medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

**Research letter:** The proportion of Alzheimer's disease attributable to Apolipoprotein E

Dylan M Williams PhD[1], Neil M Davies PhD[2], Emma L. Anderson PhD[2]

[1] MRC Unit for Lifelong Health & Ageing at UCL, University College London, London, UK

[2] Division of Psychiatry, University College London, Maple House, 149 Tottenham Court Rd, London W1T 7NF

**Corresponding author:**

Dylan M. Williams; Principal Research Fellow, MRC Unit for Lifelong Health and Ageing at UCL, University College London, Floor 5, 1-19 Torrington Place, London, WC1E 7HB, UK; Tel. +44(0) 20 7670 5713; dylan.williams@ucl.ac.uk

**Article word count:** 600

**Key words:** Apolipoprotein E, *APOE*, Alzheimer's disease, dementia, population attributable fraction

**NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

**Abstract**

*Introduction*

The extent to which genetic variation at the *APOE* locus explains the burden of late-onset Alzheimer's disease (AD) is poorly understood. We aimed to provide new estimates of the proportions of AD and all-cause dementia attributable to combined carriage of ε3 and/or ε4 alleles of *APOE*.

*Methods*

We conducted a nested case-control study using data from 171,133 participants of the UK Biobank cohort study, aged ≥60 years at baseline assessments in 2006-2010. Carriage of ε2/ε3/ε4 alleles of *APOE* were coded from genotyped or imputed microarray data for single nucleotide polymorphisms rs7412 and rs429358. AD and all-cause dementia were ascertained from baseline self-report and follow-up via linked electronic health and death records up to December 2022 (minimum/maximum follow-up: 12.2 / 16.8 years). Risks of these outcomes due to ε3 and/or ε4 carriage were modelled with multivariable logistic regression, adjusting for age at baseline, self-reported sex and ethnicity, 10 genetic principal components and genotyping array. Odds ratios and prevalence of ε3 and ε4 carriage were used to calculate population attributable fractions (PAFs) of the outcomes due to these genotypes.

*Results*

99.4% of the sample had either ε3 and/or ε4 carriage. By the end of follow-up, 3026 (1.8%) and 6634 (3.9%) of the sample had AD and all-cause dementia, respectively. The odds ratio for AD risk due to ε3 and ε4 carriage with reference to ε2 homozygotes was 3.80 (95% CI:1.58, 9.17). The equivalent risk for all-cause dementia was 1.74 (95% CI: 1.16, 2.61). PAFs for AD and all-cause dementia burden due to ε3 and ε4 exposure were 73.6% (95% CI: 36.6, 89.0) and 42.4% (95% CI: 13.6, 61.6%), respectively.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

*Conclusions*

Differences in the molecular physiology of Apolipoprotein E cause most AD and a large fraction of dementia cases. Research into this pathway should be prioritised to facilitate dementia prevention.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

**Main text**

The extent to which common variation in the *APOE* gene explains the burden of late-onset Alzheimer's disease (AD) is poorly understood. To address this, we provide new estimates for the proportions of AD and all-cause dementia cases attributable to carriage of both the common ε3 and ε4 alleles.

We conducted a nested case-control study using data from 171,133 participants of the UK Biobank cohort, aged ≥60 years at baseline in 2006-2010.[1] *APOE* ε2/ε3/ε4 genotypes were coded as one exposure representing ε3 and/or ε4 carriage, with ε2 homozygotes as reference. AD and all-cause dementia were identified from self-report at baseline and follow-up via linked electronic health and death records available up to December 2022 (minimum/maximum follow-up: 12.2 / 16.8 years). Risk of each outcome was modelled using multivariable logistic regression (online supplement). The prevalence of the exposure in the full sample and computed odds ratios were used to calculate population attributable fractions (PAFs) for AD and all-cause dementia.

Characteristics of the UK Biobank sample are displayed in table 1. 99.4% of the sample had ε3 and/or ε4 carriage. The odds ratio for AD due to ε3 or ε4 carriage was 3.80 (95% CI:1.58, 9.17; table 2). The equivalent risk for all-cause dementia was 1.74 (95% CI: 1.16, 2.61). PAFs based on these statistics were 73.6% (95% CI: 36.6, 89.0%) and 42.4% (95% CI: 13.6, 61.6%) for AD and all-cause dementia, respectively.

Thus, if interventions could obviate entirely the detrimental effects of ε3 and ε4 carriage on these outcomes in a population akin to the UKB sample, we could expect to prevent approximately three quarters of AD cases. Such a magnitude has been suggested for combined ε3 and ε4 carriage previously (perhaps accounting for as much as 95% of AD[2]) but not, to our knowledge, demonstrated empirically until now. Due to the properties of genetic inheritance, risk estimates for genetic variants such as the *APOE* alleles are not subject to reverse causation and they are unlikely to be affected by confounding. Hence,

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

PAFs for variants may provide more robust estimates of disease burden attributable to the variation in question than PAFs estimated for environmental factors.[3]

These findings should remind researchers and funders that, as the major aetiological factor determining most AD (and a large proportion of all dementia), the mechanisms linking Apolipoprotein E (ApoE) with AD should receive a proportionate amount of the attention and funding devoted to dementia research. This should include efforts to understand the distinct functional properties of the ε3 isoform that confer AD risk – relative to properties of the ε2 isoform and other protective variants – and not only further research to elucidate and mitigate ε4's effects.

Limitations of this research include incomplete ascertainment of AD and all-cause dementia cases due to limited record linkage in UKB (which does not yet fully extend to primary care and mental health service records).[4] Measurement of outcome lifetime risks was incomplete – the youngest participants in our sample were aged 73 years by the current end of follow-up. Recruitment into UKB was not representative of the general UK population and was prone to selection effects,[5] perhaps including effects from *APOE* genotypes contributing to cardiovascular and other morbidity and mortality prior to age 60 years.[6] However, we anticipate that these biases would lead to underestimated, rather than inflated, PAFs.

To conclude, the root cause of most AD is differences in the molecular physiology of ApoE, principally due to properties of the ε3 and ε4 isoforms. Establishing precisely how, when and in which cell types ApoE influences AD risk – and how its deleterious effects can be mitigated – is paramount to AD prevention and treatment.

**Acknowledgements**

This research has been conducted using the UK Biobank Resource under application number 71702.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

*Author contributions*

DMW conceived the study, undertook the analyses and drafted the manuscript. All authors contributed to the interpretation of data and the content of the manuscript, and approved its final version. DMW had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

We have no conflicts of interest to disclose.

The funders had no direct role in the production of this research.

**Funding**

The MRC Unit for Lifelong Health and Ageing at UCL is funded by the Medical Research Council (MC_UU_00019/3). NMD is supported via a Norwegian Research Council Grant number 295989. ELA is supported by a UKRI Future Leaders Fellowship (MR/W011581/1). For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

**Data sharing statement**

All data used in this research are available to researchers that register with UK Biobank and request access to them as part of an approved project: https://www.ukbiobank.ac.uk/. Data fields and script used in the analyses can be viewed at: *github link to be inserted upon M/S acceptance*.

**References**

1. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203-209.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

2. Raber J, Huang Y, Ashford JW. ApoE genotype accounts for the vast majority of AD risk and AD pathology. *Neurobiology of Aging*. 2004;25(5):641-650.

3. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*. 2020;396(10248):413-446.

4. Wilkinson T, Ly A, Schnier C, et al. Identifying dementia cases with routinely collected health data: a systematic review. *Alzheimer's & Dementia*. 2018;14(8):1038-1051.

5. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *American Journal of Epidemiology*. 2017;186(9):1026-1034. doi:10.1093/aje/kwx246

6. Nelson CP, Goel A, Butterworth AS, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nature Genet*. Sep 2017;49(9):1385-1391. doi:10.1038/ng.3913

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Table 1. Descriptive characteristics on the sub-set of UK Biobank participants in this analysis (N=171,133)

| **Characteristic** | **Value** |
|---|---|
| Age (years) at baseline, mean (SD) | 64.1 (2.8) |
| Female sex, No. (%) | 89,320 (52.1) |
| Ethnicity, No. (%) | |
| White | 165,568 (96.8) |
| Asian | 2,369 (1.4) |
| Black | 1,448 (0.9) |
| Mixed | 535 (0.3) |
| Chinese | 284 (0.2) |
| Other | 929 (0.5) |
| *APOE* genotype, No. (%) [a] | |
| ε2 carriage | 26,249 (15.3) |
| ε3 carriage | 161,855 (94.6) |
| ε4 carriage | 48,179 (28.2) |
| ε2 homozygous | 1,053 (0.6) |
| ε3 homozygous | 100,892 (59.0) |
| ε4 homozygous | 4,038 (2.4) |
| | |
| Alzheimer's disease, No. (%) | 3,026 (1.8) |
| All-cause dementia, No. (%) | 6,634 (3.9) |

[a] NB: carriage groups overlap and hence percentages sum to over 100%

medRxiv preprint doi: https://doi.org/10.1101/2023.11.16.23298475; this version posted November 17, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Table 2. Odds ratios for AD or all-cause dementia by *APOE* genotypes

| **Exposure** [a] | **No.** [b] | **AD** | | | **All-cause dementia** | | |
|---|---|---|---|---|---|---|---|
| | | OR | 95% CI | P | OR | 95% CI | P |
| ε3 and/or ε4 carriage | 171133 | 3.80 | (1.58, 9.17) | 0.003 | 1.74 | (1.16, 2.61) | 0.008 |
| | | | | | | | |
| ε2/ε3 | 21962 | 1.63 | (0.66, 3.97) | 0.29 | 0.95 | (0.63, 1.45) | 0.83 |
| ε3/ε3 | 101761 | 2.01 | (0.83, 4.85) | 0.12 | 1.11 | (0.74, 1.68) | 0.40 |
| ε2/ε4 | 5189 | 3.95 | (1.59, 9.81) | 0.003 | 2.06 | (1.33, 3.17) | 0.001 |
| ε3/ε4 | 41007 | 7.73 | (3.20, 18.66) | <0.001 | 3.14 | (2.08, 4.72) | <0.001 |
| ε4/ε4 | 5088 | 24.67 | (10.16, 59.89) | <0.001 | 9.42 | (6.21, 14.28) | <0.001 |

[a] In all models, the reference group was ε2/ε2 carriers

[b] Numbers show total sample sizes per analysis – i.e. the sum of individuals with the exposure genotype and ε2/ε2 carriers (N=1053)