Check for updates

RESEARCH ARTICLE

Polygenic risk score for Alzheimer's disease and trajectories of cardiometabolic risk factors in children [version 1; peer review: 2 approved]

Roxanna Korologou-Linden[1,2], Linda O'Keeffe[1,2], Laura D. Howe [iD] [1,2], George Davey-Smith [iD] [1,2], Hannah J. Jones [iD] [1-3], Emma L. Anderson[1,2], Evie Stergiakouli [iD] [1,2,4]

[1]Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[2]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
[3]Centre for Academic Mental Health, Population Health Sciences, University of Bristol, Bristol, UK
[4]School of Oral and Dental Sciences, University of Bristol, Bristol, UK

Abstract

Introduction: Cardiometabolic factors are implicated in the aetiology of Alzheimer's disease and may lie on the pathways linking genetic variants to Alzheimer's disease across the life course. We examined whether polygenic risk scores (PRS) were associated with cardiometabolic health indicators through childhood and adolescence.

Methods: In 7,977 participants from the Avon Longitudinal Study of Parents and Children, we tested whether a PRS for Alzheimer's disease was associated with trajectories of cardiometabolic risk factors. We examined trajectories for height at 1-18 years; lean and fat mass at 9-18 years; systolic and diastolic blood pressure at 7-18 years; glucose and C-reactive protein at 9-18 years; insulin at 10-18 years; and high and low-density lipoproteins and triglycerides birth at 18 years. We also examined birthweight and interleukin-6 (IL-6) at age 9 years and physical activity at ages 11, 12, and 15 years.

Results: No consistent associations were observed between the PRS excluding genetic variants in the apolipoprotein E gene region and cardiometabolic factors trajectories across childhood and adolescence.

Conclusions: We did not detect evidence to suggest that the PRS for Alzheimer's disease acts through childhood and adolescent cardiometabolic risk factors. Further studies should examine whether these associations emerge later in adulthood when variation in cardiometabolic risk factors is likely to be greater.

Keywords

polygenic risk scores, Alzheimer's disease, trajectories, cardiometabolic

Open Peer Review

Reviewer Status ✔✔

	Invited Reviewers	
	1	**2**
version 1 20 Aug 2019	✔ report	✔ report

1 **Emily Baker** [iD] , Cardiff University, Cardiff, UK
Valentina Escott-Price [iD] , Cardiff University, Cardiff, UK

2 **Michelle Grunin**, Case Western Reserve University, Cleveland, USA

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Avon Longitudinal Study of Parents and Children (ALSPAC) gateway.

Corresponding author: Evie Stergiakouli (e.stergiakouli@bristol.ac.uk)

Author roles: Korologou-Linden R: Formal Analysis, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **O'Keeffe L**: Methodology, Supervision, Writing – Review & Editing; **Howe LD**: Conceptualization, Investigation, Supervision, Writing – Review & Editing; **Davey-Smith G**: Conceptualization, Funding Acquisition, Writing – Review & Editing; **Jones HJ**: Resources, Writing – Review & Editing; **Anderson EL**: Conceptualization, Investigation, Supervision, Writing – Review & Editing; **Stergiakouli E**: Conceptualization, Investigation, Supervision, Writing – Review & Editing

Introduction

There is growing evidence from observational studies for a role of midlife hypertension, obesity, hyperlipidaemia[1], hyperglycaemia[2], and inflammation[3,4] in the aetiology of Alzheimer's disease. Evidence for a role of cardiometabolic risk factors in later life is less clear as observed associations are often either null or protective[1,5,6]. This age-dependent change in the direction of the associations observed for cardiometabolic risk factors has been suggested to potentially reflect selection bias, as fatal cardiometabolic disease events would reduce the risk of being diagnosed with Alzheimer's disease, which usually occurs late in life[7]. Another possible explanation for these findings is reverse causation, due to the potential presence of undiagnosed Alzheimer's disease (which has been shown to be associated with a lowering of blood pressure[5] and weight loss[6]) at the time of assessment of the cardiometabolic risk factors.

Large genetic consortia have identified common genetic variants associated with late-onset Alzheimer's disease[8], with the ε4 allele of the apolipoprotein E (ApoE) gene conferring the greatest risk, increasing the risk up to twelvefold[9]. ApoE adversely affects lipid profiles[10] and elevates the risk of coronary artery disease[11]. Although the genetic variants (apart from ApoE4) individually increase risk by a small amount, they can be summarised to construct a polygenic risk score (PRS) for Alzheimer's disease which has a good ability to correctly classify those with and without the disease (area under the receiver operating characteristic curve [AUROC] = 78.2% with a P-value threshold of ≤0.5)[12]. Apart from ApoE, the relationship of these genetic variants to cardiovascular disease is uncertain.

PRS can be used to investigate potential prodromal phenotypes and or pathways through which the variants influence Alzheimer's disease risk. There are very few studies[13,14] examining the association between cardiometabolic risk factors in early life, likely due to the long follow-up required across the life course. We tested the hypothesis that a PRS for Alzheimer's disease (beyond the ApoE gene) acts through cardiometabolic pathways in children and adolescents.

Methods

Participants

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a prospective birth cohort which recruited pregnant women with expected delivery dates between April 1991 and December 1992 from Bristol UK[15,16]. A total of 14,541 pregnant mothers were initially enrolled, from which 13,867 live births occurred in 13,761 women. Including two later rounds of recruitment, there were 15,445 eligible children. Information on health and development of children and their parents were collected from regular clinic visits and completion of questionnaires. Research clinics were held when the participants were approximately seven, nine, 10, 11, 13, 15, and 18 years old. The study website contains details of all the data that is available through a fully searchable data dictionary: http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/. Ethical approval was obtained from the ALSPAC Law and Ethics Committee and the Local Ethics Committees.

Cardiometabolic measures

Anthropometry. Birthweight was obtained from obstetric records. At the clinics between 4 months and 5 years, crown-heel length was measured using a Harpenden Neonatometer and from 25 months onwards, standing height was measured using a Leicester Height Measure. From age 7 years, all children were invited to yearly clinics, where measurements of standing height were obtained using the Harpenden Stadiometer. At each clinic at ages nine, 11, 13, 15, and 18 years, whole body less head, and central fat and lean mass were derived from whole body dual energy X-ray absorptiometry (DXA) scans evaluated using a Lunar prodigy narrow fan beam densitometer.

Cardiometabolic and inflammatory factors. At each clinic (ages seven, nine, 10, 11, 12, 15, and 18), systolic and diastolic blood pressure were measured at least twice. The mean of two measurements was used in the analysis. Insulin was obtained from cord blood at birth. Non-fasting glucose was also measured at age seven as part of the metabolomics assays, using nuclear magnetic resonance spectrometry (NMR). Fasting glucose and insulin were also available for a random 10% of the cohort at age nine years, as well as research clinics at 15 and 18 years. Triglycerides, high-density lipoprotein (HDL-c), and total cholesterol (TC) were available from cord blood and venous blood, subsequently. Non-fasting samples were available from research clinics at ages 7 and 9 years; fasting measures were available from clinics at 15 and 18 years. Non-HDL-c was calculated by deducting HDL-c from total cholesterol. Hence, the trajectories of triglycerides, HDL-c, non-HDL-c, insulin and glucose are a combination of measures from cord blood, fasting, non-fasting bloods, and NMR. Interleukin-6 (IL-6) was measured by ELISA (R&D systems, Abingdon, UK) and available at 9 years, and C-reactive protein (CRP) was measured by automated particle-enhanced immunoturbidimetric assay (Roche UK, Welwyn Garden City, UK). CRP levels were measured at 9, 15 and 18 years of age. More details of measurement assays and protocols are provided as *Extended data*[17].

Physical activity. All children who attended the clinic at ages 11, 13, and 15 years were asked to wear an accelerometer for 7 days. More details of the physical activity measures are provided as *Extended data*[17].

Genetic data

DNA samples from 9,912 ALSPAC children were genotyped on the Illumina HumanHap550-quad single nucleotide polymorphism (SNP) array genotyping platform. After quality control assessment methods, imputation, and restriction to 1 child per family, genetic data was available for 7,977 individuals. More details are available in *Extended data*[17].

Statistical analysis

Generating polygenic risk scores. PRS were computed according to the method described by the International Schizophrenia Consortium[18], based on the summary statistics from the genome-wide association study (GWAS) of Alzheimer's disease by the IGAP consortium[8]. Details about IGAP are in the

Extended data[17]. PRS were created for the significance thresholds: $5×10^{-8}$ (genome-wide significance threshold), $5×10^{-2}$, and $5×10^{-1}$, including a different number of SNPs (Table S1 of *Extended data*[17]). A list of the 17 SNPs used to generate the PRS at the genome-wide significant threshold is provided in Table S2 of the *Extended data*[17]. This threshold was chosen for the main analysis since it includes the most strongly associated SNPs. However, less stringent p-value thresholds can explain more of the Alzheimer's disease variance. Results for the other thresholds are available in the *Extended data*[17]. As *ApoE i*s involved in lipid homeostasis by mediating lipid transport[19], the main analysis used scores excluding *ApoE.*

Modelling trajectories of cardiometabolic risk factors. Children with at least one measure were included in the multilevel models to minimise selection bias, under a missing at random assumption[20–23]. All trajectories were estimated using linear spline multilevel models (two levels: measurement occasion and individual). The optimal linear spline model for each cardiometabolic risk factor was based on previous work[21–27]. All trajectories were modelled in MLwiN version 3.01[28], called from STATA version 14[29] using the runmlwin command[30]. All trajectories were centred on the lowest age (in whole years) at the first available measure and mean PRS. Variables with a skewed distribution (triglycerides, fat mass, CRP) were (natural) log transformed. Fat and lean mass were adjusted for height with the use of a time- and sex-varying height covariate derived in previous literature[22,23,27], which was centred on the mean of this derived covariate. The non-linearity of height trajectories was modelled using fractional polynomials to find the best fitting function of age. Further details of trajectories are in *Extended data*[17]. We incorporated a fixed effect for the PRS in each trajectory along with an interaction term for the PRS and the linear splines or fractional polynomials. The coefficients represent the mean difference in the risk factor at baseline per one standard deviation (SD) higher PRS, and the mean change in the risk factor per year, or per year within each linear spline period, per one SD higher PRS. Results including the logged risk factors are interpreted as mean percentage change rather than mean difference. The mean trajectory was allowed to differ between males and females by including an interaction between gender and age. Results were interpreted according to the American Statistical Association guidance[31,32].

Modelling risk factors with sparse measures. Due to the sparse measures for birthweight, insulin, IL-6, and MVPA, linear regression models were used to estimate the association between the PRS and these measures. Measures from the clinic with the highest number of participants were used for the analysis, where applicable. As insulin, IL-6 and MVPA had a skewed distribution, they were transformed using the natural logarithm. Models were adjusted for age, sex and the first three ancestry-informative principal components.

Sensitivity analyses
We examined the associations between the PRS including only the two SNPs tagging the *ApoE* region (rs7412 and rs429358) and the PRS excluding *ApoE* and cardiometabolic risk factors

at different p-value thresholds. To examine if our results were changed by the inclusion of non-fasted bloods in some of the clinics, we repeated analyses excluding all participants who reported eating four hours prior to the 15 and 18-year clinics. To test the generalisability of our results, we investigated the characteristics of individuals in the analysis of insulin (model with the fewest individuals and repeated measures) to those excluded from this analysis (individuals with no genetic or cardiometabolic data).

An earlier version of this article can be found on bioRxiv (DOI: https://doi.org/10.1101/580548).

Results
Descriptive statistics
The number of children used in the models ranged from 699 to 6,953 individuals. Descriptive and model fit statistics are in Supplementary Table S3–Table S4 of the *Extended data*[17]. The household social class of children in the analysis of glucose (model with the smallest sample size) was higher compared to children excluded due to missing data (19.6% professional vs. 12.3% professional, p<0.0001). Furthermore, the prevalence of maternal smoking during pregnancy was lower in the included sample compared to the excluded sample (17.5% vs 28.2%, p<0.0001) (Supplementary Table S5 of the *Extended data*[17]).

Main analysis
Figure 1–Figure 3 show the mean trajectories of anthropometry, blood pressure, and blood-based biomarkers measures, respectively, alongside the association between the PRS and each of these cardiometabolic risk factor trajectories. At the genome-wide significant p-value threshold, there was weak evidence of association of a PRS with birthweight (Table S6.1 of *Extended data*[17]) and height trajectories (Figure 1A). At age nine years, there was weak evidence that a PRS was associated with a higher height-adjusted fat mass (β: 0.59%; 95% confidence interval [CI]: -0.92, 2.11, Table S6.3 of *Extended data*[17]) and height-adjusted lean mass (β: 0.04 kg; 95% CI: -0.03, 0.11, Table S6.3 of *Extended data*[17]). In line with these results, there was little evidence of an association between the PRS, SBP, glucose, triglycerides, non-HDL, HDL-c and CRP (Figure 3). There was stronger evidence of an association with higher DBP (β: 0.19 mmHg; 95% CI: 0.02, 0.37) at age 7 years and lower DBP (β: -0.05 mmHg; 95% CI: -0.09, -0.003) from ages 7 to 12 years (Table S6.4). Cross-sectional models showed weak evidence of association with higher insulin, IL-6, and MVPA levels (Table S6.8–6.9 of *Extended data*[17]). The direction of effect between time periods in multilevel models was consistent only for CRP (Table S6.7 of *Extended data*[17]) and MVPA (Table S6.9 of *Extended data*[17]). For all risk factors except for DBP, the 95% CIs spanned the null.

Sensitivity analyses
Sensitivity analyses comparing the results for the PRS at different p-value thresholds used PRS excluding *ApoE*. Further analyses comparing the results with the PRS including/excluding *ApoE* were performed using the genome-wide significant threshold. The effect size for the β-coefficients was larger for
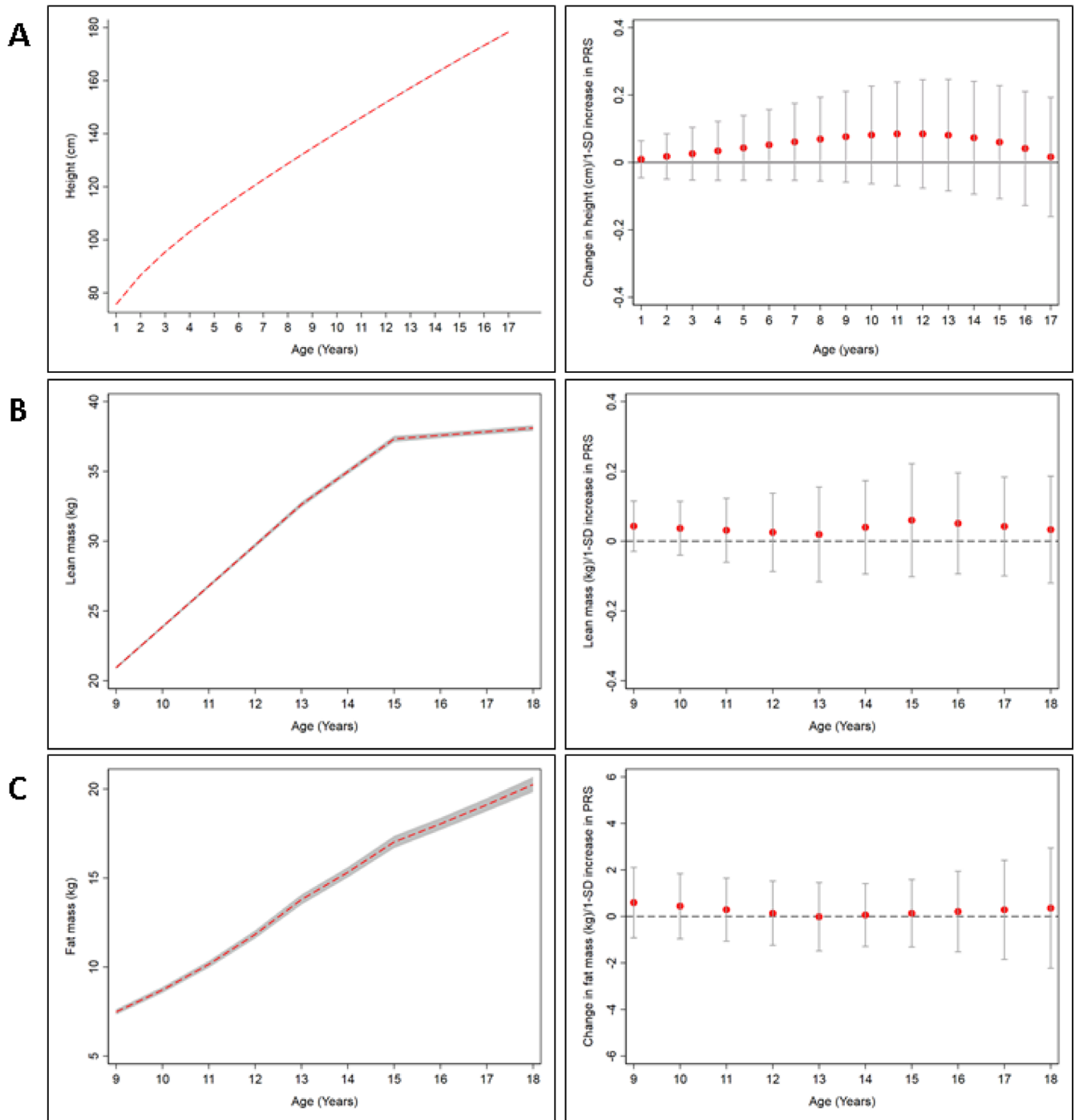
Figure 1. Mean trajectories of anthropometric measures and mean change in anthropometric measures per 1-SD increase in polygenic risk scores (PRS) from ages 9–18 years in ALSPAC. (A) Height, **(B)** lean mass, and **(C)** fat mass. The shaded areas in the left graphs and the bars in the right graphs represent 95% confidence intervals, respectively.
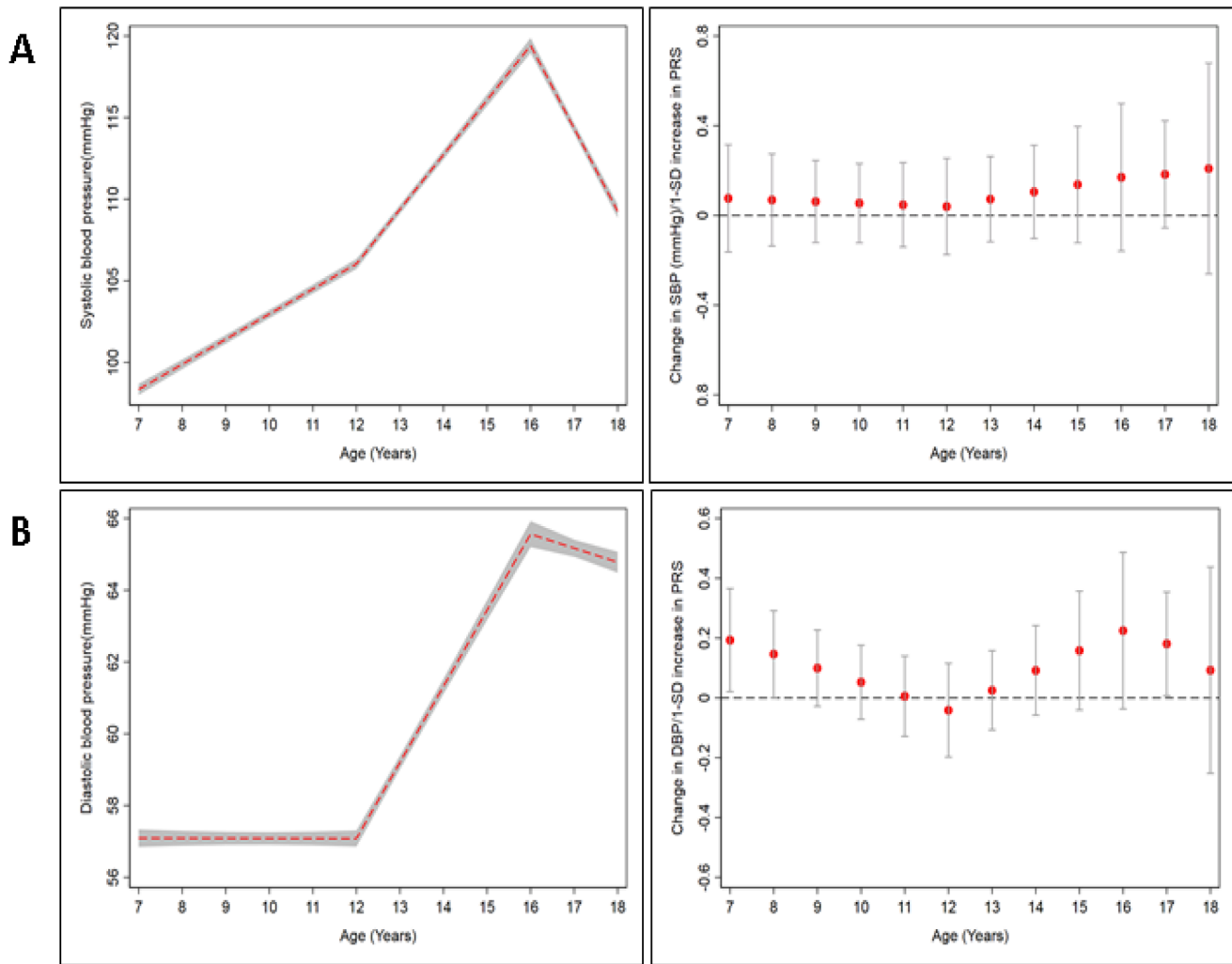
Figure 2. Mean trajectories of blood pressure measures and mean change in blood pressure measures per 1-SD increase in polygenic risk scores (PRS) from ages 7–18 years in ALSPAC. (**A**) Systolic blood pressure and (**B**) diastolic blood pressure. The shaded areas in the left graphs and the bars in the right graphs represent 95% confidence intervals, respectively.

risk factors such as height, lean mass, triglycerides and CRP at the more liberal PRS thresholds. A higher PRS was associated with shorter stature at the more liberal thresholds at ages 4–10 years (β: -0.15 cm; 95% CI: -0.20, -0.01 at age 5 years and p≤5×10⁻¹, Table S8.2 of *Extended data*[17]). A higher PRS was associated with lower lean mass at ages 9 and 15–18 years (β: -0.06 kg/yr; 95% CI: -0.11, -0.02 at age 15 years and p≤5×10⁻¹, Table S8.3 of *Extended data*[17]). The direction of effect differed between the genome-wide significant threshold and the more liberal thresholds for some risk factors (Table S7– Table S10 of *Extended data*[17]). PRS which incorporated the *ApoE* region were associated with cardiometabolic risk factors such as HDL-c, non-HDL-c, DBP, and CRP as expected by the role of *ApoE* in lipid homeostasis. There was strong evidence of association between a PRS including the *ApoE* region and higher non-HDL-c levels at birth (β: 0.03; 95% CI: 0.01, 0.04, Table S9.6 of *Extended data*[17]) and from ages 9–12 years (β: -0.003; 95% CI: -0.01, -0.0005). There was also strong

evidence that a PRS with *ApoE* was associated with lower HDL-c levels at birth (β: -0.01 mmol/l; 95% CI: -0.02, -0.005, Table S9.6 of *Extended data*[17]) and from birth to age 7 years (β: -0.002 mmol/l yr; 95% CI: -0.003, -0.0004), and higher HDL-c levels from ages 7–18 years (β: 0.001 mmol/l yr; 95% CI: 0.0002, 0.002). For a 1 SD increase in the PRS, CRP levels were 8.60% (95% CI; -11.86, -5.21, Table S9.7 of *Extended data*[17]) lower at age 9 years. To compare fasting/ non-fasting biomarkers, we used the PRS excluding *ApoE* at the genome-wide significant threshold. The associations observed between non-fasted and fasted bloods were similar for all examined blood-based biomarkers (Tables S10 of *Extended data*[17]).

Discussion

To our knowledge, this is the first study examining whether a PRS for Alzheimer's disease was associated with trajectories of cardiometabolic risk factors through childhood and adolescence. We
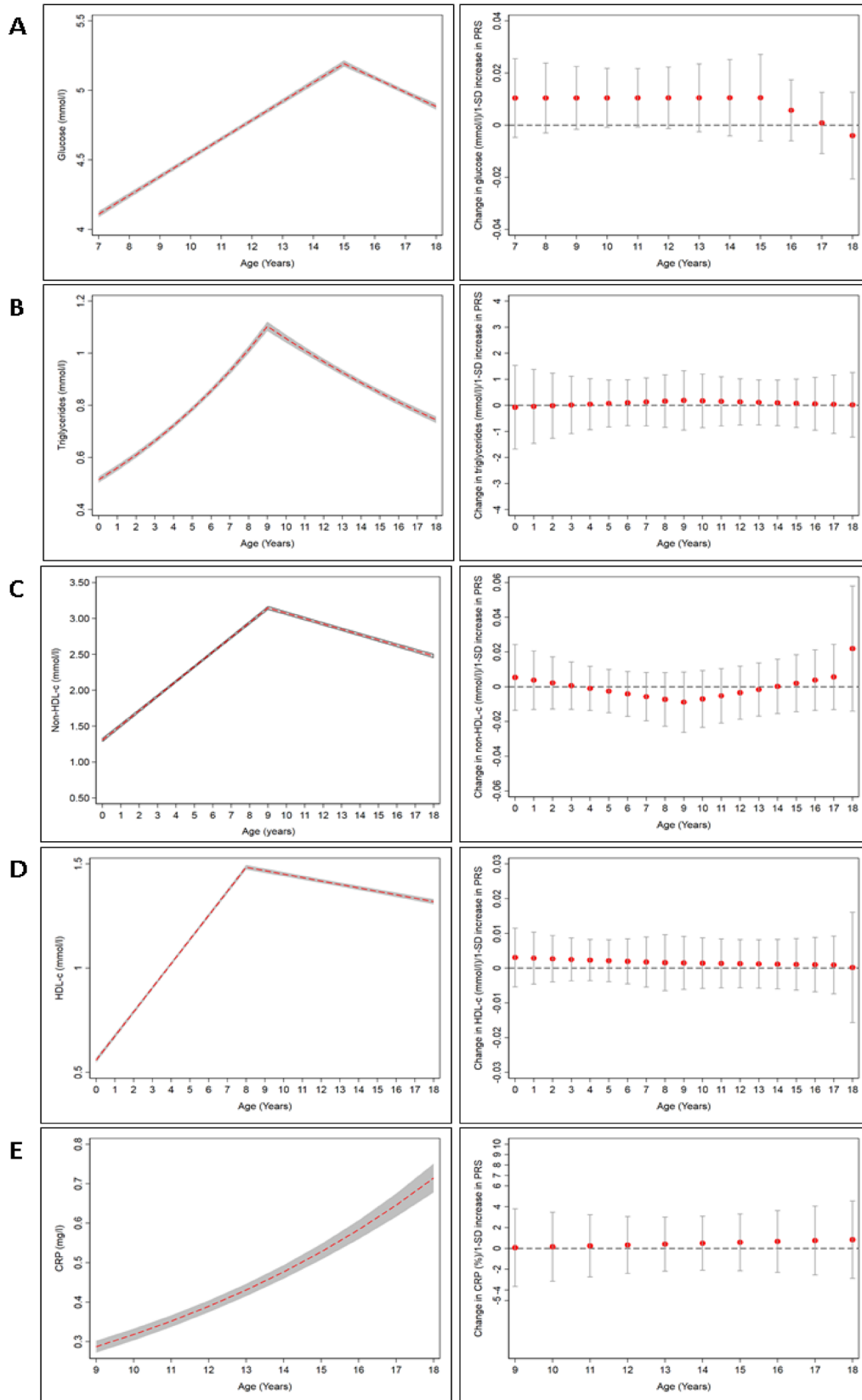
Figure 3. Mean trajectories of blood-based biomarkers and mean change per 1-standard deviation increase in PRS. (**A**) Glucose from ages 7–18 years, (**B**) triglycerides, (**C**) non-high-density lipoprotein-cholesterol (HDL-c), and (**D**) HDL-c from ages 0–18 years. (**E**) C-reactive protein (CRP) from ages 9–18 years. The shaded areas in the left graphs and the bars in the right graphs represent 95% confidence intervals, respectively.

examined these associations firstly to test whether associations may emerge early in the life course, and secondly to minimise selection bias due to attrition from cardiometabolic-associated morbidity and mortality. We did not find any consistent evidence to suggest that the PRS for Alzheimer's disease was associated with cardiometabolic risk factors in childhood or adolescence. The predominantly null results at the genome-wide significant p-value threshold may indicate that either the PRS for Alzheimer's disease does not operate through cardiometabolic pathways or that it operates partly through cardiometabolic pathways, but the associations emerge much later in life.

In our sensitivity analyses, we observed associations between the more liberal PRS and some of the cardiometabolic risk factors (such as height, lean mass, triglycerides, insulin, and CRP). For example, we found evidence of association with shorter stature from ages 4–10 years, lower lean mass at ages 9 years and from ages 15–18 years (at p-value threshold $p \leq 5 \times 10^{-1}$). As these PRS include many genetic variants which are weakly associated with Alzheimer's disease and the associations observed are dissimilar in the direction of effect for all time-points, there is a chance that the associations may be attributed to horizontal pleiotropy where the PRS is associated with a range of independent phenotypes. Furthermore, the effects of the PRS at each p-value threshold cannot be disentangled since PRS at higher p-value thresholds include all SNPs from PRS at lower p-value thresholds. We also found evidence of an association between the PRS for Alzheimer's disease (including the *ApoE* region) and HDL, non-HDL, CRP, and triglyceride levels, but the direction of effect was not consistent across all time-points.

Comparison to other studies

Longitudinal cohort studies with prospective phenotypic data through early life into adulthood and clinical diagnosis of Alzheimer's disease are scarce. Studies researching developmental factors such as height are the only source of data on early life exposures. It has been hypothesised that suboptimal foetal development, indexed by small birth weight/size may result in permanent changes in the structure, metabolism of the organs through various biological mechanisms[33]. By contrast, positive associations have been found between several cardiometabolic risk factors in midlife and Alzheimer's disease[1], with higher DBP and SBP increasing the risk of Alzheimer's disease[5], independent of the *ApoE* genotype. A positive association has also been found for BMI trajectories in midlife and Alzheimer's disease[34].

A study of PRS for Alzheimer's disease in adults with Alzheimer's disease and/or a family history of Alzheimer's disease identified an inverse association between a PRS, hypertension and stroke, but no associations for type 2 diabetes and heart disease[35]. Studies in children have only focused on *ApoE*, where there is an established linear relationship between *ApoE* and LDL in children and adults[36,37]. Although the *ApoE* region is the largest genetic risk factor for Alzheimer's disease, it is also strongly independently associated with cardiometabolic risk factors (primarily lipids), making it difficult to exclude pleiotropy as an explanation for any observed findings (i.e. *ApoE* may be associated with cardiometabolic risk independently of Alzheimer's disease. Contrary to a recent GWAS[38], we did not

find an association between *ApoE4* and physical activity. In agreement with our results, a study found a negative genetic correlation between height and dementia, suggesting our findings for height at liberal thresholds may be due to shared genetic pathways[39].

Strengths and limitations

We performed our analysis in children and adolescents to examine whether genetic risk for Alzheimer's disease operates through cardiometabolic pathways in early in the life course and to minimise selection bias due to attrition from cardiometabolic-associated morbidity and mortality. Additionally, the use of genetic scores makes confounding by social and lifestyle characteristics highly unlikely. Our study benefits from a well-characterised cohort with repeated measures of a range of cardiometabolic risk factors over a course of 18 years. The use of multilevel models enabled us to account for the clustering of repeated measurements within individuals and the correlation between repeated measurements over time.

However, as genetic scores explain a small proportion of variance in the outcome, it is highly likely we may have been underpowered to detect small polygenic effects on the examined cardiometabolic risk factors. For this reason, we did not correct our results for multiple testing, as our findings were largely null. Further limitations include the use of non-fasting and fasting bloods for some of the examined risk factors and the availability of measures from birth for only four of the risk factors. However, the estimates for fasted and non-fasted bloods were similar in our sensitivity analyses. Furthermore, PRS were derived from a GWAS of European individuals and our study was also conducted in individuals of European ancestry; hence, these results may not be generalizable to other populations.

Conclusion

We found little evidence to suggest that the combined genetic effects conferring an increased risk for Alzheimer's disease were associated with cardiometabolic risk factors, beyond the *ApoE* effect. As this is the first study to examine these associations through childhood and adolescence, these findings should be replicated in other large birth cohorts to examine whether the genetic risk for Alzheimer's disease can be captured in early childhood, or whether it becomes phenotypically manifest in adulthood.

Data availability

Underlying data

ALSPAC data access is through a system of managed open access. The steps below highlight how to apply for access to the data included in this article and all other ALSPAC data:

1. Please read the ALSPAC access policy (pdf, 800 kB) which describes the process of accessing the data and samples in detail, and outlines the costs associated with doing so.

2. You may also find it useful to browse our fully searchable research proposals database, which lists all research projects that have been approved since April 2011.

3. Please submit your research proposal for consideration by the ALSPAC Executive Committee. You will receive a

response within 10 working days to advise you whether your proposal has been approved.

Extended data

Open Science Framework: Polygenic risk score for Alzheimer's disease and trajectories of cardiometabolic risk factors in children.2019. https://doi.org/10.17605/OSF.IO/3ND8U[17].

This project contains supp_biorxiv.docx, which contains the following extended data:

- Supplementary methods

- Table S1. Number of SNPs included in PRS at different p-value thresholds including/excluding the ApoE region.

- Table S2. Genome-wide significant SNPs included in the genetic risk score based on IGAP GWAS Stage 1.

- Table S3. Descriptive statistics of cardiometabolic risk factors.

- Table S4. Details of model selection for each cardiometabolic risk factor.

- Table S5. Characteristics at birth of the mothers of children included in models of glucose (risk factor with least individuals).

- Table S6. Associations between cardiometabolic risk factors and PRS for Alzheimer's disease at $p \leq 5 \times 10^{-8}$.

- Table S7. Associations between cardiometabolic risk factors and PRS for Alzheimer's disease at $p \leq 5 \times 10^{-2}$.

- Table S8. Associations between cardiometabolic risk factors and PRS for Alzheimer's disease at $p \leq 5 \times 10^{-1}$.

- Table S9. Associations between cardiometabolic risk factors and PRS for Alzheimer's disease (including the ApoE region) at $p \leq 5 \times 10^{-8}$.

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Grant information

Acknowledgements

References

1. Baumgart B, Snyder HM, Carrillo MC, *et al.*: **Summary of the evidence on modifiable risk factors for cognitive decline and dementia: A population-based perspective.** *Alzheimers Dement.* 2015; **11**(6): 718–26.
   PubMed Abstract | **Publisher Full Text**

2. Crane PK, Walker R, Hubbard RA, *et al.*: **Glucose levels and risk of dementia.** *N Engl J Med.* 2013; **369**(6): 540–8.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Schmidt R, Schmidt H, Curb JD, *et al.*: **Early inflammation and dementia: a 25-year follow-up of the Honolulu-Asia Aging Study.** *Ann Neurol.* 2002; **52**(2): 168–74.
   PubMed Abstract | **Publisher Full Text**

4. Sundelöf J, Kilander L, Helmersson J, *et al.*: **Systemic inflammation and the risk of Alzheimer's disease and dementia: a prospective population-based study.** *J Alzheimers Dis.* 2009; **18**(1): 79–87.
   PubMed Abstract | **Publisher Full Text**

5.   Kennelly SP, Lawlor BA, Kenny RA: **Blood pressure and dementia - a comprehensive review.** *Ther Adv Neurol Disord.* 2009; **2**(4): 241–60.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6.   Gustafson DR, Luchsinger JA: **High adiposity: risk factor for dementia and Alzheimer's disease?** *Alzheimers Res Ther.* 2013; **5**(6): 57.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7.   Hernán MA, Alonso A, Logroscino G: **Cigarette smoking and dementia: potential selection bias in the elderly.** *Epidemiology.* 2008; **19**(3): 448–50.
     PubMed Abstract | **Publisher Full Text**

8.   Lambert JC, Ibrahim-Verbaas CA, Harold D, *et al.*: **Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease.** *Nat Genet.* 2013; **45**(12): 1452–8.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9.   Corder EH, Saunders AM, Strittmatter WJ, *et al.*: **Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families.** *Science.* 1993; **261**(5123): 921–3.
     PubMed Abstract | **Publisher Full Text**

10. Eichner JE, Dunn ST, Perveen G, *et al.*: **Apolipoprotein E polymorphism and cardiovascular disease: a HuGE review.** *Am J Epidemiol.* 2002; **155**(6): 487–95.
     PubMed Abstract | **Publisher Full Text**

11. Lahoz C, Schaefer EJ, Cupples LA, *et al.*: **Apolipoprotein E genotype and cardiovascular disease in the Framingham Heart Study.** *Atherosclerosis.* 2001; **154**(3): 529–37.
     PubMed Abstract | **Publisher Full Text**

12. Escott-Price V, Sims R, Bannister C, *et al.*: **Common polygenic variation enhances risk prediction for Alzheimer's disease.** *Brain.* 2015; **138**(Pt 12): 3673–84.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. Abbott RD, White LR, Ross GW, *et al.*: **Height as a marker of childhood development and late-life cognitive function: the Honolulu-Asia Aging Study.** *Pediatrics.* 1998; **102**(3 Pt 1): 602–9.
     PubMed Abstract | **Publisher Full Text**

14. Richards M, Hardy R, Kuh D, *et al.*: **Birth weight and cognitive function in the British 1946 birth cohort: longitudinal population based study.** *BMJ.* 2001; **322**(7280): 199–203.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

15. Fraser A, Macdonald-wallis C, Tilling K, *et al.*: **Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort.** *Int J Epidemiol.* 2013; **42**(1): 97–110.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. Boyd A, Golding J, Macleod J, *et al.*: **Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children.** *Int J Epidemiol.* 2013; **42**(1): 111–27.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

17. Korologou-Linden R: **Polygenic risk score for Alzheimer's disease and trajectories of cardiometabolic risk factors in children.** 2019.
     http://www.doi.org/10.17605/OSF.IO/3ND8U

18. International Schizophrenia Consortium, Purcell SM, Wray NR, *et al.*: **Common polygenic variation contributes to risk of schizophrenia and bipolar disorder.** *Nature.* 2009; **460**(7256): 748–52.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Mahley RW, Rall SC Jr: **Apolipoprotein E: far more than a lipid transport protein.** *Annu Rev Genomics Hum Genet.* 2000; **1**(1): 507–37.
     PubMed Abstract | **Publisher Full Text**

20. Tilling K, MacDonald-Wallis C, Lawlor DA, *et al.*: **Modelling childhood growth using fractional polynomials and linear splines.** *Ann Nutr Metab.* 2014; **65**(2–3): 129–38.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Howe LD, Tilling K, Matijasevich A, *et al.*: **Linear spline multilevel models for summarising childhood growth trajectories: A guide to their application using examples from five birth cohorts.** *Stat Methods Med Res.* 2016; **25**(5): 1854–74.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. O'Keeffe LM, Simpkin AJ, Tilling K, *et al.*: **Data on trajectories of measures of cardiovascular health in the Avon Longitudinal Study of Parents and Children (ALSPAC).** *Data Brief.* 2019; **23**: 103687.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23. O'Keeffe LM, Simpkin AJ, Tilling K, *et al.*: **Sex-specific trajectories of measures of cardiovascular health during childhood and adolescence: A prospective cohort study.** *Atherosclerosis.* 2018; **278**: 190–6.
     PubMed Abstract | **Publisher Full Text**

24. Morris TT, Northstone K, Howe LD: **Examining the association between early life social adversity and BMI changes in childhood: a life course trajectory analysis.** *Pediatr Obes.* 2016; **11**(4): 306–12.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

25. Howe LD, Parmar PG, Paternoster L, *et al.*: **Genetic influences on trajectories of systolic blood pressure across childhood and adolescence.** *Circ Cardiovasc Genet.* 2013; **6**(6): 608–14.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

26. Staley JR, Bradley J, Silverwood RJ, *et al.*: **Associations of blood pressure in pregnancy with offspring blood pressure trajectories during childhood and adolescence: findings from a prospective study.** *J Am Heart Assoc.* 2015; **4**(5): pii: e001422.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

27. O'Keeffe LM, Howe LD, Fraser A, *et al.*: **Associations of Y chromosomal haplogroups with cardiometabolic risk factors and subclinical vascular measures in males during childhood and adolescence.** *Atherosclerosis.* 2018; **274**: 94–103.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

28. Rabash J, Charlton C, Browne WJ, *et al.*: **MLwin version 3.01.** Centre for Multilevel Modelling. University of Bristol. 2009.

29. StataCorp: **Stata Statistical Software: Release 14.** 2016. College Station, TX; 2016.
     Reference Source

30. Centre for Multilevel Modelling U of B: **runmlwin: Stata module for fitting multilevel models in the MLwiN software.** 2016; **52**: 1–40.

31. Sterne JA, Davey Smith G: **Sifting the evidence-what's wrong with significance tests?** *BMJ.* 2001; **322**(7280): 226–31.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

32. Wasserstein RL, Lazar NA: **The ASA's Statement on p-Values: Context, Process, and Purpose.** *Am Stat.* 2016; **70**(2): 129–33.
     Publisher Full Text

33. Barker DJ, Gluckman PD, Godfrey KM, *et al.*: **Fetal nutrition and cardiovascular disease in adult life.** *Lancet.* 1993; **341**(8850): 938–41.
     PubMed Abstract | **Publisher Full Text**

34. Singh-Manoux A, Dugravot A, Shipley M, *et al.*: **Obesity trajectories and risk of dementia: 28 years of follow-up in the Whitehall II Study.** *Alzheimers Dement.* 2018; **14**(2): 178–86.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

35. Tosto G, Bird TD, Bennett DA, *et al.*: **The Role of Cardiovascular Risk Factors and Stroke in Familial Alzheimer Disease.** *JAMA Neurol.* 2016; **73**(10): 1231–7.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

36. Lehtimäki T, Porkka K, Viikari J, *et al.*: **Apolipoprotein E phenotypes and serum lipids in newborns and 3-year-old children: the Cardiovascular Risk in Young Finns Study.** *Pediatrics.* 1994; **94**(4 Pt 1): 489–93.
     PubMed Abstract

37. Bennet AM, Di Angelantonio E, Ye Z, *et al.*: **Association of apolipoprotein E genotypes with lipid levels and coronary risk.** *JAMA.* 2007; **298**(11): 1300–11.
     PubMed Abstract | **Publisher Full Text**

38. Klimentidis YC, Raichlen DA, Bea J, *et al.*: **Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including *CADM2* and *APOE*.** *Int J Obes (Lond).* 2018; **42**(6): 1161–76.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

39. Larsson SC, Traylor M, Burgess S, *et al.*: **Genetically-Predicted Adult Height and Alzheimer's Disease.** *J Alzheimers Dis.* 2017; **60**(2): 691–8.
     PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✅ ✅

Version 1

Reviewer Report 20 January 2020

https://doi.org/10.21956/wellcomeopenres.16777.r37376

✅ **Michelle Grunin**

Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA

Korologou-Linden *et al.*'s work aims to investigate whether polygenic risk scores (PRS) for Alzheimer's disease are associated with cardiometabolic measures and findings throughout the longitudinal Avon Longitudinal Study of Parents and children. They analyzed between 699 and 7,977 participants depending on the measure used, to determine if these factors throughout childhood and adolescences are associated with the PRS for Alzheimer's Disease.

The paper's statistics were very well done, including multiple different methodology depending on the factors and analyses. The results were null, but extremely interesting, and indexing this would contribute to the knowledge in the field.

A few minor comments:
1. The PRS were taken from the GWAS performed by the IGAP consortium, from their article from 2013. However IGAP has since published an updated meta-analysis (Kunkle *et al.* 2019), and some of the SNPs mentioned may not have been included in the PRS proposed in this paper. If some significant SNPs were not included, perhaps the authors could mention if this would affect the null results, despite the fact that most of the previous LOAD associated SNPs originally found in 2013 were again found in 2019 (only 5 new GWAS loci were found in 2019). In addition, the expanded and more liberal thresholds of the PRS would also change if expanded further to the 2019 article.

2. APOE was excluded from most of the analysis as the effect would be confounded by its association with cardiometabolic issues. However, why was sensitivity analyses performed including APOE if so?

3. Why do the authors think that the direction of the effect differed between GWAS thresholds and more liberal thresholds for some risk factors (Supplementary Table 7)?

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Biostatistics, Alzheimer's disease genetics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 24 September 2019

https://doi.org/10.21956/wellcomeopenres.16777.r36271

✓   **Emily Baker** (iD)
Dementia Research Institute at Cardiff, Cardiff University, Cardiff, UK
Valentina Escott-Price (iD)
Medical Research Council Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK

Korologou-Linden *et al*. aim to investigate whether polygenic risk scores for Alzheimer's disease were associated with cardiometabolic health indicators through childhood and adolescence in a sample of 7,977 participants from the Avon Longitudinal Study of Parents and Children. The results are null, however, we support the authors in wishing to publish these results. The paper is well written but requires some clarification.

In the results section of the abstract, the authors state that no associations were observed excluding APOE gene region. Do they see association with APOE gene? Please mention this in the abstract.

Genetic Data: How many SNPs in the data? How was 1 child per family selected? Randomly?

In the Statistical analysis section the authors say that they excluded APOE from the PRS (gene or region?). Do they again excluded APOE in the sensitivity analyses section? The sentence "…*the PRS excluding ApoE and cardiometabolic risk factors at different p-value threshold*" is not clear. Is it SNP association p-values or cardiometabolic risk factors p-values? Could the authors please clarify in the statistical analyses a) exact positions of APOE region which was excluded; b) why the APOE tagging SNPs where included in the sensitivity analyses; c) clarify what other models were used for the sensitivity analyses and why .

In the Results section, the sentences "*Sensitivity analyses comparing the results for the PRS at different p-value thresholds used PRS excluding ApoE. Further analyses comparing the results with the PRS including/excluding ApoE were performed using the genome-wide significant threshold*" are not clear. According to the methods section "the exclusion of APOE" model is a part of the main analysis. In addition, why were "including apoe" models only tested for GWAS significant SNPs in the PRS?

Generating polygenic risk scores: How was APOE excluded? Just the APOE SNPs, or the whole APOE locus? How was linkage disequilibrium between SNPs handled? If LD pruning was used, which parameters were used, i.e. r2 threshold.

Figures - the figures seem slightly blurred, a higher resolution would be nice. Perhaps Figures 1 & 3 could go to supplemental material?

Was the correlation between HDL, non-HDL, CRP and triglyceride levels in a consistent direction at all time points?

In the discussion section, it may also be that the effect sizes are very small and the sample size is not sufficient to reach the statistical significance (N ranged between ~700 and 7,000).

Minor points:
The following text should be reworded:
- Abstract, method: 'birth at 18 years'.

- Pg. 3 Intro paragraph 2 'can be summarised to construct a polygenic risk score'.

- Pg. 3 Methods paragraph 1, keep numbers consistent 'seven, nine, 10, …' There are other instances of this through the paper.

- Pg. 8 strengths and limitations 'in early in the life course'.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Biostatistics and Bioinformatics.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.