interactions were non-significant except one. The significant interaction was for e4 presence, sex and high alcohol intake (i.e. daily or almost daily) vs. not on reasoning scores (P=0.020). Supplementary Figure e-1 shows that the interaction was principally driven by males having a larger association between high alcohol intake and better reasoning (compared with females). While visually an e4 effect becomes slightly larger in the context of high alcohol intake, pairwise comparisons did not show this to be statistically significant (P>0.05). When all analyses were corrected for type-1 error with FDR, all significant interactions attenuated to non-significance (FDR-adjusted P-values all >0.05). The total model adjusted r^2 values ranged from 0.02 to 0.22 (i.e. 2% to 22% of total variance explained).

Additional analyses

As post-hoc analyses we additionally repeated all tests for collated (potentially protective) *APOE* e2/e2 plus e2/e3 genotypes, vs. neutral e3/e3. We also repeated the analyses with log-transformed (+1) pairs-matching error scores as an outcome. There were no significant associations or interactions once adjusted for FDR (all *q*-values P>0.100; results are available upon request).

It is possible that e4 genotype and lifestyle are not independent. Logistic regressions showed that participants who possessed the e4 allele were significantly less likely to smoke (OR = 0.95, 95% CIs = 0.93 to 0.98, P<0.001) and more likely to have a degree (OR = 1.02, 95% CIs = 1.00 to 1.03, P = 0.043) although the effect sizes were small, and carriers showed no differences in other lifestyle factors (see Supplementary Table e-4, which shows all intercorrelations).

The protective effect of alcohol intake on cognitive ability is counter-intuitive, having removed people who reported stopping due to ill health. Descriptive statistics of alcohol intake by *APOE* e4 genotype status are shown in Supplementary Table e-5.

Discussion

This study hypothesized that based on previous studies in smaller cohorts, together with biological rationale, risk factors for worse cognitive ability such as smoking history, (high) alcohol intake, obesity, and lower physical activity, would interact with *APOE* e4 genotype, such that each risk factor's association with worse cognitive scores would be larger in e4 carriers (vs. non-carriers). We also investigated the moderating role of sex[28]. We found that associations between *APOE* e4 and cognitive scores were of relatively small effect size, and only suggestive interactions with sex where e4 males scored worse than females (which did not survive correction for multiple testing; and in any case the within-sex e4 effects were not nominally significant). We also found some small, counterintuitive suggestive results e.g. that severe obesity and daily drinking could be protective. These findings could reflect: test imprecision, the generally preserved and healthy sample (i.e. selection or attrition biases), underestimation of e4's true effect (due to attrition), or that previous studies perhaps overstated the true effect. Our findings generally support a 'direct' route of *APOE* e4 genotype to cognitive decline rather than increasing vulnerability to other factors.

In this study we report negative associations between smoking and worse cognitive ability, which fits the established literature[29]; although surprisingly protective associations from high alcohol intake (i.e. daily) and obesity defined here as BMI of 40 and above (aka severely obese), even after adjusting for prevalent diseases and

accounting as much as possible for people whose alcohol intake had significantly changed in recent years due to ill health (i.e. factors which might cause reverse causality). This is more likely to reflect selection or collider bias in some way [30]: e.g. where the participants who drink more/are highly obese and respond positively to the invitation for assessment, are quite selected[31], rather than the association being causal. This is also the most likely explanation for e4 carriers having better scores (vs. non-carriers) in the context of severe obesity in this study. In any case the interactions were null after correction for type-1 error with FDR. There was no association from weighted physical activity, although the sample size for that variable was much smaller than others. There were significant associations between *APOE* e4 genotype and worse TMT-A, TMT-B, and Digit symbol substitution scores which fits previous literature that e4 genotype is deleterious for processing speed and executive function[32].

There were mostly no statistically significant interactions between lifestyle factors and *APOE* e4 genotype. The e4/cognitive associations were of quite small magnitude, compared to previous meta-analyses[33]. Power analysis estimates showed that we had relatively good power to detect an association; although it is still possible that the lack of association reflects a lack of power. Alternative interpretations include that that the UK Biobank participants have perhaps not deteriorated markedly with age or are in generally good health, and/or are slightly too young (mean age 56 at baseline) to show significant effects of *APOE* e4 genotype, which can show a larger association with cognitive function with increasing age[34] or longitudinally[32]. Further to this there may be sex effects which vary by age window: for example Neu et al.[35] found that *APOE* e3/e4 genotype was associated with earlier age at onset of Alzheimer's disease (AD) (vs. men; total N = 57,979), and Hohman et al.[36] reported significant interaction between e4

presence (vs. absence) and female (vs. male) sex on higher total cerebrospinal total and phosphorylated tau (a neuropathological marker of AD). Additional interactions which we did not assess are also possible, e.g. between *APOE* e4, sex and deprivation level, and this will be an interesting area of future research.

It is possible that the lack of interaction reflects a degree of selection bias where the sample includes 'healthier' carriers of the e4 genotype (generally reported as deleterious), and its effect in this cohort is therefore underestimated to an extent.

Our results slightly contrast with our previous findings in around n=110k UK Biobank participants, where we reported a significant deleterious interaction between e4 genotype and reasoning scores ($P<0.001$), however this (and all other tests) did not survive correction for additional covariates e.g. depression, Townsend scores, and cardiometabolic conditions in that study.

We have reported previously on potential limitations of the novel baseline tests: namely that the reasoning test includes some 'crystallized' (i.e. accumulated knowledge) items which are not strictly reasoning, and the reliabilities are poorer across time compared with more standard, validated cognitive tests[14]. We did not report on UK Biobank memory scores because our previous analysis has shown that a) the test was not reliable across time[14] and b) e4 had no major association with scores in n=110k anyway[1]. The web-based tests are more akin to existing validated cognitive batteries, but their use over the internet in this instance has not been characterised and there may be some inaccuracies due to internet connection lag etc., or computer problems in people's homes. It is possible that the interaction between e4 genotype and lifestyle risk factors has been overstated due to publication bias, particularly given many studies are quite small in terms of sample size[37]. On the other hand, the large sample size used here may increase

risk of statistically significant findings which are of such small magnitude as to not be practically or clinically significant.

The UK Biobank does not have a metric of premorbid, lifetime cognitive ability in its participants. This could be an important limitation where 'brighter' young adults are less likely to engage in unhealthy behaviours, or in midlife, people with better cognitive ability may be better able to manage their healthcare, take medications reliably etc. [38].

Genetic modification of phenotypic risk factors on cognitive ability has enormous potential implication for prevention of cognitive impairment in an ageing population. Future research may seek to investigate this question in brain imaging phenotypes (available in UK Biobank although in smaller numbers), as these factors are less 'downstream' of the effects of genetic variation compared with cognitive scores, which can be affected by state-dependent factors like stress or anxiety[39].

This study aimed to test for interactions between *APOE* e4, lifestyle and sex on cognitive abilities. We found suggestive interaction test results where men were more vulnerable to e4 genotype (in terms of cognition). Caveats to this were that the effect sizes were small, and there may be biases at play (e.g. where e4's effects are underestimated in the data). Our results therefore provide less support for the idea that e4 genotype increases vulnerability to the negative effects of lifestyle risk factors, but rather support a primarily outright association between *APOE* e4 genotype and worse cognitive ability.

Acknowledgements: This research has been conducted using the UK Biobank resource; we are grateful to UK Biobank participants. Thanks to Dr. Breda Cullen for devising exclusion criteria.

References

1. Lyall DM, Ward J, Ritchie SJ, et al. Alzheimer disease genetic risk factor APOE e4 and cognitive abilities in 111,739 UK Biobank participants. Age Ageing. doi: 10.1093/ageing/afw068

2. Lyons MJ, Genderson M, Grant MD, et al. Gene-environment interaction of ApoE genotype and combat exposure on PTSD. Am J Med Genet Part B Neuropsychiatr Genet. 2013;162:762–769.

3. Li L, Bao Y, He S, et al. The Association Between Apolipoprotein E and Functional Outcome After Traumatic Brain Injury. Medicine. 2015;94:e2028.

4. Head D, Bugg JM, Goate AM, et al. Exercise Engagement as a Moderator of the Effects of APOE Genotype on Amyloid Deposition. Arch Neurol. NIH Public Access; 2012;69:636–643.

5. Cacciottolo M, Wang X, Driscoll I, et al. Particulate air pollutants, APOE alleles and their contributions to cognitive impairment in older women and to amyloidogenesis in experimental models. Transl Psychiatry. Nature Publishing Group; 2017;7:e1022.

6. Davies G, Harris SE, Reynolds CA, et al. A genome-wide association study implicates the APOE locus in nonpathological cognitive ageing. Mol Psychiatry. 2014;19:76–87.

7. Moser VA, Pike CJ. Obesity and sex interact in the regulation of Alzheimer's disease. Neurosci Biobehav Rev. NIH Public Access; 2016;67:102–118.

8. Kivipelto M, Rovio S, Ngandu T, et al. Apolipoprotein E epsilon4 magnifies lifestyle

risks for dementia: a population-based study. J Cell Mol Med. 2008;12:2762–2771.

9. Liu C-C, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol. Nature Publishing Group; 2013;9:106–118.

10. Holtzman DM, Herz J, Bu G. Apolipoprotein E and apolipoprotein E receptors: normal biology and roles in Alzheimer disease. Cold Spring Harb Perspect Med. 2012;2:a006312.

11. Rodriguez FS, Schroeter ML, Arélin K, et al. APOE e4-Genotype and Lifestyle Interaction on Cognitive Performance: Results of the LIFE-Adult-Study. Heal Psychol. Epub 2017 Dec 7.

12. Ioannidis JPA, Munafò MR, Fusar-Poli P, Nosek BA, David SP. Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. Trends Cogn Sci. NIH Public Access; 2014;18:235–241.

13. Sudlow C, Gallacher J, Allen N, 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. Public Library of Science; 2015;12:e1001779.

14. Lyall DM, Cullen B, Allerhand M, et al. Cognitive test scores in UK biobank: Data reduction in 480,416 participants and longitudinal stability in 20,346 participants. PLoS One. 2016;11.

15. Davies G, Marioni REE, Liewald DCC, et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). Mol Psychiatry. Macmillan Publishers Limited; 2016;21:758–767.

16. Salthouse TA. What cognitive abilities are involved in trail-making performance?

453 Intelligence. NIH Public Access; 2011;39:222–232.

454 17. Salthouse TA. What Do Adult Age Differences in the Digit Symbol Substitution Test
455 Reflect ? J Gerontol Psychol Sci. 1992;47:21–128.

456 18. Fellows RP, Dahmen J, Cook D, Schmitter-Edgecombe M. Multicomponent analysis
457 of a digital Trail Making Test. Clin Neuropsychol. NIH Public Access; 2017;31:154–
458 167.

459 19. Gualtieri CT, Johnson LG. Reliability and validity of a computerized neurocognitive
460 test battery, CNS Vital Signs. Arch Clin Neuropsychol. 2006;21:623–643.

461 20. Townsend P. Townsend deprivation index. Natl. database Prim. care groups Trust.
462 1998.

463 21. Siebert S, Lyall DM, Mackay DF, et al. Characteristics of rheumatoid arthritis and its
464 association with major comorbid conditions: cross-sectional study of 502 649 UK
465 Biobank participants. RMD Open. 2016;2:e000267.

466 22. Kivimäki M, Luukkonen R, Batty GD, et al. Body mass index and risk of dementia:
467 Analysis of individual-level data from 1.3 million individuals. Alzheimer's Dement.
468 Elsevier; Epub 2017 Nov 21.

469 23. Bycroft C, Freeman C, Petkova D, et al. Genome-wide genetic data on ~500,000 UK
470 Biobank participants. doi.org. Cold Spring Harbor Laboratory; Epub 2017 Jul
471 20.:166298.

472 24. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome
473 association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–
474 575.

475 25. Pike N. Using false discovery rates for multiple comparisons in ecology and
476 evolution. Methods Ecol Evol. 2011;2:278–282.

477 26. Benjamini Y, Hochberg Y. Benjamini Y, Hochberg Y. Controlling the false discovery
478 rate: a practical and powerful approach to multiple testing. J R Stat Soc B.
479 1995;57:289–300.

480 27. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: A flexible statistical power
481 analysis program for the social, behavioral, and biomedical sciences. Behav Res
482 Methods. Springer-Verlag; 2007;39:175–191.

483 28. Nebel RA, Aggarwal NT, Barnes LL, et al. Understanding the impact of sex and
484 gender in Alzheimer's disease: A call to action. Alzheimers Dement. Elsevier;
485 2018;14:1171–1183.

486 30. Day FR, Loh P-R, Scott RA, Ong KK, Perry JRB. A Robust Example of Collider Bias in
487 a Genetic Association Study. Am J Hum Genet. Elsevier; 2016;98:392–393.

488 31. Millard LAC, Davies NM, Tilling K, Gaunt TR, Smith GD. Searching for the causal
489 effects of BMI in over 300 000 individuals, using Mendelian randomization. Epub
490 2017; bioRxiv; 236182.

491 32. Lyall DM, Harris SE, Bastin ME, et al. Are APOE ε genotype and TOMM40 poly-T
492 repeat length associations with cognitive ageing mediated by brain white matter
493 tract integrity? Transl Psychiatry. 2014;4:e449.

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

496 34. Schiepers OJG, Harris SE, Gow AJ, et al. APOE E4 status predicts age-related
497 cognitive decline in the ninth decade: longitudinal follow-up of the Lothian Birth

498 Cohort 1921. Mol Psychiatry. Macmillan Publishers Limited; 2011;17:315–324.

499 35. Neu SC, Pa J, Kukull W, et al. Apolipoprotein E Genotype and Sex Risk Factors for
500 Alzheimer Disease. JAMA Neurol. 2017;74:1178.

501 36. Hohman TJ, Dumitrescu L, Barnes LL, et al. Sex-Specific Association of
502 Apolipoprotein E With Cerebrospinal Fluid Levels of Tau. JAMA Neurol.
503 2018;75:989.

504 37. Munafò MR, Stothart G, Flint J. Bias in genetic association studies and impact factor.
505 Mol Psychiatry. 2009;14:119–120.

506 38. Calvin CM, Deary IJ, Fenton C, et al. Intelligence in youth and all-cause-mortality:
507 Systematic review with meta-analysis. Int J Epidemiol. 2011;40:626–644.

508 39. Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences.
509 Nat Rev Neurosci. 2010;11.

510

Table 1: demographic descriptive statistics.

		***APOE* e4 absent (n=232,301; 73%)**	***APOE* e4 present (n=85,886; 27%)**
Age in years	Mean (SD)	56.82 (8.00)	56.71 (8.00)
Sex	Male N (%)	106,694 (46%)	49,491 (46%)
Townsend deprivation score	Mean (SD)	-1.59 (2.92)	-1.60 (2.92)
Alcohol intake, N (%)	≤3 times per month	56,819 (53.08)	20,792 (52.94)
	Daily	50,219 (46.92)	18,484 (47.06)
	Missing	2,054	1,848
Current smoker, N (%)	Current	23,366 (15.52)	8,237 (14.88)
	Never	127,185 (84.48)	47,134 (85.12)
	Missing		
Total physical activity quintile	1st	1,634 (20%)	610 (21%)
	2nd	1,590 (20%)	560 (20%)
	3rd	1,651 (20%)	547 (19%)
	4th	1,648 (20%)	587 (20%)
	5th	1,596 (20%)	559 (20%)
	Missing	224,209	83,023
Severely obese (BMI≥40), N (%)	No	227,454 (97.91)	84,127 (97.95)
	Yes	4,847 (2.09)	1,759 (2.05)
	Missing	8,185	163
Degree, N (%)	Yes	73,820 (32%)	27,616 (32%)
	No	156,602 (68%)	57,567 (68%)
	Missing	1,879	703

Table 2: cognitive score descriptive statistics.

	***APOE* e4 absent**	***APOE* e4 present**
Reasoning scores, mean (SD)	6.20 (2.10)	6.21 (2.10)
Log transformed reaction time score, mean (SD)	6.30 (0.18)	6.30 (0.18)
Untransformed median (IQR)	535 (477-606)	535 (477-605)
Digit symbol substitution scores, mean (SD)	19.87 (5.14)	19.69 (5.26)
Log transformed Trail making test-A times, mean (SD)	3.60 (0.30)	3.60 (0.31)
Untransformed median (IQR)	35.33 (29.03 to 44.29)	35.51 (29.10 to 44.59)
Log transformed Trail making test-B times, mean (SD)	4.12 (0.32)	4.12 (0.32)
Untransformed median (IQR)	60.32 (49.07 to 75.48)	60.80 (49.33 to 76.17)

Table 3: individual associations between *APOE* e4, lifestyle and cognitive phenotypes.

	Partially adjusted	95% CI's			Fully adjusted	95% CI's		
	Standardised b	lower	upper	p	Standardised b	lower	upper	p
Log reaction time								
APOE e4	0.002	-0.006	0.009	0.678	0.002	-0.005	0.010	0.555
Smoking	**0.118**	**0.106**	**0.129**	**<0.001**	**0.070**	**0.058**	**0.082**	**<0.001**
Alcohol	**-0.108**	**-0.118**	**-0.098**	**<0.001**	**-0.075**	**-0.085**	**-0.065**	**<0.001**
Obesity	**0.082**	**0.060**	**0.105**	**<0.001**	0.022	-0.001	0.045	0.064
Physical activity	-0.004	-0.016	0.008	0.557	-0.005	-0.017	0.007	0.438
Fluid reasoning scores								
APOE e4	0.003	-0.011	0.017	0.673	<0.001	-0.013	0.013	0.964
Smoking	**-0.236**	**-0.257**	**-0.214**	**<0.001**	**-0.084**	**-0.104**	**-0.063**	**<0.001**
Alcohol	**0.289**	**0.271**	**0.307**	**<0.001**	**0.169**	**0.152**	**0.187**	**<0.001**
Obesity	**-0.137**	**-0.178**	**-0.095**	**<0.001**	-0.019	-0.058	0.021	0.355
Physical activity	0.004	-0.019	0.027	0.741	0.011	-0.011	0.033	0.311
Log TMT-A times								
APOE e4	**0.031**	**0.015**	**0.047**	**<0.001**	**0.032**	**0.016**	**0.048**	**<0.001**

Smoking	**0.084**	**0.056**	**0.113**	**<0.001**	**0.043**	**0.014**	**0.072**	**0.003**
Alcohol	**-0.081**	**-0.102**	**-0.059**	**<0.001**	**-0.051**	**-0.072**	**-0.029**	**<0.001**
Obesity	0.020	-0.036	0.076	0.487	-0.033	-0.089	0.023	0.249
Physical activity	-0.009	-0.032	0.014	0.453	-0.012	-0.035	0.012	0.323
Log TMT-B times								
APOE e4	**0.044**	**0.028**	**0.059**	**<0.001**	**0.047**	**0.032**	**0.062**	**<0.001**
Smoking	**0.197**	**0.170**	**0.225**	**<0.001**	**0.133**	**0.106**	**0.161**	**<0.001**
Alcohol	**-0.093**	**-0.114**	**-0.072**	**<0.001**	**-0.039**	**-0.060**	**-0.018**	**<0.001**
Obesity	**0.081**	**0.027**	**0.136**	**0.003**	-0.005	-0.059	0.049	0.857
Physical activity	-0.005	-0.028	0.018	0.672	-0.007	-0.030	0.015	0.524
Digit symbol scores								
APOE e4	**-0.054**	**-0.068**	**-0.040**	**<0.001**	**-0.054**	**-0.068**	**-0.040**	**<0.001**
Smoking	**-0.151**	**-0.177**	**-0.126**	**<0.001**	**-0.091**	**-0.117**	**-0.066**	**<0.001**
Alcohol	**0.115**	**0.095**	**0.134**	**<0.001**	**0.069**	**0.049**	**0.088**	**<0.001**
Obesity	**-0.117**	**-0.167**	**-0.067**	**<0.001**	-0.044	-0.094	0.006	0.085
Physical activity	0.011	-0.010	0.032	0.302	0.013	-0.008	0.033	0.222

Partially adjusted: age, sex, assessment centre, genotypic array. Fully adjusted: (also) Townsend deprivation scores, degree yes vs. no, self-report diabetes, hypertension and CHD.

Supplementary Table e-1: cardiometabolic condition frequencies.

		***APOE* e4 absent**	***APOE* e4 present**
Diabetes	Yes	11,081 (5%)	3,756 (4%)
	No	220,729 (95%)	81,962 (96%)
	Missing	491	168
Hypertension	Yes	61,426 (26%)	22,814 (27%)
	No	170,529 (74%)	62,939 (73%)
	Missing	346	133
Coronary heart disease	Yes	9,492 (4%)	4,001 (5%)
	No	81,752 (96%)	81,752 (95%)
	Missing	346	133

Supplementary Table e-2: individual two-way interactions between *APOE* e4 genotype and variables, on cognitive phenotypes.

	Partially adjusted				Fully adjusted			
	Standardised b	lower	upper	p	Standardised b	lower	upper	p
Log reaction time								
APOE e4*sex	0.02	<0.01	0.03	0.043	0.02	<0.01	0.03	0.045
APOE e4*smoking	<0.01	-0.03	0.02	0.717	<0.01	-0.03	0.02	0.832
APOE e4*alcohol	0.01	-0.01	0.03	0.415	0.01	-0.01	0.03	0.522
APOE e4*obesity	-0.04	-0.09	0.01	0.140	-0.03	-0.09	0.02	0.204
APOE e4*physical activity	<0.01	-0.03	0.03	0.911	<0.01	-0.03	0.03	0.943
Fluid reasoning scores								
APOE e4*sex	0.02	<0.01	0.05	0.082	0.03	<0.01	0.05	0.034
APOE e4*smoking	-0.03	-0.08	0.02	0.255	-0.03	-0.08	0.02	0.187
APOE e4*alcohol	0.02	-0.02	0.06	0.224	0.03	-0.01	0.06	0.162
APOE e4*obesity	0.04	-0.06	0.13	0.459	0.04	-0.05	0.13	0.437
APOE e4*physical activity	0.04	-0.01	0.09	0.148	0.02	-0.03	0.07	0.425
Log TMT-A times								
APOE e4*sex	<0.01	-0.03	0.03	0.985	<0.01	-0.03	0.03	0.977
APOE e4*smoking	-0.01	-0.07	0.06	0.877	<0.01	-0.06	0.07	0.989
APOE e4*alcohol	-0.03	-0.08	0.01	0.166	-0.03	-0.08	0.01	0.151

APOE e4*obesity	<0.01	-0.13	0.12	0.945	<0.01	-0.13	0.13	0.952
APOE e4*physical activity	0.03	-0.03	0.08	0.312	0.02	-0.03	0.08	0.368
Log TMT-B times								
APOE e4*sex	0.01	-0.02	0.04	0.581	0.01	-0.02	0.04	0.540
APOE e4*smoking	-0.01	-0.08	0.05	0.701	<0.01	-0.06	0.06	0.998
APOE e4*alcohol	-0.04	-0.08	0.01	0.109	-0.04	-0.09	0.00	0.075
APOE e4*obesity	-0.01	-0.14	0.11	0.829	-0.01	-0.14	0.11	0.860
APOE e4*physical activity	0.03	-0.03	0.08	0.343	0.02	-0.03	0.07	0.452
Digit symbol scores								
APOE e4*sex	-0.01	-0.04	0.02	0.556	-0.01	-0.04	0.02	0.544
APOE e4*smoking	0.05	-0.01	0.11	0.076	0.04	-0.01	0.10	0.135
APOE e4*alcohol	-0.02	-0.06	0.02	0.324	-0.02	-0.06	0.02	0.370
APOE e4*obesity	0.23	0.11	0.35	<0.001	0.22	0.11	0.34	<0.001
APOE e4*physical activity	0.02	-0.03	0.06	0.505	0.02	-0.03	0.07	0.390

Partially adjusted: age, sex, assessment centre, genotypic array. Fully adjusted: (also) Townsend deprivation scores, degree yes vs. no, self-report diabetes, hypertension and CHD. Each dependent variable (cognitive score) is underlined in the left-hand column.

Supplementary Table e-3: individual three-way interactions between *APOE* e4 genotype, sex, and lifestyle variables, on cognitive phenotypes.

	Partially adjusted				Fully adjusted			
	Standardised b	lower	upper	p	Standardised b	lower	upper	p
Log reaction time								
APOE *e4**sex*smoking	-0.02	-0.06	0.03	0.495	-0.02	-0.06	0.03	0.489
APOE e4*sex*alcohol	0.03	<0.01	0.07	0.039	0.03	<0.01	0.07	0.043
APOE e4*sex*obesity	-0.01	-0.12	0.10	0.845	-0.02	-0.13	0.09	0.726
APOE e4*sex*physical activity	<0.01	-0.02	0.03	0.715	<0.01	-0.02	0.03	0.810
Fluid reasoning scores								
APOE e4*sex*smoking	0.04	-0.05	0.13	0.430	0.04	-0.04	0.13	0.342
APOE e4*sex*alcohol	0.06	<0.01	0.12	0.035	0.07	0.01	0.12	0.017
APOE e4*sex*obesity	0.06	-0.13	0.26	0.521	0.08	-0.11	0.27	0.407
APOE e4*sex*physical activity	<0.01	-0.05	0.04	0.861	0.01	-0.04	0.05	0.799
Log TMT-A times								
APOE e4*sex*smoking	0.01	-0.12	0.13	0.931	0.01	-0.12	0.13	0.923
APOE e4*sex*alcohol	0.01	-0.05	0.08	0.650	0.02	-0.05	0.08	0.591
APOE e4*sex*obesity	-0.04	-0.31	0.24	0.790	-0.04	-0.31	0.23	0.761
APOE e4*sex*physical activity	0.02	-0.03	0.07	0.425	0.02	-0.03	0.06	0.459

Log TMT-B times								
APOE e4*sex*smoking	0.09	-0.03	0.21	0.152	0.08	-0.03	0.20	0.164
APOE e4*sex*alcohol	-0.01	-0.07	0.06	0.842	-0.01	-0.07	0.06	0.870
APOE e4*sex*obesity	-0.02	-0.29	0.25	0.871	-0.01	-0.27	0.26	0.960
APOE e4*sex*physical activity	0.03	-0.02	0.07	0.231	0.02	-0.02	0.07	0.295
Digit symbol scores								
APOE e4*sex*smoking	-0.01	-0.12	0.10	0.815	-0.03	-0.13	0.08	0.652
APOE e4*sex*alcohol	0.02	-0.04	0.07	0.597	0.01	-0.04	0.07	0.608
APOE e4*sex*obesity	-0.09	-0.34	0.15	0.462	-0.11	-0.36	0.13	0.366
APOE e4*sex*physical activity	<0.01	-0.04	0.05	0.860	<0.01	-0.04	0.05	0.860

Partially adjusted: age, sex, assessment centre, genotypic array. Fully adjusted: (also) Townsend deprivation scores, degree yes vs. no, self-report diabetes, hypertension and CHD. Each three-way interaction includes the two-way interactions plus main effects in the model(s). Each dependent variable (cognitive score) is underlined in the left-hand column.

Supplementary Table e-4 – lifestyle and *APOE* e4 intercorrelations.

Odds ratio (CI's)	Smoking	Alcohol	Obesity	Physical activity	Degree	*APOE* e4
Smoking status (never/previous vs. never)	-					
Alcohol (heavy vs. not)	1.88 (1.81-1.94) **	-				
Obesity (obese vs. not)	1.00 (0.92-1.09)	0.22 (0.21-0.25) **	-			
Physical activity quintile (1-5; ordinal).	0.96 (0.88-1.06)	0.97 (0.89-1.05)	1.09 (0.09-1.33)	-		
Degree (college and above vs. not)	0.51 (0.49-0.52) **	2.19 (2.14-2.24) **	0.58 (0.55-0.62) **	0.89 (0.84=0.93) **	-	
APOE e4 allele presence (vs. absence)	0.95 (0.93 to 0.98) **	1.01 (0.98-1.03)	0.98 (0.93-1.04)	0.98 (0.93-1.04)	1.02 (1.00-1.03)*	-

Odds ratios reflect logistic regressions of a 1-unit change in the independent variable (y-axis) vs. dependent variable (x-axis). *P<0.05; **P<0.001.

Supplementary Table e-5 – alcohol intake by *APOE* e4 status.

		***APOE* e4 absent**		***APOE* e4 present**	
		Light drinker	**Heavy drinker**	**Light drinker**	**Heavy drinker**
Reasoning scores (mean; SD)		5.89 (2.07)	6.54 (2.09)	5.85 (2.07)	6.56 (2.08)
	N	18,773	16,708	6,608	6,145
Reaction time (msecs; median and interquartile range)		543 (485-617)	531 (470-601)	543 (484-614)	531 (477-598)
	N	56,389	50,038	20,652	18,424
Trail making test A (secs; median and interquartile range)		35.88 (29.20 – 45.18)	35.41 (29.33-43.97)	36.28 (29.54-46.18)	35.26 (29.31-43.54)
	N	10,656	12,483	3,859	4,465
Trail making test B (secs median and interquartile range)		61.11 (49.40 – 76.80)	61.05 (49.78-76.03)	62.00 (50.06 – 78.88)	60.96 (49.77 – 75.11)
	N	10,659	12,504	3,859	4,469
Digit symbol substitution score (mean; SD)		19.62 (5.26)	19.66 (4.96)	19.50 (5.39)	19.51 (5.12)
	N	12,434	14,297	4,476	5,152

SD = standard deviation. Scores are medians for reaction time and trail making test scores because they were not normally distributed. Reasoning and Digit symbol scores are means.

Supplementary Figure e-1: three-way *APOE* e4, alcohol intake and sex plot for reasoning scores (estimated marginal means based on fully-adjusted model; see 'analysis').