

RESEARCH ARTICLE

APOE alleles are associated with sex-specific structural differences in brain regions affected in Alzheimer's disease and related dementia

Chloé Savignac¹, Sylvia Villeneuve^{2,3,4,5}, AmanPreet Badhwar^{6,7}, Karin Saltoun¹, Kimia Shafighi¹, Chris Zajner¹, Vaibhav Sharma¹, Sarah A. Gagliano Taliun^{8,9}, Sali Farhan^{2,10}, Judes Poirier^{2,4,5}, Danilo Bzdok^{1,3,11,12*}

1 Department of Biomedical Engineering, Faculty of Medicine, McGill University, Montreal, Quebec, Canada, **2** Department of Neurology and Neurosurgery, Montreal Neurological Institute (MNI), Faculty of Medicine, McGill University, Montreal, Quebec, Canada, **3** McConnell Brain Imaging Centre (BIC), MNI, Faculty of Medicine, McGill University, Montreal, Quebec, Canada, **4** Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, Quebec, Canada, **5** Centre for Studies in the Prevention of Alzheimer's Disease, Douglas Mental Health Institute, McGill University, Montreal, Quebec, Canada, **6** Department of Pharmacology and Physiology, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada, **7** Centre de recherche de l'Institut universitaire de gériatrie de Montréal (CRIUGM), Montreal, Quebec, Canada, **8** Department of Neurosciences & Department of Medicine, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada, **9** Montreal Heart Institute, Montréal, Quebec, Canada, **10** Department of Human Genetics, Faculty of Medicine, McGill University, Montreal, Quebec, Canada, **11** School of Computer Science, McGill University, Montreal, Quebec, Canada, **12** Mila—Quebec Artificial Intelligence Institute, Montreal, Quebec, Canada

* danilo.bzdok@mcgill.ca



OPEN ACCESS

Citation: Savignac C, Villeneuve S, Badhwar A, Saltoun K, Shafighi K, Zajner C, et al. (2022) *APOE* alleles are associated with sex-specific structural differences in brain regions affected in Alzheimer's disease and related dementia. *PLoS Biol* 20(12): e3001863. <https://doi.org/10.1371/journal.pbio.3001863>

Academic Editor: Laurel A. Beckett, University of California Davis School of Medicine, UNITED STATES

Received: May 16, 2022

Accepted: September 30, 2022

Published: December 13, 2022

Copyright: © 2022 Savignac et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All used data are available to other investigators online (ukbiobank.ac.uk, <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases>, <https://openpreventad.loris.ca>). The Schaefer-Yeo atlas is accessible online (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal). The individual numerical values that underlie the summary data displayed in all the main and

Abstract

Alzheimer's disease is marked by intracellular tau aggregates in the medial temporal lobe (MTL) and extracellular amyloid aggregates in the default network (DN). Here, we examined codependent structural variations between the MTL's most vulnerable structure, the hippocampus (HC), and the DN at subregion resolution in individuals with Alzheimer's disease and related dementia (ADRD). By leveraging the power of the approximately 40,000 participants of the UK Biobank cohort, we assessed impacts from the protective *APOE* ϵ 2 and the deleterious *APOE* ϵ 4 Alzheimer's disease alleles on these structural relationships. We demonstrate ϵ 2 and ϵ 4 genotype effects on the inter-individual expression of HC-DN co-variation structural patterns at the population level. Across these HC-DN signatures, recurrent deviations in the CA1, CA2/3, molecular layer, fornix's fimbria, and their cortical partners related to ADRD risk. Analyses of the rich phenotypic profiles in the UK Biobank cohort further revealed male-specific HC-DN associations with air pollution and female-specific associations with cardiovascular traits. We also showed that *APOE* ϵ 2/2 interacts preferentially with HC-DN co-variation patterns in estimating social lifestyle in males and physical activity in females. Our structural, genetic, and phenotypic analyses in this large epidemiological cohort reinvigorate the often-neglected interplay between *APOE* ϵ 2 dosage and sex and link *APOE* alleles to inter-individual brain structural differences indicative of ADRD familial risk.

supplementary figures have been publicly deposited online (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)).

Funding: DB was supported by the Brain Canada Foundation, through the Canada Brain Research Fund, with the financial support of Health Canada, National Institutes of Health (NIH R01 AG068563A, NIH R01 R01DA053301-01A1), the Canadian Institute of Health Research (CIHR 438531, CIHR 470425), the Healthy Brains Healthy Lives initiative (Canada First Research Excellence fund), Google (Research Award, Teaching Award), and by the CIFAR Artificial Intelligence Chairs program (Canada Institute for Advanced Research). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: ADRD, Alzheimer's disease and related dementia; *APOE*, Apolipoprotein E; CA, cornu ammonis; CCA, canonical correlation analysis; DG, dentate gyrus; dmPFC, dorsomedial prefrontal cortex; DN, default network; FDR, false discovery rate; GDC, gradient distortion correction; HC, hippocampus; HPA, hypothalamus–pituitary–adrenal; IGF-1, insulin-like growth factor 1; IPL, inferior parietal lobule; MTS, middle temporal sulcus; MoCA, Montreal Cognitive Assessment; MRI, magnetic resonance imaging; OFC, orbitofrontal cortex; PCA, principal component analysis; PCC, posterior cingulate cortex; PCu, precuneus; PREVENT-AD, PResymptomatic Evaluation of Experimental or Novel Treatments for Alzheimer's disease (AD); RBANS, Repeatable Battery for Assessment of Neuropsychological Status; RSC, retrosplenial cortex; SD, standard deviation; SNP, single-nucleotide polymorphism; STG, superior temporal gyrus; SPL, superior parietal lobule; STS, superior temporal sulcus; SVD, singular value decomposition; TPJ, temporoparietal junction; UKB, UK Biobank; vlPFC, ventrolateral prefrontal cortex; vmPFC, ventromedial prefrontal cortex.

Introduction

Around the globe, >50 million people are living with dementia—a global burden of >1 trillion USD\$ annually [1]. By 2050, an estimated 3-fold increase in affected individuals is projected as a result of increased longevity [2]. The anticipated explosion in the number of dementia cases will put a strain on the 82 billion hours of annual informal care provided by caretakers worldwide [1]. In contrast to this secular trend, the age-specific prevalence of dementia is expected to decrease in certain high-income countries, which can be attributable to improvement in underlying health and socioeconomic determinants [2]. A recent authoritative report on dementia prevention has identified about a dozen potentially modifiable risk factors that could explain the disparity in Alzheimer's disease and related dementia (ADRD) incidence [3]. The disparate risk dimensions include personal habits and lifestyle, physical and mental health, as well as societal and external factors. New public health policies targeted at reducing mid- to late-life risk factors (e.g., physical inactivity, social disengagement, loneliness) thus have the potential to delay dementia onset in the most disadvantaged older adults. As the global prevalence of dementia is quickly rising, there is an unprecedented need to characterise the impact of genetic predisposition (e.g., Apolipoprotein E (*APOE*) polymorphism [4]) and modifiable risk factors on ADRD-vulnerable brain structures before the onset of cognitive decline.

Over the past 2 decades, brain-imaging studies have converged on the disruption of a coherent network of higher association regions that involve key nodes of the default network (DN) in individuals with ADRD compared to healthy controls [5]. Extensive efforts have mobilised resting-state functional connectivity analyses to investigate patients with ADRD, with converging results in the DN [6]. However, delineating a definitive profile of functional connectivity deviations related to ADRD risk in healthy subjects was plagued with slow progress. Most such biomarker studies have attempted to identify functional connectivity patterns that reliably tell apart $\epsilon 4$ carriers from non-carriers. Yet, most other *APOE* variants have been largely neglected, perhaps because they occur much more infrequently in the general population. The extensive literature on altered DN connectivity in $\epsilon 4$ carriers has yet to reach consensus as reports of both increased [7] and decreased [8] connectivity within nodes of the DN have repeatedly led to contradictory conclusions. Among the few studies that could investigate concurrent connectivity alterations in the hippocampus (HC) and regions of the DN in $\epsilon 4$ carriers, the HC was typically treated as a monolithic structure [9] rather than appreciating its functional and structural heterogeneity. That is, it was studied as a single node when interrogating its coupling links to other DN nodes [10]. These inconsistencies are probably also due in part to data acquisition and preprocessing methods for functional connectivity analysis, which have made some findings in $\epsilon 4$ carriers hard to replicate [11]. Moreover, because of the overwhelming singular focus on $\epsilon 4$ carriers in the research community, the neural correlates associated with other *APOE* variants remain underspecified. Of particular appeal, illuminating the allegedly opposing effects of *APOE* $\epsilon 2$ and $\epsilon 4$ on DN and HC integrity could be crucial in guiding potential treatment avenues, given the $\epsilon 2$ -associated protective outcome on brain structure [12].

A parallel stream of literature has focused on changes in hippocampal microstructure over the course of ADRD progression, mostly by performing thorough post-mortem autopsy on patients with probable ADRD. The HC formation is known for subfield-specific vulnerability to ADRD, at least since the late 1990s [13]. Yet, the HC is still routinely treated as if it was an anatomically homogeneous structure in common brain-imaging studies [9,14,15]. By extension, such an analytical approach is blind to the distinct links between HC subregions and DN subregions. In vivo examinations in the macaque monkey have shown that the HC formation receives important axon projections from the retrosplenial cortex and posterior cingulate

cortex in the presubiculum and parasubiculum subregions [16]. Yet, the fornix, which carries the axons from the CA1 and subiculum subregions, forwards the only hippocampal output signals that directly go to the ventromedial and orbitofrontal cortex (OFC) of the DN [17,18]. Glossing over these known microanatomical nuances could explain reports of poor predictive value of hippocampal atrophy in early ADRD stages when measuring the whole HC as a single unit. In a randomised clinical trial, baseline HC volumes, manually traced and corrected for inhomogeneity, predicted conversion to ADRD over a 3-year period at 60.4% accuracy [19]. With the advent of ultra-high-resolution atlases and advanced automatic sub-segmentation techniques, assessment of the subfield-specific vulnerability of both hippocampi to ADRD progression in an observer-independent fashion is now coming into reach. Instead of relying primarily on post-mortem autopsy from patients to ultimately confirm ADRD status, we will soon be able to directly, non-invasively, quantify the level of risk of a given patient based on subfield-level granular information. From the perspective of clinical translation, coming up with individual profiles of microstructural alterations characteristic of ADRD risk could usher a principled path toward precision medicine in neurology.

For these reasons, here we opted for structural brain imaging to relate genetic risk to robust codependence principles between neocortical DN and allocortical HC at subregion granularity. Given the panoply of individual factors that may affect cortical blood flow (e.g., vigilance, mood, cortisol levels, and coffee intake), functional connectivity would likely paint a more circumstantial portrait of ADRD vulnerability. We therefore designed an analytical framework for doubly multivariate decomposition to zoom in on the structural correspondence between HC and DN subregions at the population level. The two-pronged approach was carefully tailored to derive coherent signatures of HC-DN co-variation sensitive to the subregion-specific vulnerability of these neural circuits in ADRD. We were able to quantify the level of risk by looking for structural deviation in individuals with and without family history of ADRD by deep inspection of concomitant regimes of HC-DN co-variation. Capitalising on the rich phenotyping available for 40,000 UK Biobank participants, our study could confront the effects of *APOE* ϵ 2 and ϵ 4 on inter-individual expressions of HC-DN co-variation—something out of reach in traditional brain-imaging studies involving small to medium sample sizes. In doing so, our study was also uniquely positioned to illuminate possible sex-specific associations across less prevalent *APOE* gene variants that previous brain-imaging investigations systematically ignored.

Results

Rationale

In post-mortem autopsy of patients with ADRD, structural alterations of microanatomically defined subregions composing the human HC have been described in extenso [20]. Despite such insights from rigorous invasive studies, the overwhelming majority of existing brain-imaging studies have treated the HC as a monolithic brain structure. Hence, the specific vulnerability of its heterogeneous subregions to ADRD pathology remains largely concealed today. Advances in automatic segmentation techniques for the HC using *ex vivo* brain imaging allow for subject-specific parcellations that respect the diversity of distinct subregions identified post-mortem. Capitalising on these ultra-high-resolution segmentations, we are now equipped to assess microstructural alterations of the human HC in a newly detailed way that scales to the approximately 40,000 UK Biobank (UKB) participants [21]. These advances enabled us to describe ADRD-related patterns of structural co-variation in 91 DN subregions, which were in lockstep with 38 fine-grained HC subregions. Working at a population scale made it possible for us to investigate the effect of rare genotypes on brain structure. This

approach was especially fruitful for the less common *APOE* $\epsilon 2/\epsilon 2$ genotype, which has a prevalence of <1% among the general population [22]. Given this setup, our investigation was uniquely positioned to carry out sex-specific examinations across all *APOE* gene variants that previous brain-imaging studies systematically ignored. The availability of deep profiling of the UKB participants further allowed us to chart brain-behaviour associations across the whole phenome in an impartial data-driven approach.

Population signatures of HC-DN co-variation capture subregion-level structural ties

We first delineated the structural dependencies in regional grey matter volume between the subregion atlas of the HC and that of the DN to identify deviations that jointly go hand-in-hand. We benefitted from canonical correlation analysis (CCA), a doubly multivariate pattern-learning tool (cf. methods), to identify the sources of common population variation between the full sets of 38 HC subregions and that of 91 DN subregions. This algorithmic approach finds principled signatures of structural co-variation between 2 sets of variables [23]. Patterns of shared co-variation (canonical variates, cf. methods) embed the effects of HC or DN subregion sets in a new representational space where the 2 sets were most strongly correlated with each other. Pairs of canonical variates, 1 for the HC and 1 for the DN, are what we henceforth call modes. By construction, these are ranked by importance; each mode carries unique information by being uncorrelated from each other. Each mode thus represented a different brain signature that accounted for increasingly less shared variance between the neocortical and allocortical atlas at subregion resolution.

We focused on the leading 25 modes, mode 1 being the most explanatory signature of HC-DN co-variation under the elected model. The explanatory power of a given mode was quantified by Pearson's correlation between inter-individual variation tracked by its associated HC and DN patterns (canonical correlation, cf. method). The leading signature of HC-DN co-variation (mode 1) achieved a canonical correlation of $\rho = 0.51$, whereas the second and third signatures achieved correlations of $\rho = 0.42$ and 0.39 , respectively. Canonical correlations accounted for increasingly less joint variation between the HC and DN subregions up to the last signature (mode 25), which achieved a correlation of $\rho = 0.06$. The full list of correlation coefficients for the remaining modes has been published elsewhere [24] and is openly accessible online (https://figshare.com/articles/figure/Loneliness_Supplement_July_22_docx/15060684). This multivariate decomposition served as the backbone for all subsequent analyses that aimed to elucidate how individual expressions of HC-DN co-variation varied in relation to ADRD risk.

Signatures of HC-DN co-variation illuminate concomitant deviations in ADRD risk

To interrogate the neurobiological manifestations of ADRD family history in our UKB cohort, we performed a rigorous group difference analysis that highlighted any statistically robust ADRD-related divergences in each HC-DN population signature. In doing so, we uncovered the precise subset of anatomical subregions contributing to structural HC-DN co-variation that systematically diverged in individuals with versus without family history of ADRD. An HC or DN subregion observed to have a robustly different co-variation expression in individuals with and without family history of ADRD is henceforth termed a hit. We observed a total of 28 HC and 135 DN hits across the leading 25 modes. As a general trend, HC hits were mainly located in the cornu ammonis (CA) subregions (42.9% of total divergences). Parallel DN hits were predominantly observed in the prefrontal cortex (dorsomedial prefrontal cortex

(dmPFC) and ventrolateral prefrontal cortex (vlPFC); 45.9% of total divergences) and posterior midlines structures (posterior cingulate cortex (PCC), precuneus (PCu), and retrosplenial cortex (RSC); 27.4% of total divergences).

In mode 1, we identified 12 HC hits as indicative for family history of ADRD, with the strongest subregion effects identified in CA1, CA2/3, molecular layer, and granule cell layer of the dentate gyrus (DG) (66.7% of HC divergences in mode 1). The remaining HC hits for mode 1 were either located in the parasubiculum, CA4 or HC tail (Fig 1). We revealed 34 concomitant DN hits, most of them located in the prefrontal cortex (dmPFC and vlPFC) and posterior midline structures (RSC, PCC, and PCu) that represented 55.9% and 35.3% of total DN hits in mode 1, respectively. As for mode 2, 80.0% of the 10 identified HC hits were located in the left hemisphere (S1 Fig). Of those hits, the strongest weights were found in the presubiculum and CA2/3. The remaining HC hits were identified in the CA1, CA4, hippocampal fissure, and DG. While the majority of the 30 DN divergences for mode 2 were located in the prefrontal cortices (dmPFC; 30.0%) and posterior midline structures (PCC and RSC; 26.6%), a substantial proportion of hits were located in the temporal and posterior cortices. In particular,

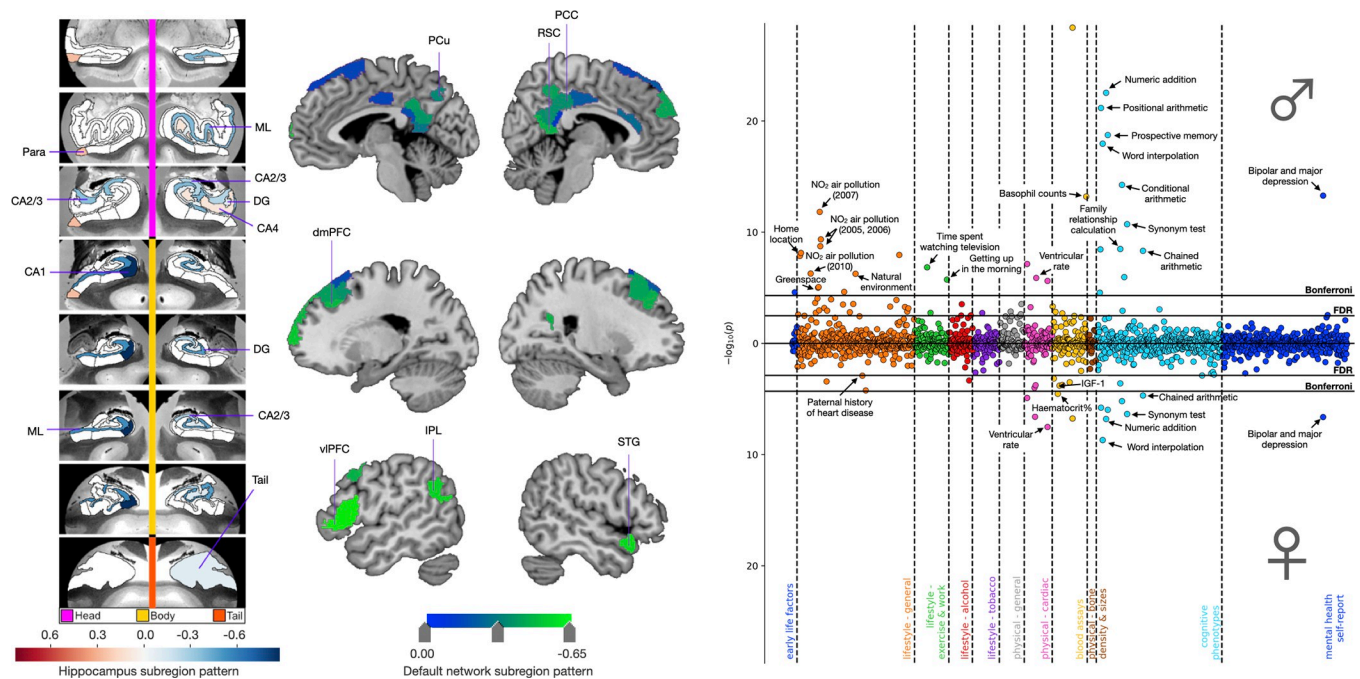


Fig 1. Cognitive, environmental, and cardiovascular phenotypes show sex-specific associations with APOE dosage in the context of mode 1. The leftmost and central panels display structural divergences in the HC and DN, respectively, on mode 1 for the group difference analysis of ADRD family history. We identified 12 HC hits, mostly located in the CA subfields and molecular layer. We also showed 34 DN hits, most of them located in the prefrontal cortex and midline structures. In separate analyses for males ($N = 17,561$) and females ($N = 19,730$), APOE dosage was regressed on HC and DN co-variation patterns from mode 1. We then used these sex-specific models to predict APOE dosage based on inter-individual expressions of mode 1. APOE dosage predicted for each individual was then correlated to 977 UKB phenotypes in separate analyses for males and females. The rightmost panel displays the Miami plot for the correlations between predicted APOE dosage in the context of mode 1 and UKB traits. The upper and lower part of the Miami plot displays the correlations for males and females, respectively. The y-axis indicates negative decimal logarithms for the p -values of each correlation represented by a dot. We highlight important brain-behaviour associations between APOE dosage pooled across subject-specific expressions of mode 1 and verbal-numerical reasoning, supplemented by male-specific correlations with environmental phenotypes. Females showed a specific profile of brain-behaviour associations with cardiovascular phenotypes (e.g., systolic and diastolic blood pressure, IGF-1, and urea) that extended beyond physical traits shared with males (e.g., cardiorespiratory fitness and ventricular and pulse rate). Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer’s disease and related dementia; APOE, Apolipoprotein E; CA, cornu ammonis; DG, granule cell layer of the dentate gyrus; dmPFC, dorsomedial prefrontal cortex; DN, default network; FDR, false discovery rate correction; HC, hippocampus; IPL, inferior parietal lobule; ML, molecular layer; Para, parasubiculum; PCC, posterior cingulate cortex; PCu, precuneus; RSC, retrosplenial cortex; STG, superior temporal gyrus; vlPFC, ventrolateral prefrontal cortex.

<https://doi.org/10.1371/journal.pbio.3001863.g001>

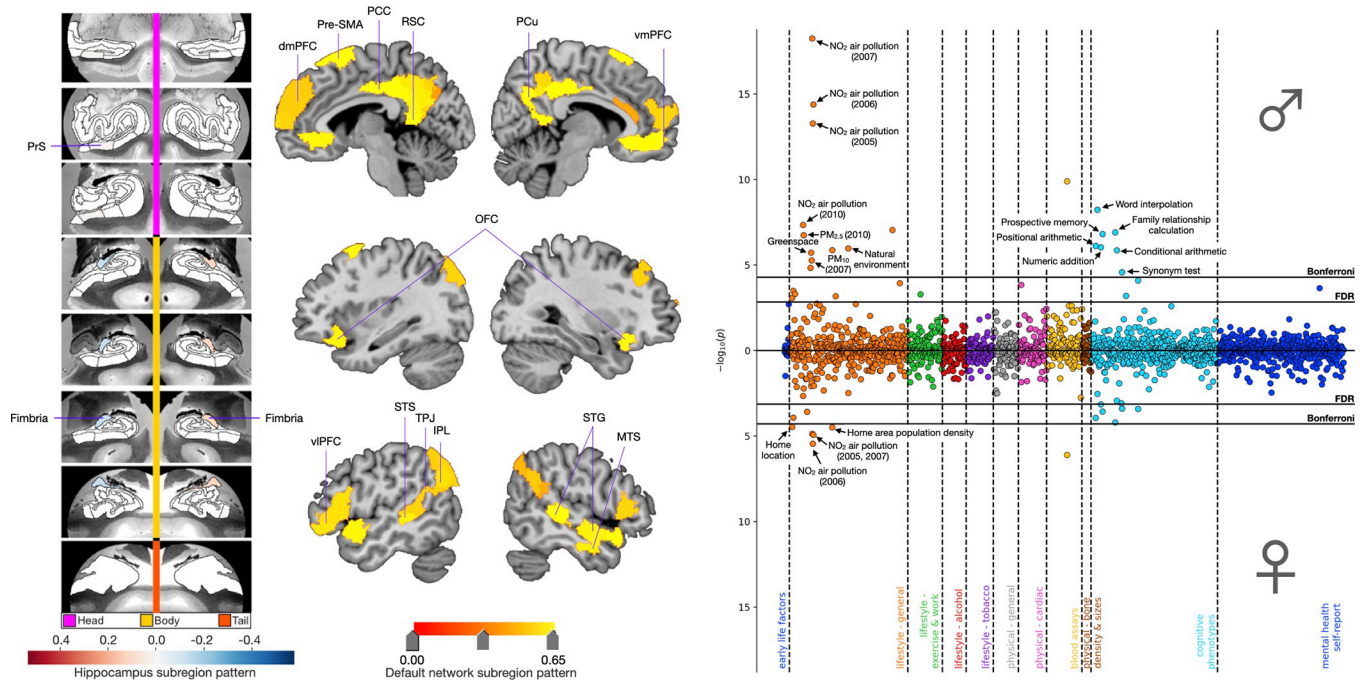


Fig 2. APOE associations for mode 3 revealed a prominence of cognitive and environmental phenotypes in males. Shown here are ADRD-related subregion divergences for mode 3 for the HC (leftmost panel) and DN (central panel). We identified focalized hits to the fimbria and presubiculum with corresponding grey matter differences across the whole DN. In males and females separately, we regressed APOE dosage on HC and DN co-variation patterns from mode 3. We then used these sex-specific models to predict APOE dosage based on inter-individual expressions of mode 3. The rightmost panel displays the Miami plot for the correlations between APOE scores in the context of mode 3 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We highlighted significant associations with environmental phenotypes that were again more prominent in males than females. We additionally showed significant correlations with the fluid intelligence battery that were male specific. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer’s disease and related dementia; APOE, Apolipoprotein E; dmPFC, dorsomedial prefrontal cortex; FDR, false discovery rate correction; IPL, inferior parietal lobe; MTS, middle temporal sulcus; OFC, orbitofrontal cortex; PCC, posterior cingulate cortex; Pre-SMA, pre-supplementary motor area; PrS, presubiculum; PCu, precuneus; RSC, retrosplenial cortex; STG, superior temporal gyrus; STS, superior temporal sulcus; TPJ, temporoparietal junction; vIPFC, ventrolateral prefrontal cortex; vmPFC, ventromedial prefrontal cortex.

<https://doi.org/10.1371/journal.pbio.3001863.g002>

23.3% of DN divergences for mode 2 were located in the temporal cortices (superior temporal sulcus (STS), middle temporal sulcus (MTS), and temporal pole) compared to 20.0% to the left posterior cortex (inferior parietal lobule (IPL) and superior parietal lobule (SPL)). Mode 3 in turn showed 3 statistically relevant HC hits to the fornix’s fimbria and presubiculum, in concordance with 56 DN divergences (Fig 2). Of the DN hits identified for mode 3, 35.7% were located in the frontal lobe (dmPFC, ventromedial prefrontal cortex (vmPFC), vIPFC, pre-supplementary motor area (pre-SMA), and OFC), 30.3% to posterior midline structures (PCC, RSC, and PCu), 17.9% to the temporal cortices (STS, MTS, and superior temporal gyrus (STG)), and 16.1% to the parietal cortices (IPL, SPL, and temporoparietal junction (TPJ)). A minority of the modes only showed HC hits, either located in the fimbria (mode 8; Fig 3) or in the hippocampus–amygdala transition area (modes 6 and 10; S2 and S3 Figs) without any concomitant DN hits. Inversely, some modes only showed DN divergences in the absence of HC hits. This was the case for mode 4 for which we identified 4 DN hits in the dmPFC (S4 Fig), mode 7 for which 9 DN hits were identified in the PFC (dmPFC and OFC; S5 Fig), mode 11 for which 1 DN hit was identified in the PCC (S6 Fig), and mode 13 for which 1 DN hit was identified in the STS (S7 Fig).

Across HC-DN co-variation signatures, we noted a prominence of HC structural deviation in the CA1, CA2/3, and fimbria for the group analysis of ADRD risk. As for the DN

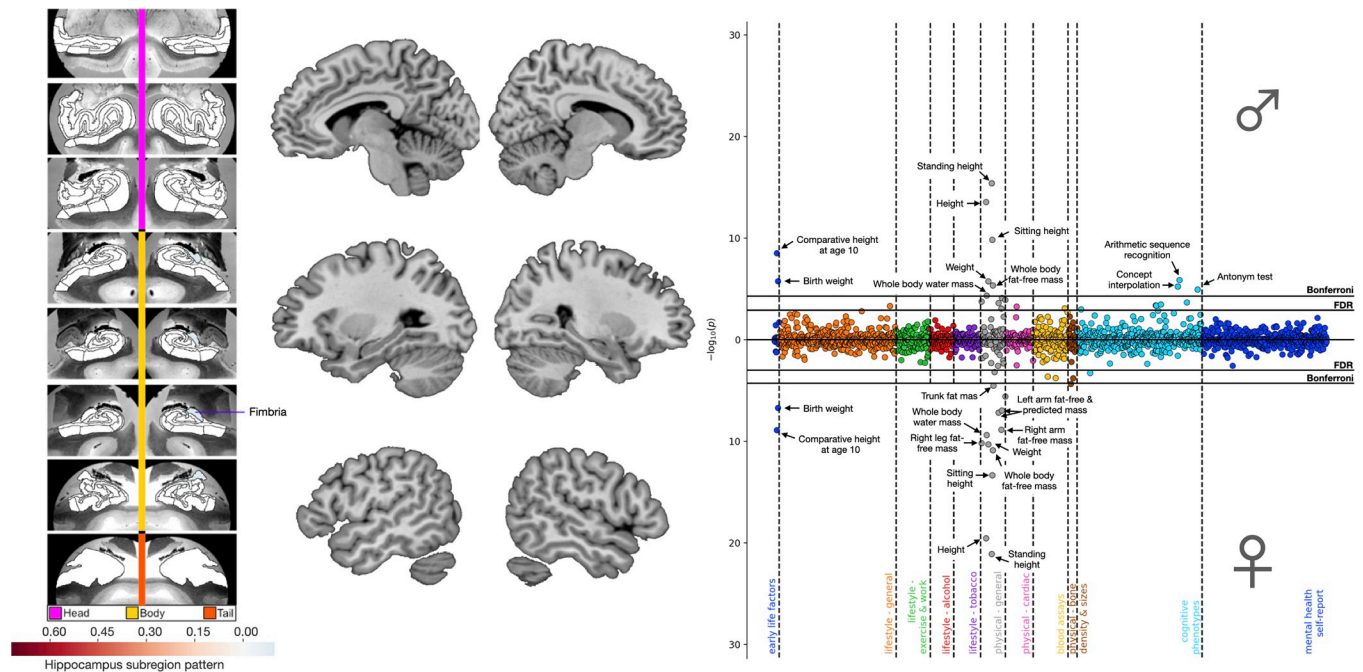


Fig 3. APOE associations for mode 8 linked lipid metabolism to deviation of the fimbria. Shown here are ADRD-related subregion divergences for mode 8 for the HC (leftmost panel) and DN (central panel). We identified a focalized divergence to the fimbria with no corresponding DN hits. In males and females separately, we regressed *APOE* dosage on HC and DN co-variation patterns from mode 8. We then used these sex-specific models to predict *APOE* dosage based on inter-individual expressions of mode 8. The rightmost panel displays the Miami plot for the correlations between *APOE* scores in the context of mode 8 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We show associations with phenotypes related to lipid metabolism and height, supplemented by male-specific associations with the fluid intelligence battery. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer's disease and related dementia; *APOE*, Apolipoprotein E; DN, default network; FDR, false discovery rate correction; HC, hippocampus.

<https://doi.org/10.1371/journal.pbio.3001863.g003>

divergences, we highlighted a constellation of structural deviations involving the prefrontal cortices and posterior midline structures. Modes 1 and 2 showed the highest relative numbers of HC divergences (i.e., 12 and 10 hits, respectively) as compared to any other modes. While the third signature of HC-DN co-variation only showed 3 statistically relevant HC hits, it showed the highest relative number of DN divergences. Together with mode 8, the focalized divergences found in the fimbria for mode 3 highlighted the importance of the fornix in ADRD risk. We further uncovered concomitant structural divergences in HC and DN subregions with known direct anatomical connections in macaque monkeys, such as the presubiculum with RSC [16] and molecular layer with OFC/vmPFC [17]. Ultimately, we revealed an intertwined collection of structural divergences in highly coupled HC and DN subregions that have been linked to ADRD risk and progression by previous research, such as the CA1, CA2/3, presubiculum, and the fornix's fimbria [13,25–28], as well as the dlPFC, OFC, PCC, and PCu [29–32].

Phenome-wide fingerprints of brain-behaviour associations uncover sex-specificity in ADRD risk

We next conducted a phenome-wide analysis to systematise domains of UKB traits in terms of their association with HC-DN signatures and ADRD risk. To quantify genetic risk, we created a bivariate dosage scale that tested for the opposing effects of *APOE* $\epsilon 2$, often suspected to confer protective benefits [33], and $\epsilon 4$, classically believed to escalate dementia risk [4]. We fitted linear regression models to relate inter-individual expressions of HC-DN co-variation from

the 25 signatures to subject-level *APOE* $\epsilon 2$ versus $\epsilon 4$ dosage. Subject-level *APOE* dosage was predicted from a collection of HC-DN signatures using these linear models and subsequently tested against 977 curated UKB phenotypes in a phenome-wide assay conducted separately in males and females. Only the top 3 modes with the most brain-behaviour associations across sexes, i.e., modes 1, 3, and 8, are presented below (Figs 1–3). The phenome-wide profiles for each of the remaining modes with statistically defensible deviations with respect to family history of ADRD are available as part of the online Supporting information (S1–S7 Figs).

The phenome-wide profile for mode 1 highlighted brain-behaviour associations with cognitive traits in addition to male-specific correlations with environmental phenotypes (Fig 1). After carrying out Bonferroni's correction for multiple comparisons, *APOE* dosage pooled across subject-specific expressions of mode 1 yielded 31 and 13 significant associations in males and females, respectively. Cognitive traits represented 35.5% of significant mode-trait associations in males and 53.8% of those identified in females. Baseline cognitive performance on the fluid intelligence battery accounted for most of the cognitive associations, with 7 questions yielding significant associations in males compared to 6 in females. Significant associations with baseline prospective memory were also identified for both sexes, measured as the correct recalling of the object previously shown to participants on the screen. The phenome-wide profiles for both sexes also included ventricular rate on electrocardiogram measured at rest, the completion status of electrocardiogram during exercise, and bipolar and major depression status. At the more lenient false discovery rate (FDR) correction, we observed additional phenotypes linked with erythrocytes count for both sexes. The second most dominant sets of associations for mode 1 centred on environmental phenotypes, such as NO₂ exposure, natural environment, and greenspace, representing 29.0% of significant mode-trait correlations identified in males. Other male-specific associations included lifestyle (time spent watching television and difficulty waking up in the morning) and physiological (hand grip strength, arm mass, and height) phenotypes. At the more lenient FDR correction, males showed additional brain-behaviour associations including exposure to particulate matter of 2.5 μm and 10 μm or less in diameter (PM_{2.5} and PM₁₀). After applying Bonferroni's correction, females showed unique associations with diastolic blood pressure and hematocrit percentage. When applying FDR correction, additional cardiovascular phenotypes showed significant associations in females, such as a paternal history of heart attack, systolic blood pressure, insulin-like growth factor 1 (IGF-1), and haemoglobin concentration. In sum, our phenotypical profiling assay highlighted important phenome-wide associations between *APOE* dosage pooled across subject-specific expressions of mode 1 and verbal-numerical reasoning, supplemented by male-specific correlations with environmental phenotypes. Females instead showed a specific profile of brain-behaviour associations with cardiovascular phenotypes that extended beyond physical traits shared with males.

In the phenome-wide profile for mode 3, we uncovered brain-behaviour associations with cognitive and environmental phenotypes, again more prominent in males than females (Fig 2). After Bonferroni's correction, *APOE* dosage in the context of mode 3 expressions yielded 19 and 6 significant mode-trait associations in males and females, respectively. Environmental phenotypes represented 52.6% of significant associations in males and 83.3% of those identified in females. Significant associations with NO₂ exposure and home area population density were observed for both sexes. Males also showed significant associations with baseline cognitive performance on 6 questions from the fluid intelligence battery as well as with baseline prospective memory. Females did not show significant associations beyond those shared with males, with the exception of home location. At the more lenient FDR correction, females showed additional associations with prospective memory and baseline cognitive performance on 5 questions from the fluid intelligence battery. As such, *APOE* dosage pooled across

subject-specific expressions of mode 3 allowed us to uncover a rich portfolio of associations with environmental and cognitive phenotypes that were more robust in males than females.

In comparison to the overlapping portfolio of brain-behaviour associations derived from modes 1 and 3, the phenome-wide profile for mode 8 emphasised a unique set of physiological phenotypes (Fig 3). After Bonferroni's correction, *APOE* dosage pooled across subject-specific expressions of mode 8 yielded 11 and 15 significant mode-trait associations in males and females, respectively. Physical phenotypes related to body mass and height represented 55.5% of significant correlations in males and 80.0% of those identified in females. After Bonferroni's correction, males showed significant associations with cognitive performance on 3 questions from the fluid intelligence battery assessed in the online follow-up. At the more lenient FDR correction, males showed further associations with cognitive performance on 2 additional questions from the fluid intelligence battery and with the maximum number of digits remembered correctly on the numeric memory test, both assessed in the online follow-up. After Bonferroni's correction, females showed significant associations with trunk fat mass and heel bone mineral density. In sum, we highlighted important phenome-wide associations between *APOE* dosage pooled across subject-specific expressions of mode 8 and proxies of cardiovascular health, supplemented by male-specific correlations with cognitive phenotypes. A formal assessment of the difference in associations between males and females for the 3 modes with the most brain-phenotypic associations across sexes (i.e., modes 1, 3, and 8) is presented in the Supporting information (S8–S10 Figs) and serves as a complement to their respective Miami plots (Figs 1–3) (cf. methods). The phenome-wide profiles derived across these 3 concomitant regimes of HC-DN co-variation emphasised sex differences in ADRD risk, with recurring associations with air pollution and verbal-numerical reasoning that were more prominent in males than females.

***APOE* gene variants are associated with distinct clusters of risk-anatomy links**

We next examined ADRD-specific clusters of risk-anatomy links across each unique *APOE* gene variant (i.e., $\epsilon 2/2$, $\epsilon 2/3$, $\epsilon 3/3$, $\epsilon 2/4$, $\epsilon 3/4$, and $\epsilon 4/4$). We computed the interactions between the subject-specific expressions of HC-DN co-variation modes (canonical variates) and each *APOE* genotype (encoded as binary variables, such that participants who do not carry a given genotype were zeroed out). In doing so, we obtained 6 new population-wide indices, 1 for each distinct *APOE* genotype that we correlated, using Spearman's coefficient, with 63 curated ADRD risk factors (a phenotype collection used previously [34]). We then performed an agglomerative clustering analysis that consisted of a nested merging of correlation coefficients with similar variance until all observations merged in a single cluster. The ensuing dendrograms indicated the distance between each cluster identified when retaining 3 levels of branching (Fig 4). A formal metric of statistical agreement between cluster models was provided as part of supplementary analyses (S11 Fig).

Our integrated analysis of risk-anatomy links showed the relatively early branching of social engagement phenotypes for $\epsilon 2/2$ (e.g., being a full- or part-time student and doing unpaid or voluntary work), $\epsilon 2/3$ (e.g., number of full siblings, looking after one's home or family, family relationship satisfaction, and number of people in household), $\epsilon 3/4$ (e.g., number of full siblings), and $\epsilon 4/4$ (e.g., being a full- or part-time student, attending adult education classes, retirement, family relationship satisfaction, lack of social support, and friendships satisfaction) genotypes. The relevance of social engagement phenotypes across most *APOE* gene variants suggests that the contribution of social behaviours to risk-anatomy links transcend genetic risk. $\epsilon 3$ carriership was characterised by the early branching of socioeconomic determinants as

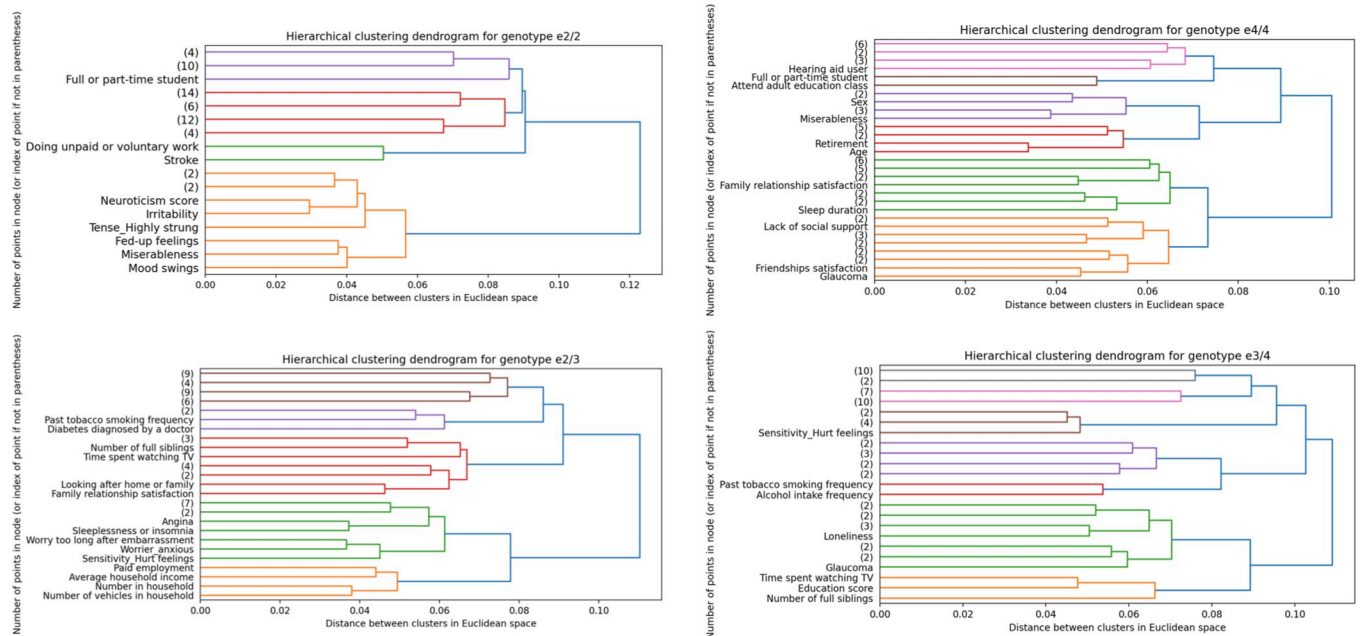


Fig 4. Neuroticism-related phenotypes show unique risk-anatomy links in $\epsilon 2$ carriers. To test for risk-anatomy links, we computed the Spearman's correlations between the population-wide HC and DN co-variation patterns, multiplied by each of the 6 *APOE* genotypes and the 63 preselected Alzheimer's disease risk factors [34]. We performed an agglomerative clustering analysis on these Spearman's correlations, which consists of repeatedly merging Spearman's correlations with similar variance until all observations are merged into a single cluster. Here are shown the dendrograms that indicate the distance between each cluster identified when retaining 3 levels of branching for *APOE* $\epsilon 2/2$ (upper left; $N = 217$), $\epsilon 2/3$ (lower left; $N = 4,625$), $\epsilon 4/4$ (upper right; $N = 822$), $\epsilon 3/4$ (lower right; $N = 8,613$). The dendrograms for $\epsilon 3/3$ and $\epsilon 2/4$ can be found in the Supporting information (S12 Fig). We showed the early emergence of social engagement phenotypes (e.g., doing unpaid or voluntary work, attending adult education classes, family relationship satisfaction, number of people in household, and number of full siblings) across the different *APOE* gene variants suggesting that the contribution of social behaviours to risk-anatomy links transcend genetic risk. $\epsilon 3$ carriership was characterised by the early branching of socioeconomic determinants (e.g., paid employment, average household income, number of vehicles in the household, time spent watching TV, and education score) as shown on the dendrograms for $\epsilon 2/3$, $\epsilon 3/4$, and $\epsilon 3/3$ (S12 Fig). While clusters of social engagement and socioeconomic determinants were shared across different *APOE* genotypes, we found that neuroticism was uniquely associated with $\epsilon 2$ carriership. Indeed, the dendrogram for $\epsilon 2/2$, $\epsilon 2/3$, and $\epsilon 2/4$ (S12 Fig) showed an early emerging cluster of neuroticism-related phenotypes (e.g., irritability, miserableness, being worried/anxious). This personality cluster was especially apparent for $\epsilon 2$ homozygotes, as reflected by the relatively high Euclidean distance of the first branching that split the neuroticism-related phenotypes from the rest of the risk factors. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/clustering_analysis (DOI: 10.5281/zenodo.7126809). *APOE*, Apolipoprotein E; DN, default network; HC, hippocampus.

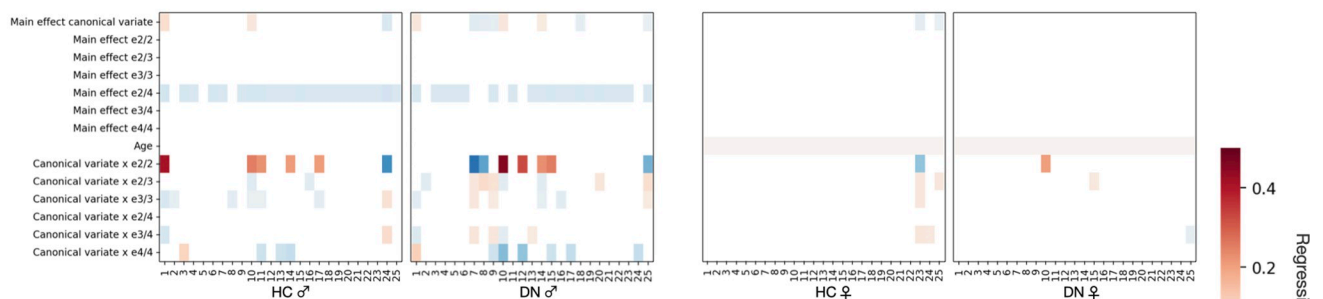
<https://doi.org/10.1371/journal.pbio.3001863.g004>

shown on the dendrograms for $\epsilon 2/3$ (e.g., past tobacco smoking frequency, time spent watching television, paid employment, average household income, and the number of vehicles in the household), $\epsilon 3/4$ (e.g., past tobacco smoking frequency, alcohol intake frequency, time spent watching television, and education score), and $\epsilon 3/3$ (time spent watching television, education score, past and current tobacco smoking frequency, alcohol consumption on a typical drinking day and alcohol intake frequency; see S12 Fig). We noted the early emergence of a personality cluster in $\epsilon 2$ carriers that comprised self-reported traits related to neuroticism as shown on the dendrograms for $\epsilon 2/2$ (e.g., irritability, miserableness, mood swings), $\epsilon 2/3$ (e.g., being worried/anxious and easily hurt), and $\epsilon 2/4$ (e.g., being worried/anxious, mood swings, and miserableness; see S12 Fig). All these personality traits have been identified as neurotic behaviour domains and are part of the neuroticism battery of the UKB (UKB data field 20127). We thus uncovered that neuroticism, which is known to be closely linked to loneliness [35], is a personality trait that shows distinct patterns of association with HC-DN co-variation expressions in $\epsilon 2$ carriers.

Sex-specific dependencies between *APOE* gene variants and signatures of HC-DN co-variation explain ADRD risk

We next directed attention to sex-specific interactions between HC-DN co-variation regimes and *APOE* genotype status that would explain inter-individual differences in ADRD risk. To this end, we tested whether HC-DN signatures systematically interacted with specific *APOE* genotypes in explaining variation in a collection of 63 ADRD risk factors (cf. above). More formally, each risk factor was individually regressed on the subject-specific expressions of HC and DN patterns for each of the 25 modes. This analysis step hence supplied 50 estimated linear models per target risk factor. Each model took as input variables the main effect of the HC or DN pattern expressions, the main effects of the 6 *APOE* genotypes, and the interaction between each *APOE* genotype with the HC or DN pattern, controlling for age. Separate analyses were carried out in the male (Fig 5, leftmost panels) and female (Fig 5, rightmost panels) subgroups of our UKB cohort. To ascertain the robustness of our findings, we compared each coefficient estimate against empirically data-derived null distributions obtained through a rigorous permutation procedure (i.e., label shuffling permutation). We only interpreted the model coefficients that emerged as statistically relevant based on the respective null distributions at 95% confidence.

Interactions of HC and DN co-variation patterns with *APOE* genotypes in estimating **doing unpaid or voluntary work** in males and females



Interactions of HC and DN co-variation patterns with *APOE* genotypes in estimating **engaging in strenuous sports** in males and females

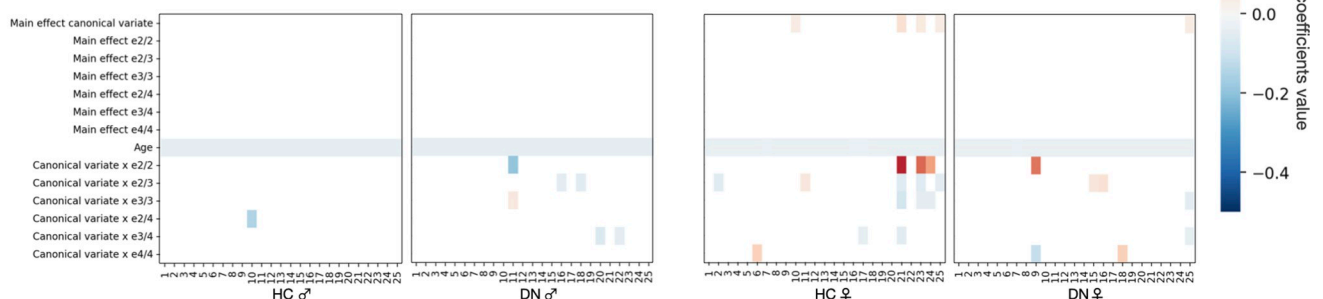


Fig 5. Brain-*APOE* $\epsilon 2/2$ interaction explains variance in social lifestyle in males and physical activity in females. We tested whether HC-DN signatures interacted with *APOE* genotypes in explaining variance on the 63 preselected ADRD risk factors. Separate analyses were run for males (leftmost plots) and females (rightmost plots). Each column on the heat maps represents the coefficients for a single linear regression model. The first 25 columns show the coefficients for HC patterns, whereas the last 25 columns show the coefficients for DN patterns. We assessed the robustness of our findings by comparing each coefficient to empirically built null distributions obtained through permutation testing. Only the coefficients that were statistically different from their respective null distributions 95% of the time are presented. We displayed the modifiable risk factors for which the strongest brain-*APOE* interactions were observed. In the top panels, we show that *APOE* $\epsilon 2/2$ preferentially interacts with HC and DN canonical variates in estimating doing unpaid or voluntary work in males. In the bottom panels, we show that *APOE* $\epsilon 2/2$ interacts with selective HC and DN canonical variates in estimating engaging in strenuous sport in females. We have thus shown that *APOE* $\epsilon 2/2$ interacts preferentially with HC-DN co-variation patterns in estimating social lifestyle in males and physical activity in females. These interactions profiles suggest that $\epsilon 2$, and not $\epsilon 4$, is driving most of the brain-*APOE* interactions in healthy individuals at risk of developing ADRD with a substantial level of sex-specificity. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/fig_5 (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer's disease and related dementia; *APOE*, Apolipoprotein E; DN, default network; HC, hippocampus.

<https://doi.org/10.1371/journal.pbio.3001863.g005>

In a comprehensive set of analyses across 63 ADRD risk factors, we identified the strongest nonlinear interaction effects in homozygote $\epsilon 2$ carriers. Notably, brain-*APOE* interactions accounted for more variance in several modifiable social and cardiovascular risk factors than did the main effects of *APOE* $\epsilon 2$ and $\epsilon 4$. Male $\epsilon 2$ homozygotes showed strong interactions with HC and DN pattern expressions for doing unpaid or voluntary work. In parallel, female $\epsilon 2$ homozygotes showed strong interactions with HC-DN pattern expressions for engagement in strenuous sports. Across the different domains of risk factors investigated, we singled out brain-*APOE* interactions specific to $\epsilon 2$ homozygotes that were not identifiable in heterozygotes and non-carriers. While we observed no appreciable sex effect for the interaction of *APOE* $\epsilon 4/4$ and HC-DN co-variation expressions, we found defensible sex-specificity for the role of *APOE* $\epsilon 2/2$. More precisely, we showed strong interactions between *APOE* $\epsilon 2/2$ and HC-DN co-variation patterns for social lifestyle factors in males and physical activity factors in females. Through our analyses of a variety of risk factors, we have thus isolated brain-*APOE* interactions unique to $\epsilon 2$ carriers that depend on sex.

After examining target risk factors, we next put to the test whether expressions of HC-DN signatures bear relations with *APOE* genotypes in explaining ADRD risk. In dedicated analyses for males (Fig 6, upper panels) and females (Fig 6, lower panels), family history of ADRD was regressed on a single HC or DN pattern, resulting in 50 different linear models per sex. Each such model was fed as input variables the main effect of the HC or DN pattern, the main effects of the *APOE* genotypes, and the interactions between each *APOE* genotype and the HC or DN pattern, controlling for age. We assessed the robustness of our findings by comparing each coefficient to empirically built null distributions obtained through permutation testing (cf. above). We focused interpretation on the model coefficients that were statistically robust against their respective null distributions at 95% confidence. We found no statistically relevant main effect of *APOE* $\epsilon 2/2$ on ADRD risk among males. For *APOE* $\epsilon 2/3$ and $\epsilon 3/3$ carriers, we found similar effects on ADRD risk in males, lowering the odds of ADRD family history by approximately 30% across the different models investigated. Likewise, *APOE* $\epsilon 2/4$ and $\epsilon 3/4$ carriers showed similar effects in tracking ADRD risk in males, elevating the odds of ADRD family history by more than 20% on average. As expected from the literature, *APOE* $\epsilon 4/4$ increased the odds of ADRD family history by more than 56% in males across the different models investigated. In females, *APOE* $\epsilon 2/2$ status decreased the odds of ADRD family history by 50% on average, while $\epsilon 2/3$ and $\epsilon 3/3$ status led to decreases of approximately 25% and 17%, respectively. In contrast, *APOE* $\epsilon 3/4$ and $\epsilon 4/4$ status lifted the odds of ADRD family history by approximately 35% and 86%, respectively. Among females, *APOE* $\epsilon 2/4$ carriers were associated with dampened ADRD risk relative to *APOE* $\epsilon 3/4$ carriers. The odds of ADRD family history associated with *APOE* $\epsilon 2/4$ were only increased by 24% in females. This approximately 10% reduction in ADRD risk, uniquely observed among females, could be taken to suggest that $\epsilon 2$ can still be protective against ADRD risk in the presence of an $\epsilon 4$ allele. Females also showed some strong brain-*APOE* interactions above and beyond the well-established risk and protective effects associated with each *APOE* genotype. Notably, the interaction of mode 9 DN pattern expressions with *APOE* $\epsilon 2/2$ status was associated with a 2-fold increase in ADRD risk. It was considerably stronger than the main risk effect conferred by *APOE* $\epsilon 4/4$. This strong interaction effect can be taken to suggest that HC-DN co-variation plays a chief role in ADRD risk, which might have been overlooked by previous analyses restricted to genetic data. In sum, we identified and annotated sex-specificity in the opposing effects of $\epsilon 2$ and $\epsilon 4$ on ADRD risk, with demonstrably stronger brain-*APOE* interactions among females.

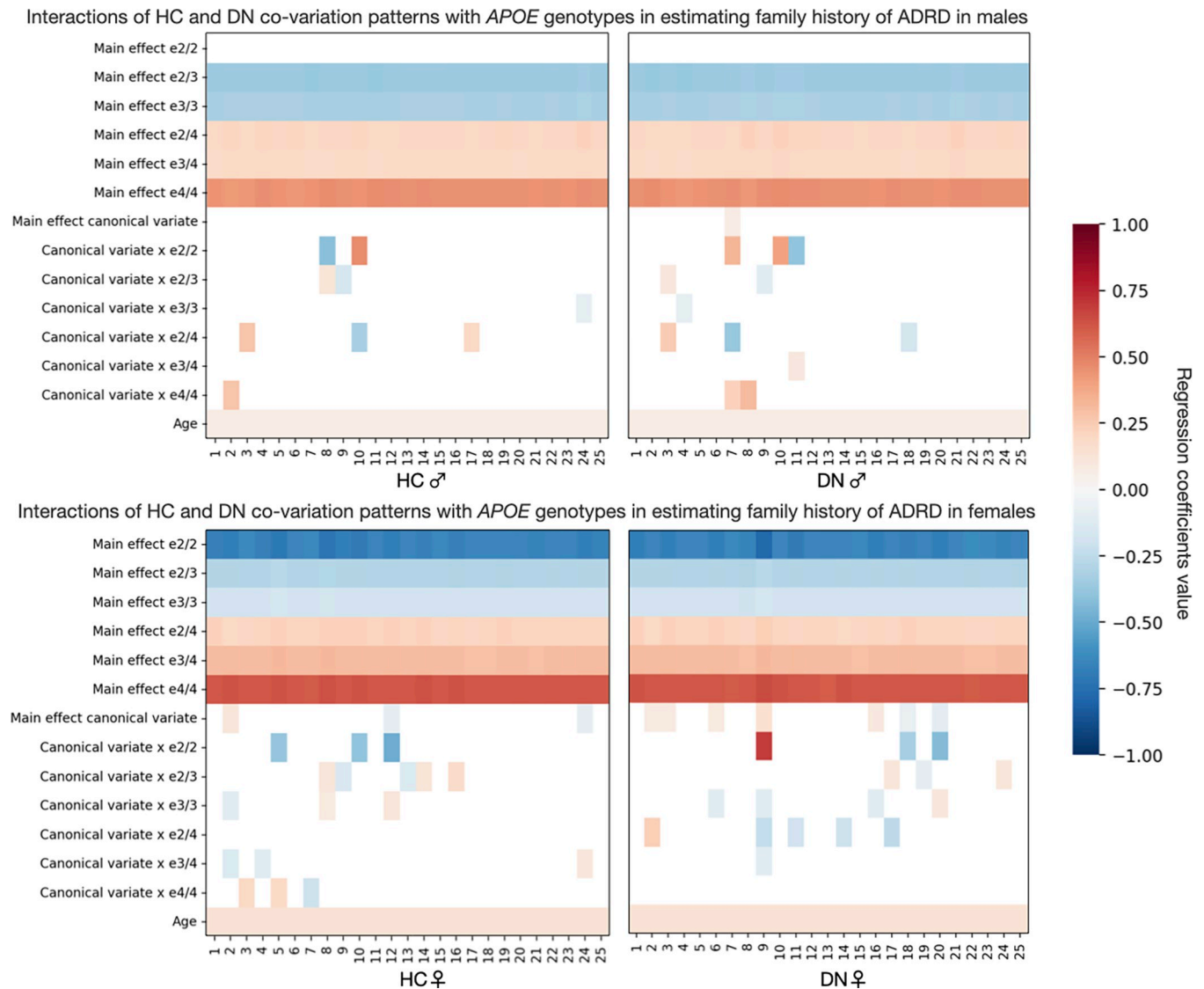


Fig 6. The protectiveness of $\epsilon 2$ is sex-dependent and modulated by HC-DN co-variation patterns. In separate analyses for males and females, we tested whether HC-DN signatures interacted with *APOE* genotypes in explaining variance in family history of ADRD. Separate analyses were run for males (higher plots) and females (lower plots). Each column on the heat maps represents the coefficients for a single linear regression model. The first 25 columns show the coefficients for HC patterns, whereas the last 25 columns show the coefficients for DN patterns. We assessed the robustness of our findings by comparing each coefficient to empirically built null distributions obtained through permutation testing. Only the coefficients that were statistically different from their respective null distributions 95% of the time are presented. We found that the main effect of *APOE* $\epsilon 2/2$ against ADRD risk was only statistically robust in females. We also showed a spectrum in the opposing effects of $\epsilon 2$ and $\epsilon 4$ among females, such that $\epsilon 2/4$ was associated with a lower increase in ADRD risk than did *APOE* $\epsilon 3/4$, which in turn was associated with lesser risk than $\epsilon 4/4$. We further found that the protectiveness of *APOE* $\epsilon 2/2$ interacts with brain structure and can even lead to an increase in ADRD risk among females with a strong expression of mode 9. These interactions profiles suggest that the protectiveness of $\epsilon 2/2$ is not only sex-specific but also modulated by HC-DN co-variation expressions. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/fig_6 (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)). ADRD, Alzheimer's disease and related dementia; *APOE*, Apolipoprotein E; DN, default network; HC, hippocampus.

<https://doi.org/10.1371/journal.pbio.3001863.g006>

Dominant principles of brain-behaviour associations uncovered a male-specific link with neuroticism

In a final suite of analyses, we conducted an exploratory principal component analysis (PCA) to disentangle the major sources of brain-behaviour variation in our UKB cohort. We first computed Pearson's correlations between the 25 pairs of expressions (i.e., canonical variates)

from the HC and those from the DN patterns and the 63 preselected ADRD risk factors. This step yielded 3,150 distinct coefficients represented by a risk by canonical variates matrix ($X_{63 \times 50}$). We then carried out a PCA to reduce the dimensionality to 3 major axes of brain-behaviour variation. These explained approximately 13.8%, approximately 9.6%, and approximately 8.2% of the total variance in the cross-correlation matrix, respectively (S13 Fig).

The leading axis of variation highlighted social phenotypes previously singled out in the clustering analysis (e.g., attending religious groups, attending adult education classes, and the number of people in the household). We also observed a strong expression of socioeconomic determinants among the first axis of brain-behaviour associations (e.g., age completed high school education, average household income, paid employment, and the number of vehicles in the household). The second most important axis mainly emphasised health-related phenotypes (e.g., stroke, hypertension, and diabetes) and lifestyle factors (e.g., alcohol intake frequency, difficulty getting up in the morning, being a morning person, and sleeplessness). The third most explanatory axis tracked neuroticism and its associated personality trait indicators (being worried/anxious, being easily hurt, and worrying too long after embarrassment) from the rest of the risk factors. We again emphasised the importance of social factors on HC-DN co-variation expressions along with other socioeconomic and lifestyle behaviours.

To certify the robustness of our findings, we performed a split-half reliability assessment of our principal component solution across 1,000 bootstrap iterations. At each iteration, we drew 37,291 participants with replacements to simulate random participant samples that we could have pulled from the same population. We then randomly split the sample in half to create 2 analogous subsets. We computed the Pearson's correlation between possible pairs of the 50 canonical variates and 63 phenotypes across participants for each random subset. We then estimated 2 PCA models in parallel, 1 for each random half subset, on the z-scored correlation coefficients matrices (63 phenotypes \times 50 canonical variates). We showed the average projection of each Pearson's correlation coefficient on the 3 principal axes of brain-behaviour associations across the 1,000 iterations. We found that the projections of the risk-anatomy link on component 1 were robust. While of lesser strength than the first axis of brain-behaviour associations, the projections for components 2 and 3 are reminiscent of the original analysis. In particular, neuroticism-related personality traits are distinctly expressed on the third axis of brain-behaviour associations, as was found in our original analysis (S14 Fig). A formal account of statistical agreement between both subsets was provided as part of the Supporting information (S15 Fig).

We then repeated the identical pattern-learning workflow sex-stratifying in males and females separately. The top 3 principal components explained approximately 12.1%, approximately 9.9%, and approximately 9.0% of the total variance in males, and approximately 13.1%, approximately 9.5%, and approximately 7.3% of the total variance in females (S16 Fig). The first axis of brain-behaviour associations was roughly the same in males and females as in our original analysis. In fact, the same set of social phenotypes was emphasised on component 1 (e.g., attending religious groups, attending adult education classes, and the number of people in the household) for both sexes. In contrast, component 2 separated neuroticism-related items (e.g., miserableness, fed-up feelings, mood swings, and being worrier/anxious) from the rest of the risk factors in males only. The fact that the neuroticism-related component was the second most important axis of brain-behaviour associations in males but was found in third place on the whole population-derived PCA suggests that the association between neuroticism-related phenotypes and HC-DN co-variation expressions was most important in males. Lastly, the third axis of brain-behaviour associations emphasised different categories of risk factors in males and females. The male-derived component 3 emphasised socioeconomic determinants (e.g., education score and the number of vehicles in the household). In contrast,

the female-derived one emphasised lifestyle risk factors (e.g., alcohol intake frequency, alcohol consumption on a typical drinking day, and past tobacco smoking frequency). Our sex-specific analysis hence revealed that the first and most robust axis of brain–behaviour associations was shared across sexes, whereas the second and third axis emphasised sexually dimorphic groups of risk factors.

We performed a bootstrap analysis of the sex-specific PCA solutions to assess the robustness of our findings. Across 1,000 bootstrap iterations, we drew 17,561 males and 19,730 females with replacements to simulate random participant samples that we could have gotten. At each iteration, we computed the Pearson’s correlation between possible pairs of the 50 canonical variates and the remaining 62 phenotypes (as sex was used for grouping) across males and females separately. We then estimated 2 PCA models in parallel, 1 for each sex, on the z-scored correlation coefficients matrices (62 phenotypes \times 50 canonical variates). A formal assessment of statistical agreement in the PCA solutions between both sexes was performed (S17 Fig). We observed a low agreement between the male- and female-derived PCA solutions, thus emphasising the sex-specificity of our derived brain–behaviour axes.

External validation

To externally validate our discovered associations between HC-DN co-variation signatures and ADRD risk factors, we have investigated whether our UKB-derived population signatures of HC-DN co-variation successfully track ADRD-related variation in unseen participants from an independent sample. We capitalised on the openly available PResymptomatic EVAluation of Experimental or Novel Treatments for Alzheimer’s disease (AD) (PREVENT-AD) dataset, one of the largest single-site prospective cohorts of pre-symptomatic individuals with a family history of Alzheimer’s disease. Our final sample included image-derived phenotypes of grey matter morphology and *APOE* SNP genotyping from 318 participants, totaling data from 799 visits. For each visit, we computed the level of expression of each of the 25 HC-DN co-variation signatures, from the UKB, for a participant from PREVENT-AD (cf. methods). To test whether distinct derived modes of HC-DN co-variation track distinct aspects of ADRD-related behaviours in unseen participants, we correlated the individual expressions of the 25 modes, represented by pairs of latent expressions of the UKB-derived brain signatures for the HC and DN sides, with a collection of 157 widely established indicators of ADRD progression (e.g., cerebrospinal fluid and blood biochemistry, cognitive and neurosensory evaluations, and health and demographic profile). We assessed the Pearson’s correlations through permutation testing. We reported only the coefficients that were robustly different from the derived empirical null distribution in at least 95% of the 1,000 permutation iterations (S18 Fig).

We found that the several categories of risk factors that emerged in the phenome-wide profiling in the UKB dataset were also flagged in the PREVENT-AD dataset. For example, we have corroborated a link between individual expressions of mode 1 in PREVENT-AD participants and depression, a phenotype that emerged as statistically significant in the phenome-wide profiling for mode 1 for males and females in the UKB. Similarly, we have replicated associations between mode 2 and verbal-numerical reasoning by linking mode 2 expressions in PREVENT-AD participants to several measures of language fluency and working memory highlighted by the Montreal Cognitive Assessment (MoCA) and Repeatable Battery for Assessment of Neuropsychological Status (RBANS), respectively. The MoCA is a cognitive screening tool specially designed to track mild cognitive impairment [35]. Performance on the MoCA has previously been associated with grey matter volumes in subregions of the HC, including the HATA, in middle-aged patients with diabetes [36]. Looking at the individual expressions of mode 6 in PREVENT-AD subjects, we found robust ties of several sub-items of the MoCa

(e.g., attention, subtraction, and language fluency) with the same HC-DN population signature that also showed HATA-specific divergence in the UKB participants.

The phenome-wide profiling for mode 6 further highlighted several indicators of vascular integrity (e.g., carotid intima-media thickness)—a cue to cardiovascular system implication that also emerged in PREVENT-AD participants as reflected by a correlation between mode 6 (on the HC side) and atrial fibrillation. Similarly, the phenome-wide profiling for mode 8 in the UKB highlighted several phenotypes related to body mass, while the expression of mode 8 in PREVENT-AD participants was related to arthritis, a joint disorder worsened by age and weight. In addition to replicating the UKB findings, we found complementarity in the associations between the HC-DN signatures and PREVENT-AD phenotypes such that distinct modes track different domains of ADRD risk. For example, DN variation captured by modes 6 and 8 tracks several global indices of the RBANS, a cognitive battery designed to monitor cognitive decline over time. Notably, only mode 6 tracked the visuospatial dimension of the test, as reflected by correlation with sub-items of the figure drawing tests. Further, only individual expressions of mode 6 in PREVENT-AD participants were also correlated to cognitive performance on the MoCA. These patterns of associations, specific to mode 6, reflect a sensitivity to general cognitive ability in PREVENT-AD participants, who all have a family history of ADRD. We found similar patterns of robust associations to PREVENT-AD phenotypes up to the 25th and last mode of HC-DN co-variation that showed noticeable associations with tau CSF levels on the HC side and cardiovascular factors (e.g., systolic blood pressure, pulse, and *APOE* ϵ 4/ ϵ 4 genotype) on the DN side. We have thus shown that HC-DN signatures robustly track different aspects of ADRD risk in a cohort fully independent from the one in which the co-variation patterns have originally been derived. We have thus corroborated and extended the characterisation of our population-derived limbic-neocortical co-variation signatures by linking them with several known indicators of ADRD risk based on new data.

Discussion

Longstanding research has insisted on the alteration of the DN and HC in early ADRD development (see, for example, [14]). However, brain-imaging investigations seldom had the opportunity to incorporate rare genotypes such as *APOE* ϵ 2/ ϵ 2. At the same time, common epidemiological studies that have reported the protective effect of carrying an ϵ 2 allele are not typically equipped to perform an adequately powered brain-imaging examination at a scale of thousands of people. We overcame several shortcomings by capitalising on *APOE* genotyping and structural brain scans from approximately 40,000 UK Biobank participants. Our mission-tailored analytical framework was specially designed for disentangling ADRD-specific differences in brain structure at the population level. Revisiting ADRD through this lens, we uncovered sex-specific associations between rarely investigated *APOE* gene variants and microstructurally defined HC-DN signatures hardly ever discerned in a prospective human cohort. Our collective findings paint a more concrete picture of the antagonistic effects of *APOE* ϵ 2 and ϵ 4 on population-wide HC-DN signatures, along with their interlocking divergences between men and women.

Epidemiological studies, without access to brain-imaging assessments, have provided evidence suggesting that an ϵ 2 allele typically acts to protect against late-onset Alzheimer's disease [22,33] and against A β accumulation [37–42]. A β accumulation in ϵ 2 carriers could be delayed by 30 to 50 years compared to ϵ 4 carriers, who start showing A β positivity in their early 40s [12,40,43]. The protective qualities of ϵ 2 status have been noted even in the presence of an ϵ 4 allele [12]. Nonetheless, the sex-specific impact of *APOE*, especially its ϵ 2 gene variants, on brain structure could seldom be investigated at the population level. By deriving an envelope

of distinct HC-DN signatures at a fine-grained resolution among thousands of healthy adults, we were able to uncover brain-*APOE* interactions systematically overlooked by traditional brain-imaging studies. Stratifying our population cohort by sex and *APOE* gene variants, we were in a position to conclude that the protective effect of *APOE* $\epsilon 2/2$ on ADRD risk was not statistically robust among males, even in a sample of approximately 20,000 participants. In contrast, we demonstrated a spectrum of $\epsilon 2$ and $\epsilon 4$ effects among females such that *APOE* $\epsilon 2/4$ was associated with milder ADRD risk than $\epsilon 3/4$, which in turn was associated with milder ADRD risk than $\epsilon 4/4$. Resilience towards cognitive decline generally observed among $\epsilon 2$ carriers could arise from relatively higher baseline *APOE* steady-state levels in regions including the HC and frontal cortex as compared to $\epsilon 4$ carriers and $\epsilon 3$ homozygotes [44–47]. Isoform-specific effects related to the *APOE* protein could be further enhanced by microglia-driven homeostatic responses to $A\beta$ accumulation [48,49]. In fact, $\epsilon 2$ carriers are biologically more efficient at scavenging $A\beta$ [50]. As a result, $A\beta$ positivity in $\epsilon 2$ carriers with normal cognition is generally detected in much older age (approximately 95 years) as in $\epsilon 4$ carriers (40 to 55 years) [40]. Older $\epsilon 2$ carriers with amyloid pathology are likewise less likely to be diagnosed with dementia than $\epsilon 3$ homozygotes of the same age [51]. Cell proliferation and survival in the HC are thought to be particularly modulated by estrogens [52–54] that could have a downstream impact on microglial and astrocytic *APOE* synthesis [55]. The presence of an estrogen-dependent enhancer in the promoter region of the *APOE* gene is thus bound to favour female $\epsilon 2$ carriers [56]. These previous elements of evidence are in line with our present finding suggesting that the protective effect of *APOE* $\epsilon 2$ on ADRD risk is sex specific and also unique to particular HC-DN co-variation patterns. Notably, we found that female $\epsilon 2$ homozygotes with a high expression of mode 9 had twice the odds of having a family history of ADRD. We have thus shed light on important nuances in the predominant genetic account of ADRD by questioning the protectiveness of $\epsilon 2$ when placed in relation to sex and brain structure.

We expanded upon the discovered sex differences in ADRD risk by highlighting a female-specific constellation of brain-behaviour associations with cardiovascular traits. As the neuroprotective effect of estrogen weakens with older age, women become more vulnerable to neurovascular disorders that can ultimately lead to dementia [57]. Cardiovascular risk factors that are exacerbated in females following menopause, such as trunk fat mass, have been associated with chronic neuroinflammation and microstructural alteration of the fornix [58,59]—the main output tract from the HC that carries direct neural signals toward partner regions of the midline DN [60]. Building on existing literature, we identified ADRD-related divergences in the fimbria of the fornix in healthy participants for mode 8 that we have linked to selective brain-behaviour associations with proxies of cardiovascular health (e.g., water mass, fat-free mass, and weight). For the same HC-DN signature, we found a female-specific association with trunk fat mass, a correlate of estrogen declines [61]. This observation supports a link between cardiovascular health, female sex, and microstructural alteration of the fornix. Despite the protective effect of *APOE* $\epsilon 2$ against ADRD previously discussed, carrying an $\epsilon 2$ allele has been associated with elevated risks for cardio- and neurovascular disorders [62–66]. *APOE* $\epsilon 2$ is indeed limited in its ability to mediate the vascular clearance of cholesterol metabolites and triglycerides that could in turn precipitate the risks of cholesterol pathologies such as hyperlipoproteinemia and cardiovascular sequelae [67]. The variability of the protective effect of physical activity on dementia risk when stratifying participants by $\epsilon 4$ status might be taken to suggest that *APOE* $\epsilon 2$ is driving the relationship between physical activity and cognitive performance [68–72]. Hypothetically, engaging in physical activity could be particularly beneficial to older female $\epsilon 2$ carriers in counteracting the rising risk of neurovascular complications resulting from the combined effect of *APOE* $\epsilon 2$ and decreased estrogen levels. Bringing support for this claim, we have shown specific interactions between HC-DN signatures and *APOE*

$\epsilon 2/2$ genotype in explaining variation in physical activity—an effect that we found exclusive to females. The specificity of this effect to $\epsilon 2$ homozygotes is consistent with previous findings that have associated $\epsilon 2$ with increased longevity in centenarians [73]. Given that almost 90% of centenarians are females, the sex-specificity of our results is consistent with a genotype-driven behaviour that favours longevity via exercise in female $\epsilon 2$ homozygotes.

Epidemiological studies have provided evidence that traffic-related air pollution and residence near major roadways are associated with decreased cognitive abilities [74–82] and a higher risk of developing dementia [83–92]. The detrimental effect of air pollution on cognition could even be exacerbated in *APOE* $\epsilon 4$ carriers [82,85]. Our phenome-wide assay tied mode 1 expressions to blood markers (e.g., erythrocytes, haemoglobin, and haematocrit) and air pollution. This phenome-wide profiling supports an interplay between environmental stressors, vascular integrity, and dementia. Mode 1 also showed 19 DN hits in the PFC—a sub-region in which vascular and perivascular white matter damage has been specifically observed in humans and canines chronically exposed to high levels of air pollutants [93]. Accumulation of nanoscale particulate matter in endothelium cells, basement membranes, axons, and dendrites coincided with prefrontal white matter damage, which is in line with deficits in the blood–brain barrier [93]. Such pollution-driven prefrontal white matter damage is thought to be intensified in *APOE* $\epsilon 4$ carriers [94]. Autopsy samples from patients with Alzheimer’s disease have further shown reduced pericyte coverage in CA1 and PFC (Brodmann area 9/10). These were 2 subregions in which we showed ADRD-related structural divergences in mode 1, as compared to healthy blood vessels in controls [95]. Evidence suggests that alteration of pericytes in cortex and hippocampus subregions could be modulated by $\epsilon 4$ carriership [96]. We have thus identified subregions that are consistent with early vascular leakage in the aging brain, such as CA1 and PFC, as manifesting ADRD-related structural deviation in the same HC-DN signature associated with air pollution in our phenome-wide analysis pooled across *APOE* genotypes. In doing so, we extend the alleged role of vascular integrity in protecting the brain from environmental stressors that might precipitate ADRD onset in *APOE* $\epsilon 4$ carriers.

In a similar vein, in vitro analyses have suggested that exposure to air pollution can trigger microglial activation, which in turn can cause oxidative stress [97,98]. Pollution-triggered oxidative stress could be particularly detrimental to males as they are thought to display lower expression of antioxidant enzymes responsible for scavenging reactive oxygen species [99,100]. As a result, male mice show up to 4-fold higher rates of oxidative toxicity in astrocytes, neurons, and mitochondria compared to female mice [99,101]. Our results suggest that the association between HC-DN co-variation and air pollution is male specific, building on experimental findings primarily from rodent species. Parts of the DN are thought to be among the earliest sites of A β accumulation [29] and consume some of the highest oxygen levels in the entire brain [102]. As such, the DN sticks out as a hotspot for both oxidative stress and ADRD pathology. A previous study has indeed found widespread glucose hypometabolism in the DN of ADRD patients that was associated with increased levels of CSF lactate, a marker of mitochondrial damage, in the OFC and mPFC as compared to cognitively healthy controls [103]. Recent evidence suggests that A β_{1-42} acts on reactive oxygen species to induce glucose hypometabolism [104]. One could argue that the combined effect of air pollution and amyloid pathology could be particularly detrimental in exacerbating ADRD risk among males. In line with an effect on escalating ADRD risk, specifically in males, we have linked ADRD-related structural deviation in the OFC and mPFC with a profile of associations with environmental phenotypes for mode 3. As was the case for mode 1, these associations were more prominent in males than females. In addition to emphasising a male-specific vulnerability to neurotoxicity, our phenome-wide analysis pointed towards a female-specific resilience to pollution-mediated impairment and subsequent neuronal death. For example, our phenome-wide profile for

mode 1, derived for females, did not show statistically relevant associations with air pollutants but displayed a significant correlation with IGF-1. Estrogen and IGF-1 are thought to exert synergetic, non-additive effects on neurite outgrowth and survival, presumably by acting on a single neuroendocrine pathway [105]. IGF-1 is secreted by neurons and glia and possibly acts as a neurotrophic factor regulating neuroendocrine function in the central nervous system [105]. Subcutaneous injection of IGF-1 has previously been associated with increased neurogenesis in the adult rat brain [106,107]. In mode 1, in addition to a female-specific association with IGF-1, we have shown HC hits in the granule cell layer of the DG and in CA4, which are 2 subfields in which neurogenesis has been observed in rodents [106,108,109] and primates [110]. Together with its associated divergences in HC-DN co-variation expressions, the phenome-wide profile for mode 1 shed light on a female-specific resilience towards pollution-induced impairment and subsequent neuronal death. While scarcely reported in human subjects, these sex-specific divergences in vulnerability to neurotoxicity—observed here for both modes 1 and 3—are hence in accordance with experimental findings from animal models.

Building on the knowledge that ADRD and verbal-numerical reasoning overlap in underlying genetic architecture [111], we showed significant brain-behaviour associations between ADRD risk and baseline cognitive performance on the fluid intelligence battery for top modes 1, 2, and 3. While previous investigations of fluid intelligence and ADRD in the UKB were often limited to genetic evidence [111–114], we highlighted distinct HC-DN signatures related to verbal-numerical reasoning at the population level. In doing so, we found prominent ADRD-related structural divergences in the left CA1, CA2/3, presubiculum, and fimbria, which are among the first and notorious regions to be affected by ADRD pathology [13,25–27]. Some authors have claimed that white matter disruption may trigger grey matter degradation in the HC and higher-order neocortex by activating a maladaptive neuroinflammatory response [115]. Changes in fornix microstructure have indeed been reported in individuals at risk of ADRD before the onset of clinical symptoms [26] and subsequently identified as an accurate predictor of progression from mild cognitive impairment to ADRD [27]. Consistent with the early involvement of the fornix in ADRD-associated cognitive deficits, we showed structural divergence in the fornix's fimbria and 56 DN regions for mode 3, which were accompanied by a profile of associations with questions from the fluid intelligence battery.

Recent brain-imaging evidence has extended the concept of a hippocampally mediated cognitive map to interpersonal relationships by highlighting the involvement of the DN, and hence the fornix, in schematic representations of the self and others. Notably, fMRI results from Tavares and colleagues suggest that the HC tracks how we represent others in a social hierarchy while the PCC/PCu, key hubs of the DN, tracks the social distance between ourselves and others [116]. Consistent with a reliance on the HC-DN pathway for human-defining aspects of spatiotemporal processing, we found a brain-behaviour association with navigating family relationships, a subtest of the fluid intelligence battery, that was significant in males for mode 3. We have thus provided a plausible link between verbal-numerical reasoning and ADRD risk that was accompanied by alterations in HC and DN subregion co-variation regimes involved in episodic processing.

By exploring risk-anatomy links across the different *APOE* gene variants, we have tied social engagement measures to subject-specific expressions of HC-DN co-variation signatures. Notably, we found that the contribution of social behaviours to risk-anatomy links went beyond genetic risk and was prominent across the different *APOE* genotypes. In older age, a decrease in social activity possibly related to unemployment and/or retirement could increase feelings of loneliness and consequently escalate the risk of cognitive decline and ADRD [117]. Social disengagement has indeed been associated with the incidence of cognitive decline among older adults [118–120]. In contrast, engaging in social activities has been linked with

up to a 40% decrease in ADRD risk [68,118,121]. While social support has been associated with a dampened stress response [122], loneliness is thought to affect not only neuroendocrine but also immune functions [123,124]. Volunteering and having student status, 2 social engagements that have repeatedly been flagged in our analyses, could possibly downplay the pathological stress response observed in lonely older adults. Our study has thus uncovered risk-anatomy links that are consistent with the involvement of social factors as potentially preventing or exacerbating ADRD risk.

Our clustering analyses also uncovered that neurotic behaviours show unique ties to HC-DN co-variation expressions in $\epsilon 2$ carriers. Neuroticism, which is intimately related to loneliness [125], could predispose individuals to ADRD by weakening strong social support ties and increasing chronic stress through dysregulation of the hypothalamus–pituitary–adrenal (HPA) axis [123,126]. In fact, the HC subiculum, presubiculum, and parasubiculum are believed to have direct connections to the hypothalamus via the fornix [127]. These connections could possibly provide a pathway through which the subjective appraisal of one's relationships, which can in turn result in loneliness or neuroticism if social needs are unfulfilled, is conveyed to the HPA to affect the stress response. Prospective cohort studies have indeed linked neuroticism to higher risks of developing cognitive impairments [128] and dementia [129–131]. Yet, no effects of $\epsilon 4$ dosage on cognitive decline have been observed in neurotic individuals in these previous reports [128,130]. The absence of a relationship between *APOE* and neurotic traits reported by previous studies might arise from restricting analyses to $\epsilon 4$ carriers [128,130]. Indeed, the combined analysis of $\epsilon 4$ and the K variant of *BCHE*, another genetic risk factor associated with ADRD, revealed an intriguing association between the combined risk alleles, increased basal levels of serum glucocorticoids, cognitive performance, and lower self-esteem in older adults [132]. The ramifications of neuroticism for ADRD risk, which might be underscored by *APOE* $\epsilon 2$, have been overlooked in all these studies. Recent evidence has also shown that having a positive outlook on aging, such as a sense of purpose, amplified the protective effect of *APOE* $\epsilon 2$ against cognitive decline [133]. The protective effect of *APOE* $\epsilon 2$ on cognition was enhanced for individuals with positive beliefs about aging and reduced for those with negative beliefs to the point where $\epsilon 2$ carriers no longer showed a significant cognitive advantage [133]. Our results add elements to this literature by suggesting that having a negative outcome on life, which is characteristic of a neurotic personality type, is especially detrimental to $\epsilon 2$ carriers as reflected by unique patterns of brain–behaviours associations with specific HC and DN subregions. The opposing health effects of neuroticism and social activity are possibly reflected in the brain, as social and neurotic phenotypes were divided into 2 main groups when clustered based on their correlation with HC-DN co-variation regimes for $\epsilon 2$ homozygotes. Our study thus reinforces the detrimental effect of neuroticism on ADRD risk and characterised its unique interplay with HC-DN co-variation expressions in $\epsilon 2$ homozygotes.

In sum, the typically protective benefits conferred by *APOE* $\epsilon 2$ regarding ADRD risk have mainly been discussed in epidemiological cohorts that were not designed to incorporate inter-individual differences in high-resolution brain structure assessments. In contrast, neuroimaging investigations of healthy participants before the onset of ADRD-associated clinical symptoms have focused on characterising the functional correlates of $\epsilon 4$ carriership. Our present study has reconciled these 2 research streams by contrasting profiles of brain–behaviours associations characteristic of *APOE* $\epsilon 2$ and $\epsilon 4$ in a large epidemiological cohort of approximately 40,000 participants. In doing so, we were uniquely positioned to illuminate sex-specific associations with modifiable risk factors that were unique to $\epsilon 2$ and $\epsilon 4$ homozygotes. Key risk factors relevant to $\epsilon 2$ carriers included neuroticism, social disengagement, and physical inactivity. In contrast, environmental phenotypes that repeatedly emerged in our results as being linked to ADRD risk could be especially detrimental to $\epsilon 4$ carriers. These distinct risk factors could guide potential clinical interventions and governmental policies.

Methods

Population data source

The UK Biobank (UKB) is a large-scale data-collection initiative that offers in-depth information on approximately 500,000 participants recruited from across Great Britain (<https://www.ukbiobank.ac.uk/>). This rich epidemiological cohort comprises a wide variety of resources, including physical and cognitive assessments, as well as demographic and health records. In addition to the availability of genetic data for most participants through a genotyping array (and more recently through whole-exome sequencing), the UKB provides multimodal imaging scans that are routinely augmented and will extend to approximately 100,000 participants by the end of 2022. The present study was based on the data release from February/March 2020. To ensure reproducibility, we adopted the uniform preprocessing pipelines designed and carried out by FMRIB, Oxford University, United Kingdom [134]. Building on a uniform quality control workflow enables a better comparison to other and future UKB research. At the time of data release, expert-curated image-derived phenotypes of grey matter morphology (T1-weighted magnetic resonance imaging) were available for 38,292 participants. Grey matter phenotypes from these participants were used to compute dominant regimes of structural correspondence between the hippocampus (HC) and default network (DN) and identify anatomical subregions that systematically differentiate individuals with and without a family history of ADRD. As for all subsequent analysis steps, we focused on the 37,291 participants with both *APOE* single-nucleotide polymorphisms (SNPs) genotyping (rs429358 and rs7412) and brain-imaging measures (47% men and 53% women). When recruited, these participants were aged 40 to 70 years (mean age 54.8, standard deviation [SD] 7.5 years). The demographic information for the UKB participants included in the present study, grouped per *APOE* genotypes, can be found in Table 1. The present analyses were conducted under UK Biobank application number 25163. UK Biobank participants gave written, informed consent for the study, which

Table 1. UK Biobank demographic information.

	ε3/3	ε3/4	ε2/3	ε2/4	ε4/4	ε2/2
N (%)	22,129 (59.3)	8,613 (23.1)	4,625 (12.4)	885 (2.4)	822 (2.2)	217 (0.6)
Age, Mean ± SD	54.9 ± 7.5	54.5 ± 7.4	55.0 ± 7.5	55.0 ± 7.5	54.3 ± 7.3	54.6 ± 7.5
Sex, n (%)						
Females	11,579 (52.3)	4,634 (53.8)	2,464 (53.3)	489 (55.3)	447 (54.4)	117 (53.9)
Males	10,550 (47.7)	3,979 (46.2)	2,161 (46.7)	396 (44.7)	375 (45.6)	100 (46.1)
Family history of ADRD, n (%)						
Maternal	3,516 (15.9)	1,972 (22.9)	695 (15.0)	204 (23.1)	227 (27.6)	27 (12.4)
Paternal	1,871 (8.5)	1,078 (12.5)	382 (8.3)	100 (11.3)	136 (16.5)	18 (8.3)
Both	328 (1.5)	235 (2.7)	77 (1.7)	20 (2.3)	39 (4.7)	2 (0.9)
Household income, n (%)						
Less than 18,000 £	2,786 (12.6)	1,077 (12.5)	570 (12.3)	110 (12.4)	103 (12.5)	24 (11.1)
18,000 to 30,999 £	4,980 (22.5)	1,851 (21.5)	1,067 (23.1)	206 (23.3)	168 (20.4)	43 (19.8)
31,000 to 51,999 £	6,602 (29.8)	2,639 (30.6)	1,379 (29.8)	262 (29.6)	245 (29.8)	72 (33.2)
52,000 to 100,000 £	6,086 (27.5)	2,413 (28.0)	1,314 (28.4)	238 (26.9)	240 (29.2)	63 (29.0)
Greater than 100,000 £	1,675 (7.5)	633 (7.3)	278 (6.4)	69 (7.7)	66 (8.0)	15 (6.9)
Age completed full-time education, Mean ± SD	17.0 ± 2.4	17.0 ± 2.4	17.0 ± 2.4	16.9 ± 2.4	16.8 ± 2.5	16.9 ± 2.0
Fluid intelligence score, Mean ± SD	6.2 ± 2.2	6.2 ± 2.1	6.2 ± 2.2	6.3 ± 2.3	6.2 ± 2.2	6.1 ± 2.2

Distribution of the demographic information from the UK Biobank participants included in the present study grouped per *APOE* genotypes. ADRD, Alzheimer's disease and related dementia; SD, standard deviation.

<https://doi.org/10.1371/journal.pbio.3001863.t001>

was approved by the Research Ethics Committee under application 11/NW/0382. Further information on the consent procedure can be found elsewhere (<http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=200>).

Target phenotype for ADRD risk

We used the self-reported family history of ADRD as a simple but accurately measurable estimate of ADRD risk. ADRD is the terminology adopted and recommended by the National Institute on Aging, one of the US Federal Government's National Institutes of Health, to characterise the umbrella of symptoms, diagnoses, and risk factors characteristic of Alzheimer's disease (<https://www.nia.nih.gov/health/alzheimers>). The acronym "ADRD" acknowledges the known heterogeneity of clinical diagnoses of dementia. Additionally, one can only ultimately confirm Alzheimer's disease at the highest degree of certainty based on post-mortem tissue analysis. In the UKB resource, maternal (UKB data field 20110) and paternal (UKB data field 20107) history of ADRD was ascertained as part of the initial assessment (2006 to 2010). As per UKB protocol, participants were asked, "Has/did your mother ever suffer from Alzheimer's disease or dementia?" and "Has/did your father ever suffer from Alzheimer's disease or dementia?". This exact phenotype has been successfully treated as a reliable estimate of maternal/paternal history of late-onset Alzheimer's disease by previously published genome-wide association studies conducted in the UKB cohort that successfully recovered well-known genetic risk loci for this diagnosis [135–137]. There were a total of 9,776 (25.5%) participants with self-reported parental history of ADRD within the brain-imaging cohort of 38,292 participants. Of those with family risk, 6,820 UKB participants reported an occurrence of ADRD on their mother's side and 3,675 participants on their father's side. A minority of participants reported both maternal and paternal history of ADRD (719 individuals).

Most genome-wide association studies have adopted a case-control framework that focused on the difference in allele frequency between patients with ADRD and healthy controls [138,139]. While useful in identifying risk loci associated with clinical diagnosis, this approach might not be best suited to derive a reliable estimate of ADRD liability in the general population. When dealing with late-onset diseases, such as ADRD, using "proxy cases," i.e., the relatives of affected individuals, could allow for a more complete characterisation of disease risk among individuals before the onset of clinical symptoms [136]. It was a key advantage that working with proxy cases also allowed us to boost the sample size and, thus, the statistical power of our quantitative analyses to identify more suitable effects. In particular, self-report of family history of ADRD in the UKB, precisely the same phenotype at the core of the present investigation, was found to replicate established risk loci from case-control investigations as well as identify novel loci [136,137].

Brain-imaging and preprocessing procedures

Magnetic resonance imaging (MRI) scanners (3T Siemens Skyra) were matched at several dedicated data collection sites with the same acquisition protocols and standard Siemens 32-channel radiofrequency receiver head coils. Brain-imaging data were defaced, and any sensitive meta-information was removed to protect the anonymity of the study participants. Automated processing and quality control pipelines were deployed [134,140]. Noise was removed utilising 190 sensitivity features to improve the homogeneity of the imaging data. This approach allowed for the reliable identification and exclusion of problematic brain scans, such as due to excessive head motion.

The structural MRI data were acquired as high-resolution T1-weighted images of brain anatomy using a 3D MPRAGE sequence at 1 mm isotropic resolution. Preprocessing included gradient distortion correction (GDC), field of view reduction using the Brain Extraction Tool

[141] and FLIRT [142,143], as well as nonlinear registration to MNI152 standard space at 1 mm resolution using FNIRT [144]. All image transformations were estimated, combined, and applied by a single interpolation step to avoid unnecessary interpolation. Tissue-type segmentation into cerebrospinal fluid, grey matter, and white matter was applied using FAST (FMRIB's Automated Segmentation Tool, [145]) to generate full bias-field-corrected images. In turn, SIENAX [146] was used to derive volumetric measures normalised for head sizes.

Parcellation of the DN was anatomically guided by the Schaefer-Yeo reference atlas [147]. We extracted a total of 400 parcels among the 7 canonical networks, 91 of which were defined as belonging to the DN. Volume extraction for 38 HC subregions was conducted using FreeSurfer automatic sub-segmentation [21], which drew on an ultra-high-resolution (approximately 0.1 mm isotropic) probabilistic atlas. As part of the FreeSurfer 7.0 suite, HC sub-segmentation was refined by carefully considering surrounding anatomical structures.

As a preliminary procedure, these MRI-derived measures were cleaned to remove inter-individual variation in brain region volumes that could be explained by nuisance variables. Building on previous UK Biobank research [148,149], we regressed out the following variables of no interest from each brain-derived volume measure: body mass index, head size, head motion during task-related brain scans, head motion during task-unrelated brain scans, head position and receiver coil in the scanner (x , y , and z), position of scanner table, as well as the data acquisition site, in addition to age, age², sex, sex*age, and sex*age². Sex was acquired from the National Health Service (NHS) central registry and updated by the participant if incorrect (UKB data field 31). The nuisance-cleaned volumetric measures served as the basis of our primary co-decomposition analysis—seeking to quantify how the 91 DN subregions co-deviate with the 38 HC subregions in the context of ADRD risk.

Population co-variation between hippocampus subregions and default-network subregions

At the heart of our analysis workflow, we derived dominant regimes of structural correspondence that provide insights into how structural variation among the finely segregated HC can track structural variation among the finely segregated DN. We employed canonical correlation analysis (CCA), a doubly multivariate statistical technique, to identify population “signatures” of HC-DN co-variation. CCA was a natural choice of method as it is specially designed to disentangle patterns of joint correlation between 2 high-dimensional variable sets [23,150,151]. The first variable set, X , was constructed from subject-level grey matter volume in DN subregions (number of participants \times 91 DN parcels matrix). The second variable set, Y , was constructed from HC subregion volumes (number of participants \times 38 HC parcels matrix). The 2 variable sets can be formally described as follows:

$$X \in \mathbb{R}^{n \times p}$$

$$Y \in \mathbb{R}^{n \times q},$$

where n denotes the number of observations or UKB participants, p is the number of DN subregions, and q is the number of HC subregions. Subregion volumes from both variable sets were z-scored across participants to zero mean (i.e., centering) and unit variance (i.e., rescaling). CCA then addressed the problem of maximising the linear correlation between low-rank projections from 2 variable sets or data matrices [23]. The 2 sets of linear combinations of the original variables are obtained by optimising the following target function:

$$L_X = XV \quad L_Y = YU$$

$$l_{X,l} = Xv_l \quad l_{Y,l} = Yu_l$$

$$\text{corr}(l_{X,l}, l_{Y,l}) \propto l_{X,l}^T l_{Y,l} = \max,$$

where V and U denote the respective contributions of X and Y , L_X and L_Y denote the respective latent “modes” expression of joint variation (i.e., canonical variates) based on patterns derived from X and patterns derived from Y , $l_{X,l}$ is the l th column of L_X , and $l_{Y,l}$ is the l th column of L_Y .

Our CCA application thus sought to identify linear combinations of X and Y that optimise their low-rank projections in the derived latent embedding. Such an approach resulted in pairs of latent vectors with subject-specific expressions $l_{X,l}$ and $l_{Y,l}$ (i.e., canonical variates) with maximised joint correlation. Corresponding pairs of latent vectors were found by iteratively decomposing the data matrices X and Y into k components, where k denotes the number of modes given the model specification. In other words, CCA searched for the canonical vectors u and v that maximise the symmetric relationship between the data matrices of DN subregion volumes (X) and HC subregion volumes (Y). In doing so, CCA identified the 2 concomitant projections Xv_l and Yu_l that optimised the correspondence between structural variation in the segregated DN and HC.

Put differently, each principled signature of HC-DN co-variation, or mode, represents the cross-correlation between a constellation of within-DN volumetric variation and a constellation of within-HC volumetric variation that co-occurred in conjunction with each other. The set of k modes are mutually uncorrelated by construction (orthogonality) [23]. They are also naturally rank-ordered based on the amount of variance explained between the embedded allocortical and neocortical volume sets [23]. The first and strongest mode, therefore, explained the largest fraction of joint variation between (linear) combinations of HC subregions and (linear) combinations of DN subregions. Each ensuing cross-correlation signature captured a fraction of structural variation that is not explained by one of the $k-1$ other modes. The Pearson’s correlation between a pair of canonical variates (i.e., canonical correlation) is commonly used to quantify the linear correspondence between HC subregions and DN subregions for a given mode. The 2 variable sets were entered into CCA after a confound-removal procedure based on previous UK Biobank research (cf. above).

Group difference analysis

After constructing population signatures of conjoint HC-DN co-variation, we performed a rigorous group difference analysis to single out microstructural divergences in specific anatomical subregions with respect to ADRD family history. For each of the derived modes of HC-DN co-variation, we aimed to isolate anatomical subregions that show statistically defensible deviation in individuals with and without a family history of ADRD. To do so, we carried out a principled test that assessed any statistically relevant differences in the solution vector obtained from the CCA (i.e., canonical vectors, cf. above) of individuals at ADRD risk compared to the control group without ADRD family history (cf. above for target phenotype).

Following previous UK Biobank research [24,152], we robustly characterised the difference between individuals with and without a family history of ADRD by carrying out a bootstrap difference test of the CCA solution at hand [153]. This approach aimed to identify consistent patterns of deviation that differentiate subjects with and without a family history of ADRD. We first proceeded by constructing several alternative datasets that we could have gotten (with the same sample size), which capture the underlying population variation. For each of the 100 bootstrap iterations, these alternative datasets were built by randomly pulling participant

samples with replacements. In each such bootstrap iteration, we estimated 2 CCA models in parallel by fitting 1 separate model to each of the 2 groups. In doing so, we carried out $2 * 100$ separate model estimations of the doubly multivariate correspondence between HC subregions and DN subregions.

To compare the CCA solution in individuals with and without a family history of ADRD, we matched corresponding modes based on sign invariance and mode rank order. Canonical vectors of a given mode that carried opposite signs were aligned by multiplying 1 with -1 . The importance rank of the CCA modes was adjusted by sorting Pearson's correlation coefficients between pairs of corresponding canonical vectors (i.e., canonical correlations) from strongest to weakest. To estimate a quantity of group difference in relation to ADRD risk, we performed the elementwise subtraction of the corresponding canonical vector entries of a given mode k between the 2 groups. Pooling outcomes across the 100 bootstrap iterations, we thus aggregated the difference estimate for each canonical vector entry, thereby quantifying the uncertainty deviation for each particular HC or DN subregion.

By probing the underlying population variation, we were able to quantify the degree of uncertainty within each of our derived modes of HC-DN co-variation. For each identified population signature, we therefore isolated statistically defensible group differences in microanatomically defined HC and DN subregions. ADRD-related structural divergences were determined by whether the two-sided confidence interval included zero or not according to the 10/90% bootstrap-derived distribution of difference estimates [149]. In doing so, we obtained a nonparametric estimate of how ADRD risk is manifested in specific subregions for each of the 25 examined HC-DN signatures.

SNP genotyping: 6 variants of APOE gene

We capitalised on our large sample size to demystify the HC-DN co-variation expressions associated with $\epsilon 2$ allele and $\epsilon 4$ allele homozygotes compared to their heterozygous counterparts for the $\epsilon 2$, $\epsilon 3$, or $\epsilon 4$ alleles. Genotype-level sampling and quality control procedures for the UKB are available online (<https://biobank.ndph.ox.ac.uk/showcase/label.cgi?id=263>). APOE genotypes were determined based on 2 SNPs: rs429358 and rs7412. APOE $\epsilon 4$ was determined as the combination of rs429358(C) and rs7412(C). APOE $\epsilon 2$ was determined as the combination of rs429358(T) and rs7412(T). APOE $\epsilon 3$ was determined based on rs429358(T) and rs7412(C). A total of 37,291 participants had both APOE genotyping and brain-imaging-derived measures. Among those participants, 9,525 (25.5%) reported a family history of ADRD. We observed 6 different APOE gene variants in our population sample: $\epsilon 3/3$ (59.3%), $\epsilon 3/4$ (23.1%), $\epsilon 2/3$ (12.4%), $\epsilon 2/4$ (2.4%), $\epsilon 4/4$ (2.2%), and $\epsilon 2/2$ (0.6%), which correspond to frequencies expected from a population primarily composed of people from European descent [22]. Contrasting the effect of $\epsilon 2$ versus $\epsilon 4$ allele dosage on inter-individual expressions of HC-DN co-variation enabled us to quantify the degree to which distinct APOE allelic combinations are characteristic of ADRD risk (cf. next section). In doing so, we aimed to interrogate gradual dosage effects in brain-APOE associations rather than simply look at $\epsilon 4$ carrier versus non-carrier status.

Phenome-wide analysis of brain-behaviour associations in relation to $\epsilon 2$ versus $\epsilon 4$ dosage

We performed a rich annotation of the HC-DN co-variation signatures by means of their phenome-wide association with UKB traits. We were interested in how $\epsilon 2$ versus $\epsilon 4$ allele dosage is manifested in inter-individual expressions of HC-DN co-variation and how these manifestations, in turn, relate to UKB traits among a variety of predefined risk categories. We benefited

from a rich portfolio of phenotypes encompassing lifestyle, cognitive, mental, and physical health assessments to ascribe profiles of brain-behaviour associations to each of the 25 modes of HC-DN co-variation.

We started with a raw collection of approximately 15,000 phenotypes that we fed into the FMRIB UKB Normalisation, Parsing And Cleaning Kit (FUNPACK version 2.5.0; <https://zenodo.org/record/4762700#.YQrpui2caJ8>). FUNPACK was used to extract phenotype information covering 11 major categories, including cognitive and physiological assessments, physical and mental health records, blood assays, as well as sociodemographic and lifestyle factors. We removed any brain-imaging-derived information. The diet category was additionally excluded from downstream analyses as it contained only 4 phenotypes. FUNPACK was designed to perform automatic refinement on the UKB data, which included removing “do not know” responses and filling the blank left by unanswered sub-questions. For example, the amount of alcohol drunk on a typical drinking day for a participant who indicated not drinking would be scored as zero drinks, even though this sub-question was not actually asked at assessments. FUNPACK’s output consisted of a collection of 3,330 curated phenotypes which were then fed into PHEnome Scan ANalysis Tool (PHESANT [154], <https://github.com/MRCIEU/PHESANT>) for further refinement. In addition to data cleaning and normalisation, PHESANT categorised the data as belonging to 1 of 4 datatypes: categorical ordered, categorical unordered, binary, and numerical. Categorical unordered variables were one-hot encoded, such that each possible response was represented by a binary column (true or false). The final curated inventory comprised 977 phenotypes spanning 11 FUNPACK-defined categories.

We next checked for statistically robust associations between HC-DN signatures and the portfolio of 977 extracted phenotypes with respect to ADRD genetic risk. We used a one-step stacking strategy [155,156] to predict genetic risk as a function of individual expressions of HC-DN co-variation. Data stacking consists of using a “base” model, often linear regression [156], to express an input vector in a lower-dimensional space. The output of the base model, which often consists of a single variable, can then be used as a single predictor in a new “stacking” model. Therefore, data stacking addressed the problem of selecting a single best predictor out of a combination of highly correlating input variables—which in our case were the corresponding HC and DN co-variation patterns. Such an approach allowed us to re-express a whole signature of HC-DN co-variation in terms of the degree it tracked the associated risk conferred by *APOE*. We formed a single continuous number representing how much a given HC-DN signature reflects $\epsilon 2$ versus $\epsilon 4$ dosage for a given individual. Investigations of *APOE* $\epsilon 4$ dosage effects have been prevalent in brain imaging research [114,157,158].

The Alzheimer’s disease research community has widely endorsed encoding $\epsilon 4$ dosage in a stepwise fashion, i.e., based on the number of allele copies carried by a given patient [112,114,157,158]. By adopting such target variable representation, Lyall and colleagues have found a significant interaction between *APOE* genotype dosage and coronary artery disease in estimating verbal-numerical scores from the fluid intelligence battery in the UK Biobank [112]. Lyall and colleagues, however, missed looking at $\epsilon 2$ dosage despite the well-established association between the $\epsilon 2$ allele and neurovascular diseases [62,63]. More recently, *APOE* $\epsilon 4$ dosage, stepwise encoded as 0, 1, or 2, was shown to be significantly associated with right hippocampal volume and white matter intensity in the UK Biobank [114]. The authors, however, did not benefit from investigating HC anatomical segmentations besides the standard head/body/tail subdivision [114]. Again, *APOE* $\epsilon 2$ dosage was not considered in this previous work even though neuroimaging evidence has lent support for a dose-dependent increase in hippocampal volume of 769.3 mm³ per copy of the $\epsilon 2$ allele, on average [158].

Consequently, the present study builds on the widely shared belief that the $\epsilon 2$ and $\epsilon 4$ alleles have largely opposing effects on Alzheimer’s risk and pathophysiology [38,159,160]. We

sought an analogous composite dosage scale that readily captures opposite effects in modelling the HC and DN volume variation dependent on the copy number of $\epsilon 2$ and $\epsilon 4$ alleles. We thus created a bivariate dosage scale by summing up positive “ $\epsilon 2$ ” and negative “ $\epsilon 4$ ” alleles, such that a homozygous individual carrying *APOE* $\epsilon 2/2$ would have a score of +2 and one carrying *APOE* $\epsilon 4/4$ a score of -2. The neutral *APOE* $\epsilon 3$ allele, usually considered as a baseline risk in epidemiological studies [22], was scored as 0. Using a bivariate dosage scale made it possible to investigate the antagonistic effects of $\epsilon 2$ and $\epsilon 4$ in a single model. In doing so, we stayed faithful to our overarching goal of unravelling their adversarial impact on HC-DN co-variation.

Aiming to capture possible sex-specific effects, we regressed the $\epsilon 2$ versus $\epsilon 4$ dosage on inter-individual expressions of a given mode in males and females separately. We thus estimated 2 * 25 different base models, 1 for each HC-DN signature and each sex, that each had 2 parameters: the pair of co-variation expressions (i.e., canonical vectors, cf. above) associated with the HC and DN patterns. We used these 25 regression models to explain the subject-level $\epsilon 2$ versus $\epsilon 4$ dosage as a function of HC-DN co-variation expressions. For each subject and mode combination, we asked what would the expected $\epsilon 2$ versus $\epsilon 4$ dosage be given this subject's specific expression of HC-DN co-variation? For each subject, we hence used the regression model to explain a range from -2 to +2 for each mode, which represented the $\epsilon 2$ versus $\epsilon 4$ dosage associated with their individual expression of HC-DN co-variation. For each mode, we selected the 5th and top 95th percentiles to identify the top 5% and lower 5% of individuals who were more versus less likely to develop ADRD based on the derived $\epsilon 2$ versus $\epsilon 4$ dosage risk. We focused on the extreme of the dosage distribution to target the brain-*APOE* associations especially linked to $\epsilon 2$ and $\epsilon 4$. The analogous approach is widely adopted in genome-wide analyses to remove associations not directly linked to the target genotype [161,162].

For each sex separately and for a given mode, the designated participants were put to a test of association with the 977 curated UKB phenotypes, with appropriate correction for multiple comparisons. The Pearson's correlation between a phenotype and genetic risk predicted based on a specific HC-DN signature revealed both the association strength and accompanying statistical significance of the given mode-trait association. For each HC-DN signature, 2 widely used procedures were carried out to adjust for the multitude of associations being assessed. First, we adjusted for the number of tested phenotypes by using Bonferroni's correction for multiple comparisons ($0.05/977 = 5.11 \times 10^{-5}$). Second, we used the FDR, another popular adjustment, although less stringent than Bonferroni's correction. The FDR [163] was set as 5% [140,164,165] and computed for each HC-DN signature in accordance with standard practice [166]. For the sake of visualisation, we used Miami plots to compare the profiles of brain-behaviour associations derived from males and females. For visualisation purposes, phenotypes in Miami plots were coloured and grouped according to the category membership defined by FUNPACK.

Clustering of risk factors based on their correlation with HC-DN co-variation expressions

We next systematically explored nonlinear associations between established ADRD risk phenotypes and HC-DN co-variation expressions across the different *APOE* gene variants. Our goal was to probe for clusters of risk factors that are interrelated with the derived patterns of HC and DN co-variation. To this end, we used a hierarchical clustering approach that allowed us to assess the relative importance of ensuing clusters in each of the different *APOE* genotypes to explore gradual *APOE* dosage effects on risk-anatomy links.

We adopted a targeted approach by focusing on a set of 63 risk factors (collection of phenotypes used previously [34]), including classical cardiovascular and demographic traits, as well

as social richness indicators recently linked to ADRD in the UKB cohort. The first step of the clustering analysis consisted of multiplying the z-scored canonical variates by each of the 6 one-hot encoded *APOE* genotypes (i.e., $\epsilon 2/2$, $\epsilon 2/3$, $\epsilon 3/3$, $\epsilon 2/4$, $\epsilon 3/4$, and $\epsilon 4/4$) such that participants without a given genotype were zeroed out. The 6 ensuing matrices (number of participants \times 50 canonical variates) represented the individual expressions of HC-DN co-variation signatures for participants with a given *APOE* genotype, whereas other participants were scored as 0s. We then computed Spearman's correlation between these 6 genotype-specific matrices and the z-scored risk factor matrix (37,291 participants \times 63 risk factors) to investigate risk-anatomy links. Spearman's correlation is a nonparametric measure of statistical dependence between the rankings of 2 variables that can be used to capture monotonic nonlinear phenomena. The Spearman's correlation coefficients reduce to the Pearson's correlation between the rank values of 2 variables and hence range from -1 (inversely proportional association) to $+1$ (proportional association). We obtained a new cross-association matrix $X \in \mathbb{R}^{63 \times 50}$ which represented the Spearman's correlation between the 63 risk factors and the 50 canonical variates for each of the 6 *APOE* genotypes. The obtained Spearman's correlation coefficients thus carried the nonlinear association strength of a given risk-anatomy link for a particular *APOE* genotype.

For each of the 6 *APOE* genotypes, we performed an agglomerative hierarchical clustering analysis on X to regroup risk factors based on their 50 associations with HC-DN co-variation pattern expressions. We used Ward's minimum variance method [167] to compute the linkage matrix between the Spearman's correlation coefficients of each risk-anatomy link in Euclidian space. Ward's minimum variance criterion consists in minimising the total within-cluster variance defined as the error sum of squares:

$$d_{ij} = d(\{X_i\}, \{X_j\}) = \|X_i - X_j\|^2,$$

where d_{ij} represents the squared Euclidean distance between 2 points (or cluster of points) i and j . At each step, the pair of coefficients or preceding candidate clusters that give the minimum increase in within-cluster variance is selected for merging. The procedure was performed recursively until all coefficients were merged into a single cluster. For each of the 6 *APOE* genotypes, we could thus create a dendrogram that represented the distance in Euclidian space between the clusters retained after 3 levels of branching. The level of branching refers to the number of divisions from the final merge. The dendrograms allowed us to visualise the clustering results for each of the 6 *APOE* genotypes at the same level of branching and identify meaningful clusters of risk-anatomy links that are shared or unique. To provide a more direct assessment of the degree of dissimilarity, we have compared the spread between nodes in the analogous dendrograms for each *APOE* genotype. We used Pearson's correlation to examine the Euclidean distance between the 2 descendent links across corresponding hierarchical merging steps in the 6 genotype-specific cluster models.

Regression of ADRD risk on HC-DN signatures and *APOE* gene variants

We next tested whether specific *APOE* genotypes showed interaction effects with signatures of HC-DN co-variation in explaining inter-individual differences in ADRD risk. As our goal was to highlight previously overlooked sex effects, we conducted our interaction analyses in males and females separately. In doing so, we aimed to characterise brain-*APOE* interactions in relation to their sex-specific impact on ADRD risk.

A first series of analyses consisted in regressing each of the previously investigated ADRD risk factors on *APOE* genotypes, co-variation patterns from the HC and DN sides (i.e., canonical variates), and the interaction between *APOE* genotypes and co-variation patterns,

controlling for age. Aiming to capture possible sex-specific effects, we conducted separate analyses on males and females. We, therefore, looked at 61 ADRD risk factors, while age was used as a covariate and sex was the grouping factor for stratification. Each of the 25 modes of HC-DN co-variation was represented by 2 regression models: 1 for its HC pattern and 1 for its DN pattern. We thus formed 50 univariate regression models, in males and females, for each of the 61 risk factors. In each of these models, a given risk factor was regressed on 1 HC or DN canonical variate, the 6 *APOE* genotypes ($\epsilon 2/2$, $\epsilon 2/3$, $\epsilon 3/3$, $\epsilon 2/4$, $\epsilon 3/4$, and $\epsilon 4/4$), and 6 interaction terms capturing the nonlinear association between each of the 6 *APOE* genotypes and the given HC or DN pattern, controlling for age. Each regression model thus aimed at explaining variance in one of the 61 risk factors for a given sex based on these 14 parameters.

As a conjoint analysis across the regression models, we performed a rigorous permutation analysis to assess the robustness of each of the 14 regression coefficients. In as many as 61,000 iterations (i.e., 61 risk factors * 1,000 iterations), we randomly shuffled the outcome variable (i.e., a given risk) across participants. We recomputed the otherwise identical regression model based on the data with randomised outcomes. We recorded the regression coefficients from each of the 61,000 iterations and used them to build empirical null distributions on which we performed two-tail statistical tests. We considered statistically relevant coefficients that differ from their respective null distributions in at least 95% of the iterations, which ensured that we were at least 5% certain that the effect was robustly different from zero. This threshold remains arbitrary as our post hoc interaction analyses were merely descriptive and designed to provide a coarse portrait of gene-brain interactions rather than claiming statistical significance. For that reason, we have made publicly available masked permutations plots at the 0.0001, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 0.8, 0.95 percentiles for the coefficient estimates of each regression model for males (https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/fig_5/permutation_analysis/males/masked_plots) and females (https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/fig_5/permutation_analysis/females/masked_plots).

A second series of analyses consisted in regressing the family history of ADRD on a set of explanatory input variables including (i) *APOE* genotypes; (ii) co-variation patterns from the HC and DN sides (i.e., canonical variates); and (iii) the interaction between *APOE* genotypes and co-variation patterns, controlling for age. For each sex, we built separate logistic models for each of the 25 HC and 25 DN canonical variates, for a total of 50 models per sex. In each model, the family history of ADRD (encoded as 0 for no and 1 for yes) was regressed on 1 HC or DN canonical variate, the 6 *APOE* genotypes ($\epsilon 2/2$, $\epsilon 2/3$, $\epsilon 3/3$, $\epsilon 2/4$, $\epsilon 3/4$, and $\epsilon 4/4$), and 6 interaction terms capturing the nonlinear association between each of the 6 *APOE* genotypes and the given HC or DN pattern, controlling for age. We thus obtained a total of 100 logistic models that sought to explain variance in the family history of ADRD as a function of these 14 parameters. We performed the analogous permutation analysis (described above) to assess the robustness of each of the 14 regression coefficients derived from these 100 logistic models. We have made publicly available the permutation distributions at the 0.0001, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 0.8, 0.95 percentiles for the coefficients of each regression model (https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/fig_6/permutation_analyses).

Latent factor analysis of brain-behaviour associations

To finally distill latent factor embeddings of brain-behaviour associations from our HC-DN population signatures, we used the classical linear dimensionality reduction method PCA [168]. PCA was a natural choice of method to uncover linearly independent groupings of risk factors with similar relatedness to HC-DN co-variation patterns. Latent factors uncovered by

the PCA are naturally ordered from most to least important which allows us to select candidate principles of brain–behaviour association that account for the most inter-individual variance.

We started by computing the Pearson's correlation between the z-scored canonical variate matrix (number of participants \times 50 canonical variates) and the z-scored risk factor matrix (number of participants \times 63 risk factors). We obtained a new matrix $M \in \mathbb{R}^{63 \times 50}$, which represented the Pearson's correlation coefficients between the 63 risk factors and the 50 canonical variates. We next decomposed M into latent factor groupings by using singular value decomposition (SVD). Every correlation coefficient in M had already been z-scored to abide by zero mean and unit variance prior to computing the SVD, as per common practice [169]. More formally, solving the SVD problem took the following form:

$$M = U S V^T,$$

Where U is a 63×63 orthonormal matrix, S is a 63×50 diagonal matrix carrying the singular values, and V is a 50×50 orthonormal matrix carrying singular vectors.

We retained the top 3 singular vectors and expressed our correlation matrix in terms of the dot product $US \in \mathbb{R}^{63 \times 3}$ to be able to represent the latent-factor projections of M onto the new 3D latent space. In doing so, we obtained the distinct expression levels of the 63 risk factors for each of the top 3 brain–behaviour association axes (i.e., principal component expressions). These 3 axes are by construction orthonormal and rank-ordered, representing an uncorrelated partition of the overall variance in brain–behaviour association. The leading axis captured the largest fraction of variance and was, therefore, the most explanatory, as reflected by its associated singular value.

We then conducted an acid test of the robustness of the PCA solution by performing a rigorous split-half reliability assessment across 1,000 bootstrap iterations. At each iteration, we drew 37,291 participants with replacements to simulate random participant samples that we could have pulled from the same population. We then derived 2 random subsets of equal size ($N = 18,645$) from the original sample and re-computed the Pearson's correlation matrix M for each random subset separately. SVD was then performed on both matrices in parallel according to the procedure described above. We retained the same number of top 3 singular vectors and expressed each correlation matrix in terms of its projection onto its corresponding latent space. In doing so, we were able to compare the expression levels of each risk factor along the 3 main axes of brain–behaviour associations derived from each random subset. If the PCA solution is robust, similar groups of risk factors should be emphasised along corresponding dimensions, which, in turn, should explain similar fractions of the total variance. We also provided a more formal assessment of statistical agreement between both random subsets by computing the Pearson's correlation between the weights of the 3 first principal components for random subsets 1 and 2 across the 1,000 iterations. Higher Pearson's correlations are indicative of a substantial degree of agreement between both subsets, which in turn attests to the robustness of the original PCA solution.

Based on the desire to audit our cohort analysis for sex-specific associations, we computed the Pearson's correlation matrix M in males and females separately and repeated the PCA procedure described above for each group. Once more, we retained the top 3 singular vectors and expressed the correlation matrices in terms of their projection onto their corresponding latent embedding. We compared the expression levels of the risk factors along corresponding latent dimensions to highlight sex-specific brain–behaviour associations. In the absence of major sex differences, similar groups of risk factors should be emphasised along analogous dimensions, which should correspondingly explain similar fractions of the total variance.

We performed a similar bootstrap analysis of the sex-specific PCA solutions to formally assess the robustness of our findings. Across 1,000 bootstrap iterations, we drew 17,561 males and 19,730 females with replacements to simulate random participant samples that we could have gotten from the original population. At each iteration, re-computed the Pearson's correlation matrix for each random subset separately and repeated the analogous SVD decomposition. As for the split-half reliability assessment, we Pearson's correlated the weights of the 3 first principal components for male- and female-derived solutions in each of the 1,000 iterations. Lower Pearson's correlations would suggest a higher degree of sex-specificity in the PCA solutions.

External validation

Using the openly available PREVENT-AD (PResymptomatic EVAluation of Experimental or Novel Treatments for Alzheimer's disease (AD); [170]) cohort, we have performed a rigorous test of the external validation for our HC-DN co-variations signatures derived from the UKB cohort. The PREVENT-AD cohort is composed of older individuals with a known family history of Alzheimer's disease that were cognitively unimpaired at the time of enrollment from 2011 to 2017 (mean age 63, SD 5 years) [170]. Participants of the PREVENT-AD initiative have undergone extensive annual health and cognitive assessments for up to 5 years. This resource creates a unique opportunity to monitor longitudinal trajectories of brain-imaging assessments, cerebral fluid biochemistry, neurosensory capacities, and medical charts in pre-symptomatic individuals at Alzheimer's risk. Our independent PREVENT-AD sample consisted of 386 participants (27% men, 73% women) with the following APOE genotype distribution: $\epsilon 3/3$ (51.2%), $\epsilon 3/4$ (33.1%), $\epsilon 2/3$ (10.5%), $\epsilon 2/4$ (3.0%), $\epsilon 4/4$ (2.1%). Further information on the PREVENT-AD cohort and access to the open data inventory can be found online (<https://prevent-alzheimer.net>).

The PREVENT-AD resources provide structural brain-imaging scanning (T1-weighted images of brain anatomy) for up to 4 years of follow-up for 362 participants, totaling 980 participant assessment visits. For the brain-imaging data from each participant visit, we first performed a full FreeSurfer reconstruction followed by subcortical volumetric sub-segmentation of the 38 hippocampal subfields, analogous to the UKB brain-imaging preprocessing pipeline. We next parsed the structural brain scans according to the Schaefer-Yeo parcellation (400 parcels, 7 networks) to obtain the analogous 91 parcels defined as belonging to the DN (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal/Parcellations/project_to_individual). Age, age², sex, sex*age, and sex*age² were regressed out from each brain-derived grey matter volume measure as part of the deconfounding procedure. The final brain-imaging sample consisted of 344 participants with a total of 916 individual visits (64 visits were excluded based on errors in the preprocessing pipeline). Of the remaining visits, 117 came from participants without APOE SNP genotyping and were hence excluded.

In so doing, we extracted the same collection of brain-image-derived phenotypes of grey matter morphology as in the UKB. We were thus in a position to compute the expression of the 25 UKB-derived modes of HC-DN co-variation based on grey matter measurements for the 91 DN and 38 HC subregions in PREVENT-AD participants. For each visit, we obtained 25 pairs of subject-specific expressions of each of the 25 brain signatures of HC-DN structural co-variation (i.e., canonical variates), which served as a basis for our external validation analyses in unseen subjects.

Across MRI visits, we tested whether 25 different signatures of HC-DN co-variations are associated with different subsets among the rich palette of PREVENT-AD phenotypes designed to track ADRD progression in pre-symptomatic individuals.

To do so, using Pearson's correlation, we computed the association strength between the individual expressions of the 25 modes of HC-DN co-variation and 157 PREVENT-AD phenotypes that spanned CSF and blood samples, comprehensive cognitive and functional assessments, as well as demographic and health records. To assess the robustness of the correlation coefficients, we randomly permuted the PREVENT-AD phenotypes across participants in 1,000 iterations and recomputed the Pearson's correlation coefficients. Recording the results from these 1,000 iterations, we built an empirical null distribution for each correlation coefficient. We reported only the coefficients that were robustly different from their respective empirical null distributions in at least 95% of the 1,000 permutation iterations.

Supporting information

S1 Fig. ADRD-related divergences in HC and DN subregions for mode 2 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 2 for the HC (leftmost panel) and DN (central panel). We identified 10 HC hits, most of them located in the left hemisphere. The strongest HC divergences were observed for the presubiculum, hippocampal fissure, and CA2/3. We found corresponding DN hits in posterior midline structure (posterior cingulate cortex and retrosplenial cortex), the dorsomedial prefrontal cortex, and the posterior and temporal cortices. In males and females separately, we regressed *APOE* dosage on HC and DN co-variation patterns from mode 2. We then used these sex-specific models to predict *APOE* dosage based on inter-individual expressions of mode 2. The right panel displays the Miami plot for the correlations between *APOE* scores in the context of mode 2 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We found significant associations with the fluid intelligence battery that were unique to males. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer's disease and related dementia; CA, cornu amonis; DG, granule cell layer of the dentate gyrus; dmPFC, dorsomedial prefrontal cortex; DN, default network; FDR, false discovery rate correction; HC, hippocampus; IPL, inferior parietal lobule; MTS, middle temporal sulcus; PrS, presubiculum; PCC, posterior cingulate cortex; RSC, retrosplenial cortex; STS, superior temporal sulcus. (TIFF)

S2 Fig. ADRD-related divergences in HC and DN subregions for mode 6 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 6 for the HC (leftmost panel) and DN (central panel). We identified 1 HC hit to the hippocampus-amygdala transition area with no concurrent DN divergences. In males and females separately, we regressed *APOE* dosage on HC and DN co-variation patterns from mode 6. We then used these sex-specific models to predict *APOE* dosage based on inter-individual expressions of mode 6. The right panel displays the Miami plot for the correlations between *APOE* scores in the context of mode 6 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We found significant associations with physical phenotypes and blood assays that were unique to females. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer's disease and related dementia; DN, default network; FDR, false discovery rate correction; HATA, hippocampus-amygdala transition area; HC, hippocampus; IMT, intima-medial thickness. (TIFF)

S3 Fig. ADRD-related divergences in HC and DN subregions for mode 10 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode

10 for the HC (leftmost panel) and DN (central panel). We identified 1 HC hit to the hippocampus-amygdala transition area with no concurrent DN divergences. In males and females separately, we regressed *APOE* dosage on HC and DN co-variation patterns from mode 10. We then used these sex-specific models to predict *APOE* dosage based on inter-individual expressions of mode 10. The right panel displays the Miami plot for the correlations between *APOE* scores in the context of mode 10 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We found 1 significant association with sitting height unique to males. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)). ADRD, Alzheimer's disease and related dementia; DN, default network; FDR, false discovery rate correction; HATA, hippocampus-amygdala transition area; HC, hippocampus.

(TIFF)

S4 Fig. ADRD-related divergences in HC and DN subregions for mode 4 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 4 for the HC (leftmost panel) and DN (central panel). We identified 4 DN hits to the dorsomedial prefrontal cortex with no concurrent HC divergences. In males and females separately, we regressed *APOE* dosage on HC and DN co-variation patterns from mode 4. We then used these sex-specific models to predict *APOE* dosage based on inter-individual expressions of mode 4. The right panel displays the Miami plot for the correlations between *APOE* scores in the context of mode 4 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We found 1 significant association with receiving an attendance, disability or mobility allowance that was unique to females. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)). ADRD, Alzheimer's disease and related dementia; dmPFC, dorsomedial prefrontal cortex; DN, default network; FDR, false discovery rate correction; HC, hippocampus.

(TIFF)

S5 Fig. ADRD-related divergences in HC and DN subregions for mode 7 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 7 for the HC (leftmost panel) and DN (central panel). We identified 9 DN hits to the frontal lobe with no concurrent HC divergences. In males and females separately, we regressed *APOE* dosage on HC and DN co-variation patterns from mode 7. We then used these sex-specific models to predict *APOE* dosage based on inter-individual expressions of mode 7. The right panel displays the Miami plot for the correlations between *APOE* scores in the context of mode 7 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We found 1 significant association with diastolic blood pressure that was unique to females. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)). ADRD, Alzheimer's disease and related dementia; dmPFC, dorsomedial prefrontal cortex; DN, default network; FDR, false discovery rate correction; HC, hippocampus; OFC, orbitofrontal cortex.

(TIFF)

S6 Fig. ADRD-related divergences in HC and DN subregions for mode 11 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 11 for the HC (leftmost panel) and DN (central panel). We identified 1 DN hit to the posterior cingulate cortex with no concurrent HC divergences. In males and females separately, we regressed *APOE* dosage on HC and DN co-variation patterns from mode 11. We then used these sex-specific models to predict *APOE* dosage based on inter-individual expressions of

mode 11. The right panel displays the Miami plot for the correlations between *APOE* scores in the context of mode 11 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We found 1 significant association with the standing height that was unique to females. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer's disease and related dementia; DN, default network; FDR, false discovery rate correction; HC, hippocampus; PCC, posterior cingulate cortex. (TIFF)

S7 Fig. ADRD-related divergences in HC and DN subregions for mode 13 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 13 for the HC (leftmost panel) and DN (central panel). We identified 1 DN hit to the superior temporal sulcus with no concurrent HC divergences. In males and females separately, we regressed *APOE* dosage on HC and DN co-variation patterns from mode 13. We then used these sex-specific models to predict *APOE* dosage based on inter-individual expressions of mode 13. The right panel displays the Miami plot for the correlations between *APOE* scores in the context of mode 13 and the portfolio of UKB phenotypes for males (upper half) and females (lower half). We found significant associations with physical measurements related to height as well as feelings of guilt that were unique to females. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer's disease and related dementia; DN, default network; FDR, false discovery rate correction; HC, hippocampus; STS, superior temporal sulcus. (TIFF)

S8 Fig. Difference in associations between males and females for the phenome-wide profiling of mode 1. Absolute difference in *p*-values for the 33 brain-phenotype associations that passed the Bonferroni correction for multiple comparisons in either males or females in the original phenome-wide profiling of mode 1. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). (PNG)

S9 Fig. Difference in associations between males and females for the phenome-wide profiling of mode 3. Absolute difference in *p*-values for the 20 brain-phenotype associations that passed the Bonferroni correction for multiple comparisons in either males or females in the original phenome-wide profiling of mode 3. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). (PNG)

S10 Fig. Difference in associations between males and females for the phenome-wide profiling of mode 8. Absolute difference in *p*-values for the 18 brain-phenotype associations that passed the Bonferroni correction for multiple comparisons in either males or females in the original phenome-wide profiling of mode 8. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/Miami_Plots (DOI: 10.5281/zenodo.7126809). (PNG)

S11 Fig. Similarity between the 6 genotype-specific clustering models. We computed Pearson's correlation of the distance between the 2 descendent links of corresponding hierarchical

merging steps among the cluster analyses for the 6 APOE genotypes (i.e., $\epsilon 2/2$, $\epsilon 2/3$, $\epsilon 3/3$, $\epsilon 2/4$, $\epsilon 3/4$, and $\epsilon 4/4$). These derived distances made it possible to formally compare the cluster nodes of analogous dendrograms for each genotype-specific cluster model. We show that $\epsilon 2$ carriers are most similar to each other, as reflected by an agglomeration of strong Pearson's correlation coefficients in the top left corner of the heat map. The most dissimilar cluster models were $\epsilon 2/4$ and $\epsilon 3/4$, followed by $\epsilon 2/4$ and $\epsilon 3/3$, and lastly by $\epsilon 3/3$ and $\epsilon 4/4$. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/clustering_analysis (DOI: 10.5281/zenodo.7126809).

(TIFF)

S12 Fig. $\epsilon 3$ carriership shows risk-anatomy links with socioeconomic determinants, while $\epsilon 2$ carriership is associated with neuroticism. We multiplied the population-wide HC and DN co-variation patterns by APOE genotypes $\epsilon 3/3$ ($N = 22,129$) and $\epsilon 2/4$ ($N = 885$) such that participants who do not carry a given genotype were zeroed out. We then computed the Spearman's correlations between these 2 new vectors and the 63 preselected Alzheimer's disease risk factors to test for risk-anatomy links. We performed an agglomerative clustering analysis on these Spearman's correlations, which consists in repeatedly merging Spearman's correlations with similar variance together until all observations are merged into a single cluster. Here are shown the dendrograms, which indicate the distance between each cluster identified when retaining 3 levels of branching for APOE $\epsilon 3/3$ (leftmost panel) and $\epsilon 2/4$ (rightmost panel). We found the early branching of socioeconomic determinants $\epsilon 3/3$ (time spent watching television, education score, past and current tobacco smoking frequency, alcohol consumption on a typical drinking day, and alcohol intake frequency) in the clustering model for $\epsilon 3/3$. For $\epsilon 2/4$, we found that neuroticism-related behaviours (e.g., being worried/anxious, mood swings, and miserableness) were singled out from the other risk-anatomy links at the first branching, as was observed for other $\epsilon 2$ carriers. We thus confirm the association between $\epsilon 3$ carriership and socioeconomic determinants and between $\epsilon 2$ carriership and neurotic personality traits. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/tree/master/clustering_analysis (DOI: 10.5281/zenodo.7126809).

(TIFF)

S13 Fig. Latent factors of brain-behaviour associations emphasise satisfaction with social relationships, socioeconomic status, and neuroticism-related traits. We conducted an exploratory PCA to disentangle latent factor of brain-behaviour association in our UK Biobank sample. We first computed the Pearson's correlations between the 25 pairs of co-variation patterns from the HC and DN sides and the 63 preselected ADRD risk factors. We then ran singular value decomposition on the risk by canonical variates matrix ($X_{63 \times 50}$) and retained the 3 first PCs that explained approximately 13.8%, approximately 9.6%, and approximately 8.2% of the total variance in the data, respectively. The upper plot displays the projections of the Pearson's correlations onto each of the 3 main axes of brain-behaviour associations. The lower plot displays the eigenvectors for the top 10 HC and DN co-variation patterns. The first axis of brain-behaviour associations emphasises phenotypes from the social cluster previously identified on the clustering analysis of risk-anatomy links (Fig 4), e.g., attending religious group, attending adult education classes, and number of people in household. The second axis rather accented health-related phenotypes and lifestyle factors. Lastly, the third axis of brain-behaviour associations separated neuroticism-related items (being worried/anxious, being easily hurt, and worrying too long after embarrassment) from the rest of the risk factors. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/blob/master/PCA (DOI: 10.5281/zenodo.7126809). ADRD, Alzheimer's disease and related dementia; DN, default network; HC, hippocampus; PCA, principal

component analysis.
(TIFF)

S14 Fig. Reliability assessment of the principal component solution. We assessed the robustness of the derived brain–behaviour association axes by performing a split-half reliability assessment of our principal component solution across 1,000 bootstrap iterations. At each iteration, we drew 37,291 participants with replacements to simulate random participant samples that we could have pulled from the same population. We then derived 2 random subsets of equal size ($N = 18,645$) from the original sample. For each subset, we re-computed the Pearson’s correlation between all possible combinations of the 50 canonical variates and 63 target indicators. We then estimated 2 PCA models in parallel, one for each random half subset, on the z-scored correlation coefficients matrices. We show the average projections of the Pearson’s correlation coefficients on the 3 first axes of brain–behaviour associations. We found that the projections on component 1 were robust and consistent across subsets. The projections on the first axis of brain–behaviour associations accurately depicted those of the original PCA solution, with the same set of social phenotypes (e.g., attending religious group, attending adult education classes, and the number of people in the household) and socioeconomic determinants (e.g., age completed high school education, average household income, and the number of vehicles in the household) emphasised. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/blob/master/PCA (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)).

(TIFF)

S15 Fig. Statistical agreement between the PCA solutions for random subsets 1 and 2. We computed the Pearson’s correlation between the weights of the 3 first principal components for random subsets 1 and 2 across 1,000 bootstrap iterations. The weights of the first 2 components were robust, as reflected by a substantial degree of agreement between both subsets on components 1 (mean Pearson’s rho: 0.59, 90% CI: [0.38, 0.74]) and 2 (mean Pearson’s rho: 0.51, 90% CI: [0.15, 0.77]). In contrast, we showed volatility in the weights associated with component 3, as reflected by a wider and right-skewed distribution (mean Pearson’s rho: 0.25, 90% CI: [0.02, 0.56]). Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/blob/master/PCA (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)).

(TIFF)

S16 Fig. Neuroticism-related items expressed distinctive brain–behaviour associations in males and females. We repeated the PCA in males (left; $N = 17,561$) and females right; $N = 19,730$) separately. In each sex, we first computed the Pearson’s correlations between the 25 pairs of co-variation patterns from the HC and DN sides and the 63 preselected ADRD risk factors. We then ran singular value decomposition on the risk by canonical variates matrix ($X_{63 \times 50}$) and retained the 3 first PCs. The PCs obtained from males had explained variance of approximately 14.6%, approximately 11.9%, and approximately 9.6%, respectively. The PCs obtained from females had explained variance of approximately 14.6%, approximately 11.9%, and approximately 7.4%, respectively. The upper plots display the projections of the Pearson’s correlations onto each of the 3 axes of brain–behaviour associations for the 2 sexes. The lower plots display the eigenvectors for the top 10 HC and DN co-variation patterns. The projections of the Pearson’s correlations onto the 2 first axes of brain–behaviour association were roughly the same in males and females. In contrast, neuroticism-related items were only emphasised on the third axis of brain–behaviour association in males. We thus supplemented our population analysis by showing that the relationship between neuroticism and patterns of HC-DN co-variation was mainly male specific. Data underlying this figure can be found at <https://>

github.com/dblabs-mcgill-mila/HCDMNCOV_AD/blob/master/PCA (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)). ADRD, Alzheimer's disease and related dementia; DN, default network; HC, hippocampus; PCA, principal component analysis.

(TIFF)

S17 Fig. Statistical agreement between the PCA solutions for males and females. We computed the Pearson's correlation between the weights of the first 3 principal components for the sex-specific PCA solutions across 1,000 bootstrap iterations. We observed a low agreement between the male- and female-derived PCA solutions on all 3 components, as reflected by the widespread of the distributions and small average values. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/blob/master/PCA (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)).

(TIFF)

S18 Fig. HC-DN signatures tracked different aspects of ADRD risk in independent PREVENT-AD participants. We externally validated our UKB-derived population signatures of HC-DN co-variation by investigating their mapping to ADRD-related risk factors in an unseen, independent participant sample. We tracked subject-specific expressions of the 25 modes of HC-DN co-variation in PREVENT-AD participants to a collection of 157 widely established indicators of ADRD progression. We computed the Pearson's correlation between the HC and DN pattern expressions and the PREVENT-AD phenotypes for each mode. Only the Pearson's correlation coefficients that were statistically different from their respective null distributions 95% of the time are present. We replicated several phenotypic associations highlighted in the UKB, such as with mode 1 and depression, mode 2 and verbal-numerical reasoning, and mode 6 and vascular integrity. We also showed that our modes of HC-DN co-variation track meaningful aspects of ADRD progression up to the 25th and last signature, for which we found associations with tau CSF levels on the HC side and cardiovascular factors (e.g., systolic blood pressure, pulse, and APOE ϵ 4/4 genotype) on the DN side. We thus showed that HC-DN signatures robustly link to different aspects of ADRD risk in a completely independent cohort from the one in which the co-variation patterns have originally been derived. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/blob/master/external_validation (DOI: [10.5281/zenodo.7126809](https://doi.org/10.5281/zenodo.7126809)). ADRD, Alzheimer's disease and related dementia; DN, default network; HC, hippocampus.

(TIFF)

Author Contributions

Conceptualization: Chloé Savignac, Danilo Bzdok.

Data curation: Chloé Savignac, Karin Saltoun, Kimia Shafighi, Vaibhav Sharma, Danilo Bzdok.

Formal analysis: Chloé Savignac, Karin Saltoun, Vaibhav Sharma.

Funding acquisition: Danilo Bzdok.

Investigation: Chloé Savignac, Chris Zajner, Danilo Bzdok.

Methodology: Chloé Savignac, Karin Saltoun, Danilo Bzdok.

Project administration: Danilo Bzdok.

Resources: Danilo Bzdok.

Software: Chloé Savignac, Karin Saltoun, Vaibhav Sharma.

Supervision: Danilo Bzdok.

Validation: Chloé Savignac.

Visualization: Chloé Savignac, Karin Saltoun, Chris Zajner.

Writing – original draft: Chloé Savignac, Sylvia Villeneuve, AmanPreet Badhwar, Karin Saltoun, Kimia Shafighi, Chris Zajner, Sarah A. Gagliano Taliun, Sali Farhan, Judes Poirier, Danilo Bzdok.

Writing – review & editing: Chloé Savignac, Sylvia Villeneuve, AmanPreet Badhwar, Karin Saltoun, Kimia Shafighi, Chris Zajner, Vaibhav Sharma, Sarah A. Gagliano Taliun, Sali Farhan, Judes Poirier, Danilo Bzdok.

References

1. Patterson C. World alzheimer report. 2018: 2018.
2. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *Lancet Public Health*. 2022. [https://doi.org/10.1016/S2468-2667\(21\)00249-8](https://doi.org/10.1016/S2468-2667(21)00249-8) PMID: 34998485
3. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet*. 2020; 396(10248):413–46. [https://doi.org/10.1016/S0140-6736\(20\)30367-6](https://doi.org/10.1016/S0140-6736(20)30367-6) PMID: 32738937
4. Poirier J, Bertrand P, Poirier J, Kogan S, Gauthier S, Poirier J, et al. Apolipoprotein E polymorphism and Alzheimer's disease. *Lancet*. 1993; 342(8873):697–99. [https://doi.org/10.1016/0140-6736\(93\)91705-q](https://doi.org/10.1016/0140-6736(93)91705-q) PMID: 8103819
5. Herholz K, Salmon E, Perani D, Baron J-C, Holthoff V, Frölich L, et al. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage*. 2002; 17(1):302–16. <https://doi.org/10.1006/nimg.2002.1208> PMID: 12482085
6. Badhwar A, Tam A, Dansereau C, Orban P, Hoffstaedter F, Bellec P. Resting-state network dysfunction in Alzheimer's disease: a systematic review and meta-analysis. *Alzheimers Dement*. 2017; 8:73–85. <https://doi.org/10.1016/j.dadm.2017.03.007> PMID: 28560308
7. Song H, Long H, Zuo X, Yu C, Liu B, Wang Z, et al. APOE effects on default mode network in Chinese cognitive normal elderly: relationship with clinical cognitive performance. *PLoS ONE*. 2015; 10(7): e0133179. <https://doi.org/10.1371/journal.pone.0133179> PMID: 26177270
8. Goveas JS, Xie C, Chen G, Li W, Ward BD, Franczak MB, et al. Functional network endophenotypes unravel the effects of apolipoprotein E epsilon 4 in middle-aged adults. *PLoS ONE*. 2013; 8(2):e55902. <https://doi.org/10.1371/journal.pone.0055902> PMID: 23424640
9. Westlye ET, Lundervold A, Rootwelt H, Lundervold AJ, Westlye LT. Increased Hippocampal Default Mode Synchronization during Rest in Middle-Aged and Elderly APOE ε4 Carriers: Relationships with Memory Performance. *J Neurosci*. 2011; 31(21):7775–83.
10. Filippini N, MacIntosh BJ, Hough MG, Goodwin GM, Frisoni GB, Smith SM, et al. Distinct patterns of brain activity in young carriers of the APOE-ε4 allele. *Proc Natl Acad Sci U S A*. 2009; 106(17):7209–14.
11. Mentink LJ, Guimarães JPOFT, Faber M, Sprooten E, Olde Rikkert MGM, Haak KV, et al. Functional co-activation of the default mode network in APOE ε4-carriers: A replication study. *Neuroimage*. 2021; 240:118304.
12. Insel PS, Hansson O, Mattsson-Carlsson N. Association Between Apolipoprotein E ε2 vs ε4, Age, and β-Amyloid in Adults Without Cognitive Impairment. *JAMA Neurol*. 2021; 78(2):229–35.
13. Braak E, Braak H. Alzheimer's disease: transiently developing dendritic changes in pyramidal cells of sector CA1 of the Ammon's horn. *Acta Neuropathol*. 1997; 93(4):323–25. <https://doi.org/10.1007/s004010050622> PMID: 9113196
14. Shafer AT, Beason-Held L, An Y, Williams OA, Huo Y, Landman BA, et al. Default mode network connectivity and cognition in the aging brain: the effects of age, sex, and APOE genotype. *Neurobiol Aging*. 2021; 104:10–23. <https://doi.org/10.1016/j.neurobiolaging.2021.03.013> PMID: 33957555
15. Salami A, Pudas S, Nyberg L. Elevated hippocampal resting-state connectivity underlies deficient neurocognitive function in aging. *Proc Natl Acad Sci U S A*. 2014; 111(49):17654. <https://doi.org/10.1073/pnas.1410233111> PMID: 25422457

16. Kobayashi Y, Amaral DG. Macaque monkey retrosplenial cortex: III. Cortical efferents. *J Comp Neurol*. 2007; 502(5):810–33. <https://doi.org/10.1002/cne.21346> PMID: 17436282
17. Aggleton JP, Wright NF, Rosene DL, Saunders RC. Complementary Patterns of Direct Amygdala and Hippocampal Projections to the Macaque Prefrontal Cortex. *Cereb Cortex*. 2015; 25(11):4351–73. <https://doi.org/10.1093/cercor/bhv019> PMID: 25715284
18. Barbas H, Blatt GJ. Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. *Hippocampus*. 1995; 5(6):511–33. <https://doi.org/10.1002/hipo.450050604> PMID: 8646279
19. Fleisher A, Sun S, Taylor C, Ward C, Gamst A, Petersen R, et al. Volumetric MRI vs clinical predictors of Alzheimer disease in mild cognitive impairment. *Neurology*. 2008; 70(3):191–99. <https://doi.org/10.1212/01.wnl.0000287091.57376.65> PMID: 18195264
20. Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol*. 1991; 82(4):239–59. <https://doi.org/10.1007/BF00308809> PMID: 1759558
21. Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, et al. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. *Neuroimage*. 2015; 115:117–37. <https://doi.org/10.1016/j.neuroimage.2015.04.042> PMID: 25936807
22. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of Age, Sex, and Ethnicity on the Association Between Apolipoprotein E Genotype and Alzheimer Disease: A Meta-analysis. *JAMA*. 1997; 278(16):1349–56.
23. Wang H-T, Smallwood J, Mourao-Miranda J, Xia CH, Satterthwaite TD, Bassett DS, et al. Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists. *Neuroimage*. 2020; 216:116745. <https://doi.org/10.1016/j.neuroimage.2020.116745> PMID: 32278095
24. Zajner C, Spreng RN, Bzdok D. Loneliness is linked to specific subregional alterations in hippocampus-default network covariation. *J Neurophysiol*. 2021; 126(6):2138–57. <https://doi.org/10.1152/jn.00339.2021> PMID: 34817294
25. Padurariu M, Ciobica A, Mavroudis I, Fotiou D, Baloyannis S. Hippocampal neuronal loss in the CA1 and CA3 areas of Alzheimer's disease patients. *Psychiatr Danub*. 2012; 24(2):152–8. PMID: 22706413
26. Gold BT, Powell DK, Andersen AH, Smith CD. Alterations in multiple measures of white matter integrity in normal women at high risk for Alzheimer's disease. *Neuroimage*. 2010; 52(4):1487–94. <https://doi.org/10.1016/j.neuroimage.2010.05.036> PMID: 20493952
27. Mielke MM, Okonkwo OC, Oishi K, Mori S, Tighe S, Miller MI, et al. Fornix integrity and hippocampal volume predict memory decline and progression to Alzheimer's disease. *Alzheimers Dement*. 2012; 8(2):105–13. <https://doi.org/10.1016/j.jalz.2011.05.2416> PMID: 22404852
28. Badhwar A, Lerch JP, Hamel E, Sled JG. Impaired structural correlates of memory in Alzheimer's disease mice. *Neuroimage Clin*. 2013; 3:290–300. <https://doi.org/10.1016/j.nicl.2013.08.017> PMID: 24273714
29. Palmqvist S, Schöll M, Strandberg O, Mattsson N, Stomrud E, Zetterberg H, et al. Earliest accumulation of β -amyloid occurs within the default-mode network and concurrently affects brain connectivity. *Nat Commun*. 2017; 8(1):1–13.
30. Damoiseaux JS, Seeley WW, Zhou J, Shirer WR, Coppola G, Karydas A, et al. Gender modulates the APOE ϵ 4 effect in healthy older adults: convergent evidence from functional brain connectivity and spinal fluid tau levels. *J Neurosci*. 2012; 32(24):8254–62.
31. Reiman EM, Chen K, Alexander GE, Caselli RJ, Bandy D, Osborne D, et al. Correlations between apolipoprotein E ϵ 4 gene dose and brain-imaging measurements of regional hypometabolism. *Proc Natl Acad Sci U S A*. 2005; 102(23):8299–302.
32. Reiman EM, Caselli RJ, Yun LS, Chen K, Bandy D, Minoshima S, et al. Preclinical Evidence of Alzheimer's Disease in Persons Homozygous for the ϵ 4 Allele for Apolipoprotein E. *N Engl J Med*. 1996; 334(12):752–58.
33. Corder E, Saunders AM, Risch N, Strittmatter W, Schmechel D, Gaskell P, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat Genet*. 1994; 7(2):180–84. <https://doi.org/10.1038/ng0694-180> PMID: 7920638
34. Shafiqi K, Villeneuve S, Rosa-Neto P, Badhwar A, Poirier J, Sharma V, et al. Social isolation is linked to classical risk factors of Alzheimer's disease-related dementias. *bioRxiv*. 2021.
35. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc*. 2005; 53(4):695–99. <https://doi.org/10.1111/j.1532-5415.2005.53221.x> PMID: 15817019

36. Li M, Li Y, Liu Y, Huang H, Leng X, Chen Y, et al. Altered Hippocampal Subfields Volumes Is Associated With Memory Function in Type 2 Diabetes Mellitus. *Front Neurol*. 2021; 12:756500. <https://doi.org/10.3389/fneur.2021.756500> PMID: 34899576
37. Lim YY, Mormino EC. Initiative AsDN. APOE genotype and early β -amyloid accumulation in older adults without dementia. *Neurology*. 2017; 89(10):1028–34.
38. Nagy ZS, Esiri MM, Jobst KA, Johnston C, Litchfield S, Sim E, et al. Influence of the apolipoprotein E genotype on amyloid deposition and neurofibrillary tangle formation in Alzheimer's disease. *Neuroscience*. 1995; 69(3):757–61. [https://doi.org/10.1016/0306-4522\(95\)00331-c](https://doi.org/10.1016/0306-4522(95)00331-c) PMID: 8596645
39. Reiman EM, Arboleda-Velasquez JF, Quiroz YT, Huentelman MJ, Beach TG, Caselli RJ, et al. Exceptionally low likelihood of Alzheimer's dementia in APOE2 homozygotes from a 5,000-person neuropathological study. *Nat Commun*. 2020; 11(1):667. <https://doi.org/10.1038/s41467-019-14279-8> PMID: 32015339
40. Jansen WJ, Ossenkuppe R, Knol DL, Tijms BM, Scheltens P, Verhey FRJ, et al. Prevalence of Cerebral Amyloid Pathology in Persons Without Dementia: A Meta-analysis. *JAMA*. 2015; 313(19):1924–38. <https://doi.org/10.1001/jama.2015.4668> PMID: 25988462
41. Kim YJ, Seo SW, Park SB, Yang JJ, San Lee J, Lee J, et al. Protective effects of APOE e2 against disease progression in subcortical vascular mild cognitive impairment patients: a three-year longitudinal study. *Sci Rep*. 2017; 7(1):1–8.
42. Lippa CF, Smith TW, Saunders AM, Hulette C, Pulaski-Salo D, Roses AD. Apolipoprotein E-epsilon 2 and Alzheimer's disease: Genotype influences pathologic phenotype. *Neurology*. 1997; 48(2):515–19. <https://doi.org/10.1212/wnl.48.2.515> PMID: 9040748
43. Shu H, Shi Y, Chen G, Wang Z, Liu D, Yue C, et al. Opposite Neural Trajectories of Apolipoprotein E ϵ 4 and ϵ 2 Alleles with Aging Associated with Different Risks of Alzheimer's Disease. *Cereb Cortex*. 2016; 26(4):1421–29.
44. Shinohara M, Kanekiyo T, Yang L, Linthicum D, Shinohara M, Fu Y, et al. APOE2 eases cognitive decline during aging: clinical and preclinical evaluations. *Ann Neurol*. 2016; 79(5):758–74. <https://doi.org/10.1002/ana.24628> PMID: 26933942
45. Riddell DR, Zhou H, Atchison K, Warwick HK, Atkinson PJ, Jefferson J, et al. Impact of Apolipoprotein E (ApoE) Polymorphism on Brain ApoE Levels. *J Neurosci*. 2008; 28(45):11445–53. <https://doi.org/10.1523/JNEUROSCI.1972-08.2008> PMID: 18987181
46. Ulrich JD, Burchett JM, Restivo JL, Schuler DR, Verghese PB, Mahan TE, et al. In vivo measurement of apolipoprotein E from the brain interstitial fluid using microdialysis. *Mol Neurodegener*. 2013; 8(1):13. <https://doi.org/10.1186/1750-1326-8-13> PMID: 23601557
47. Beffert U, Cohn JS, Petit-Turcotte C, Tremblay M, Aumont N, Ramassamy C, et al. Apolipoprotein E and β -amyloid levels in the hippocampus and frontal cortex of Alzheimer's disease subjects are disease-related and apolipoprotein E genotype dependent. *Brain Res*. 1999; 843(1):87–94.
48. Krasemann S, Madore C, Cialic R, Baufeld C, Calcagno N, El Fatimy R, et al. The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases. *Immunity*. 2017; 47(3):566–81.e9. <https://doi.org/10.1016/j.immuni.2017.08.008> PMID: 28930663
49. Huynh T-PV, Wang C, Tran AC, Tabor GT, Mahan TE, Francis CM, et al. Lack of hepatic apoE does not influence early A β deposition: observations from a new APOE knock-in model. *Mol Neurodegener*. 2019; 14(1):1–23.
50. Deane R, Sagare A, Hamm K, Parisi M, Lane S, Finn MB, et al. apoE isoform-specific disruption of amyloid β peptide clearance from mouse brain. *J Clin Invest*. 2008; 118(12):4002–13.
51. Berlau DJ, Corrada MM, Head E, Kawas CH. APOE ϵ 2 is associated with intact cognition but increased Alzheimer pathology in the oldest old. *Neurology*. 2009; 72(9):829–34.
52. Galea LA, Wainwright SR, Roes M, Duarte-Guterman P, Chow C, Hamson D. Sex, hormones and neurogenesis in the hippocampus: hormonal modulation of neurogenesis and potential functional implications. *J Neuroendocrinol*. 2013; 25(11):1039–61. <https://doi.org/10.1111/jne.12070> PMID: 23822747
53. Galea L, McEwen B. Sex and seasonal changes in the rate of cell proliferation in the dentate gyrus of adult wild meadow voles. *Neuroscience*. 1999; 89(3):955–964.
54. Tanapat P, Hastings NB, Reeves AJ, Gould E. Estrogen stimulates a transient increase in the number of new neurons in the dentate gyrus of the adult female rat. *J Neurosci*. 1999; 19(14):5792–801. <https://doi.org/10.1523/JNEUROSCI.19-14-05792.1999> PMID: 10407020
55. Stone DJ, Rozovsky I, Morgan TE, Anderson CP, Hajian H, Finch CE. Astrocytes and Microglia Respond to Estrogen with Increased apoE mRNA in Vivo and in Vitro. *Exp Neurol*. 1997; 143(2):313–18. <https://doi.org/10.1006/exnr.1996.6360> PMID: 9056393

56. Lambert J, Coyle N, Lendon C. The allelic modulation of apolipoprotein E expression by oestrogen: potential relevance for Alzheimer's disease. *J Med Genet.* 2004; 41(2):104–12. <https://doi.org/10.1136/jmg.2003.005033> PMID: 14757857
57. Mahley RW, Huang Y, Rall SC. Pathogenesis of type III hyperlipoproteinemia (dysbetalipoproteinemia): questions, quandaries, and paradoxes. *J Lipid Res.* 1999; 40(11):1933–49.
58. Metzler-Baddeley C, Mole JP, Leonaviciute E, Sims R, Kidd EJ, Ertefai B, et al. Sex-specific effects of central adiposity and inflammatory markers on limbic microstructure. *Neuroimage.* 2019; 189:793–803. <https://doi.org/10.1016/j.neuroimage.2019.02.007> PMID: 30735826
59. Sochocka M, Donskow-Lysoniewska K, Diniz BS, Kurpas D, Brzozowska E, Leszek J. The Gut Microbiome Alterations and Inflammation-Driven Pathogenesis of Alzheimer's Disease—a Critical Review. *Mol Neurobiol.* 2019; 56(3):1841–51. <https://doi.org/10.1007/s12035-018-1188-4> PMID: 29936690
60. Kernbach JM, Yeo BT, Smallwood J, Margulies DS, De Schotten MT, Walter H, et al. Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants. *Proc Natl Acad Sci U S A.* 2018; 115(48):12295–300. <https://doi.org/10.1073/pnas.1804876115> PMID: 30420501
61. Yamatani H, Takahashi K, Yoshida T, Takata K, Kurachi H. Association of estrogen with glucocorticoid levels in visceral fat in postmenopausal women. *Menopause.* 2013; 20(4). <https://doi.org/10.1097/gme.0b013e318271a640> PMID: 23149864
62. Nelson PT, Pious NM, Jicha GA, Wilcock DM, Fardo DW, Estus S, et al. APOE-ε2 and APOE-ε4 Correlate With Increased Amyloid Accumulation in Cerebral Vasculature. *J Neuropathol Exp Neurol.* 2013; 72(7):708–15.
63. Yu L, Boyle PA, Nag S, Leurgans S, Buchman AS, Wilson RS, et al. APOE and cerebral amyloid angiopathy in community-dwelling older persons. *Neurobiol Aging.* 2015; 36(11):2946–53. <https://doi.org/10.1016/j.neurobiolaging.2015.08.008> PMID: 26341746
64. McKay GJ, Patterson CC, Chakravarthy U, Dasari S, Klaver CC, Vingerling JR, et al. Evidence of association of APOE with age-related macular degeneration—a pooled analysis of 15 studies. *Hum Mutat.* 2011; 32(12):1407–16. <https://doi.org/10.1002/humu.21577> PMID: 21882290
65. Zhao N, Liu C-C, Van Ingelgom AJ, Linares C, Kurti A, Knight JA, et al. APOE ε2 is associated with increased tau pathology in primary tauopathy. *Nat Commun.* 2018; 9(1):1–11.
66. Ghebremedhin E, Schultz C, Botez G, Rüb U, Sassin I, Braak E, et al. Argyrophilic grain disease is associated with apolipoprotein E ε2 allele. *Acta Neuropathol.* 1998; 96(3):222–24.
67. Mahley RW. Apolipoprotein E: from cardiovascular disease to neurodegenerative disorders. *J Mol Med.* 2016; 94(7):739–46. <https://doi.org/10.1007/s00109-016-1427-y> PMID: 27277824
68. Wang H-X, Karp A, Winblad B, Fratiglioni L. Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the Kungsholmen project. *Am J Epidemiol.* 2002; 155(12):1081–87. <https://doi.org/10.1093/aje/kwi155.12.1081> PMID: 12048221
69. Podewils LJ, Guallar E, Kuller LH, Fried LP, Lopez OL, Carlson M, et al. Physical activity, APOE genotype, and dementia risk: findings from the Cardiovascular Health Cognition Study. *Am J Epidemiol.* 2005; 161(7):639–51. <https://doi.org/10.1093/aje/kwi092> PMID: 15781953
70. Rovio S, Kåreholt I, Helkala E-L, Viitanen M, Winblad B, Tuomilehto J, et al. Leisure-time physical activity at midlife and the risk of dementia and Alzheimer's disease. *Lancet Neurol.* 2005; 4(11):705–11. [https://doi.org/10.1016/S1474-4422\(05\)70198-8](https://doi.org/10.1016/S1474-4422(05)70198-8) PMID: 16239176
71. Grande G, Vanacore N, Maggiore L, Cucumo V, Ghirelli R, Galimberti D, et al. Physical activity reduces the risk of dementia in mild cognitive impairment subjects: a cohort study. *J Alzheimers Dis.* 2014; 39(4):833–39. <https://doi.org/10.3233/JAD-131808> PMID: 24296815
72. Tan ZS, Spartano NL, Beiser AS, DeCarli C, Auerbach SH, Vasani RS, et al. Physical activity, brain volume, and dementia risk: the Framingham study. *J Gerontol A Biol Sci Med Sci.* 2017; 72(6):789–95. <https://doi.org/10.1093/gerona/glw130> PMID: 27422439
73. Schächter F, Faure-Delanef L, Guénot F, Rouger H, Froguel P, Lesueur-Ginot L, et al. Genetic associations with human longevity at the APOE and ACE loci. *Nat Genet.* 1994; 6(1):29–32. <https://doi.org/10.1038/ng0194-29> PMID: 8136829
74. Ranft U, Schikowski T, Sugiri D, Krutmann J, Krämer U. Long-term exposure to traffic-related particulate matter impairs cognitive function in the elderly. *Environ Res.* 2009; 109(8):1004–11. <https://doi.org/10.1016/j.envres.2009.08.003> PMID: 19733348
75. Wellenius GA, Boyle LD, Coull BA, Milberg WP, Gryparis A, Schwartz J, et al. Residential Proximity to Nearest Major Roadway and Cognitive Function in Community-Dwelling Seniors: Results from the MOBILIZE Boston Study. *J Am Geriatr Soc.* 2012; 60(11):2075–80. <https://doi.org/10.1111/j.1532-5415.2012.04195.x> PMID: 23126566

76. Weuve J, Puett RC, Schwartz J, Yanosky JD, Laden F, Grodstein F. Exposure to Particulate Air Pollution and Cognitive Decline in Older Women. *Arch Intern Med*. 2012; 172(3):219–27. <https://doi.org/10.1001/archinternmed.2011.683> PMID: 22332151
77. Ailshire JA, Crimmins EM. Fine Particulate Matter Air Pollution and Cognitive Function Among Older US Adults. *Am J Epidemiol*. 2014; 180(4):359–66. <https://doi.org/10.1093/aje/kwu155> PMID: 24966214
78. Ailshire JA, Clarke P. Fine Particulate Matter Air Pollution and Cognitive Function Among U.S. Older Adults. *J Gerontol B Psychol Sci Soc Sci*. 2014; 70(2):322–28. <https://doi.org/10.1093/geronb/gbu064> PMID: 24906394
79. Tonne C, Elbaz A, Beevers S, Singh-Manoux A. Traffic-related air pollution in relation to cognitive function in older adults. *Epidemiology*. 2014; 25(5):674–81. <https://doi.org/10.1097/EDE.000000000000144> PMID: 25036434
80. Gatto NM, Henderson VW, Hodis HN, St. John JA, Lurmann F, Chen J-C, et al. Components of air pollution and cognitive function in middle-aged and older adults in Los Angeles. *Neurotoxicology*. 2014; 40:1–7. <https://doi.org/10.1016/j.neuro.2013.09.004> PMID: 24148924
81. Tzivian L, Dlugaj M, Winkler A, Weinmayr G, Hennig F, Fuks KB, et al. Long-Term Air Pollution and Traffic Noise Exposures and Mild Cognitive Impairment in Older Adults: A Cross-Sectional Analysis of the Heinz Nixdorf Recall Study. *Environ Health Perspect*. 2016; 124(9):1361–68. <https://doi.org/10.1289/ehp.1509824> PMID: 26863687
82. Cacciottolo M, Wang X, Driscoll I, Woodward N, Saffari A, Reyes J, et al. Particulate air pollutants, APOE alleles and their contributions to cognitive impairment in older women and to amyloidogenesis in experimental models. *Transl Psychiatry*. 2017; 7(1):e1022. <https://doi.org/10.1038/tp.2016.280> PMID: 28140404
83. Jung C-R, Lin Y-T, Hwang B-F. Ozone, Particulate Matter, and Newly Diagnosed Alzheimer's Disease: A Population-Based Cohort Study in Taiwan. *J Alzheimers Dis*. 2015; 44:573–84. <https://doi.org/10.3233/JAD-140855> PMID: 25310992
84. Wu Y-C, Lin Y-C, Yu H-L, Chen J-H, Chen T-F, Sun Y, et al. Association between air pollutants and dementia risk in the elderly. *Alzheimers Dement*. 2015; 1(2):220–28. <https://doi.org/10.1016/j.dadm.2014.11.015> PMID: 27239507
85. Schikowski T, Vossoughi M, Vierkötter A, Schulte T, Teichert T, Sugiri D, et al. Association of air pollution with cognitive functions and its modification by APOE gene variants in elderly women. *Environ Res*. 2015; 142:10–16. <https://doi.org/10.1016/j.envres.2015.06.009> PMID: 26092807
86. Oudin A, Forsberg B, Adolfsson AN, Lind N, Modig L, Nordin M, et al. Traffic-Related Air Pollution and Dementia Incidence in Northern Sweden: A Longitudinal Study. *Environ Health Perspect*. 2016; 124(3):306–12. <https://doi.org/10.1289/ehp.1408322> PMID: 26305859
87. Chen H, Kwong JC, Copes R, Tu K, Villeneuve PJ, Van Donkelaar A, et al. Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: a population-based cohort study. *Lancet*. 2017; 389(10070):718–26. [https://doi.org/10.1016/S0140-6736\(16\)32399-6](https://doi.org/10.1016/S0140-6736(16)32399-6) PMID: 28063597
88. Zanobetti A, Dominici F, Wang Y, Schwartz JD. A national case-crossover analysis of the short-term effect of PM 2.5 on hospitalizations and mortality in subjects with diabetes and neurological disorders. *Environ Health*. 2014; 13(1):1–11.
89. Kioumourtoglou M-A, Schwartz JD, Weisskopf MG, Melly SJ, Wang Y, Dominici F, et al. 5 exposure and neurological hospital admissions in the northeastern United States. *Environ Health Perspect*. 2016; 124(1):23–9.
90. Carey IM, Anderson HR, Atkinson RW, Beevers SD, Cook DG, Strachan DP, et al. Are noise and air pollution related to the incidence of dementia? A cohort study in London, England. *BMJ Open*. 2018; 8(9):e022404. <https://doi.org/10.1136/bmjopen-2018-022404> PMID: 30206085
91. Chang K-H, Chang M-Y, Muo C-H, Wu T-N, Chen C-Y, Kao C-H. Increased risk of dementia in patients exposed to nitrogen dioxide and carbon monoxide: a population-based retrospective cohort study. *PLoS ONE*. 2014; 9(8):e103078. <https://doi.org/10.1371/journal.pone.0103078> PMID: 25115939
92. Chen H, Kwong JC, Copes R, Hystad P, van Donkelaar A, Tu K, et al. Exposure to ambient air pollution and the incidence of dementia: a population-based cohort study. *Environ Int*. 2017; 108:271–7. <https://doi.org/10.1016/j.envint.2017.08.020> PMID: 28917207
93. Calderón-Garcidueñas L, Reynoso-Robles R, Vargas-Martínez J, Gómez-Maqueo-Chew A, Pérez-Guillé B, Mukherjee PS, et al. Prefrontal white matter pathology in air pollution exposed Mexico City young urbanites and their potential impact on neurovascular unit dysfunction and the development of Alzheimer's disease. *Environ Res*. 2016; 146:404–17. <https://doi.org/10.1016/j.envres.2015.12.031> PMID: 26829765

94. Calderón-Garcidueñas L., Mora-Tiscareño A., Franco-Lira M., Zhu H., Lu Z., Solorio E., . . . & D'Angiulli A. (2015). Decreases in short term memory, IQ, and altered brain metabolic ratios in urban apolipoprotein $\epsilon 4$ children exposed to air pollution. *Journal of Alzheimer's Disease*, 45(3), 757–770.
95. Sengillo JD, Winkler EA, Walker CT, Sullivan JS, Johnson M, Zlokovic BV. Deficiency in Mural Vascular Cells Coincides with Blood–Brain Barrier Disruption in Alzheimer's Disease. *Brain Pathol.* 2013; 23(3):303–10. <https://doi.org/10.1111/bpa.12004> PMID: 23126372
96. Bell R. D., Winkler E. A., Singh I., Sagare A. P., Deane R., Wu Z., . . . & Zlokovic B. V. (2012). Apolipoprotein E controls cerebrovascular integrity via cyclophilin A. *Nature*, 485(7399), 512–516.
97. Koutsilieris E, Scheller C, Tribl F, Riederer P. Degeneration of neuronal cells due to oxidative stress—microglial contribution. *Parkinsonism Relat Disord.* 2002; 8(6):401–6. [https://doi.org/10.1016/s1353-8020\(02\)00021-4](https://doi.org/10.1016/s1353-8020(02)00021-4) PMID: 12217627
98. Roqué PJ, Dao K, Costa LG. Microglia mediate diesel exhaust particle-induced cerebellar neuronal toxicity through neuroinflammatory mechanisms. *Neurotoxicology.* 2016; 56:204–14. <https://doi.org/10.1016/j.neuro.2016.08.006> PMID: 27543421
99. Giordano G, Tait L, Furlong CE, Cole TB, Kavanagh TJ, Costa LG. Gender differences in brain susceptibility to oxidative stress are mediated by levels of paraoxonase-2 expression. *Free Radic Biol Med.* 2013; 58:98–108. <https://doi.org/10.1016/j.freeradbiomed.2013.01.019> PMID: 23376469
100. Costa LG, de Laat R, Dao K, Pellacani C, Cole TB, Furlong CE. Paraoxonase-2 (PON2) in brain and its potential role in neuroprotection. *Neurotoxicology.* 2014; 43:3–9. <https://doi.org/10.1016/j.neuro.2013.08.011> PMID: 24012887
101. Borrás C, Sastre J, García-Sala D, Lloret A, Pallardó FV, Viña J. Mitochondria from females exhibit higher antioxidant gene expression and lower oxidative damage than males. *Free Radic Biol Med.* 2003; 34(5):546–52. [https://doi.org/10.1016/s0891-5849\(02\)01356-4](https://doi.org/10.1016/s0891-5849(02)01356-4) PMID: 12614843
102. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. A default mode of brain function. *Proc Natl Acad Sci U S A.* 2001; 98(2):676–82. <https://doi.org/10.1073/pnas.98.2.676> PMID: 11209064
103. Liguori C, Chiaravalloti A, Sancesario G, Stefani A, Sancesario GM, Mercuri NB, et al. Cerebrospinal fluid lactate levels and brain [18F] FDG PET hypometabolism within the default mode network in Alzheimer's disease. *Eur J Nucl Med Mol Imaging.* 2016; 43(11):2040–49. <https://doi.org/10.1007/s00259-016-3417-2> PMID: 27221635
104. Malkov A, Popova I, Ivanov A, Jang S-S, Yoon SY, Osypov A, et al. A β initiates brain hypometabolism, network dysfunction and behavioral abnormalities via NOX2-induced oxidative stress in mice. *Commun Biol.* 2021; 4(1):1–12.
105. Daftary SS, Gore AC. IGF-1 in the brain as a regulator of reproductive neuroendocrine function. *Exp Biol Med.* 2005; 230(5):292–306. <https://doi.org/10.1177/153537020523000503> PMID: 15855296
106. Åberg MAI, Åberg ND, Hedbäcker H, Oscarsson J, Eriksson PS. Peripheral Infusion of IGF-I Selectively Induces Neurogenesis in the Adult Rat Hippocampus. *J Neurosci.* 2000; 20(8):2896–903. <https://doi.org/10.1523/JNEUROSCI.20-08-02896.2000> PMID: 10751442
107. Trejo JL, Carro E, Torres-Alemán I. Circulating Insulin-Like Growth Factor I Mediates Exercise-Induced Increases in the Number of New Neurons in the Adult Hippocampus. *J Neurosci.* 2001; 21(5):1628–34. <https://doi.org/10.1523/JNEUROSCI.21-05-01628.2001> PMID: 11222653
108. Altman J, Das GD. Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *J Comp Neurol.* 1965; 124(3):319–35. <https://doi.org/10.1002/cne.901240303> PMID: 5861717
109. Kaplan MS, Hinds JW. Neurogenesis in the adult rat: electron microscopic analysis of light radioautographs. *Science.* 1977; 197(4308):1092–94. <https://doi.org/10.1126/science.887941> PMID: 887941
110. Kornack DR, Rakic P. Continuation of neurogenesis in the hippocampus of the adult macaque monkey. *Proc Natl Acad Sci U S A.* 1999; 96(10):5768–73. <https://doi.org/10.1073/pnas.96.10.5768> PMID: 10318959
111. Hagenaaers SP, Harris SE, Davies G, Hill WD, Liewald DCM, Ritchie SJ, et al. Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N = 112 151) and 24 GWAS consortia. *Mol Psychiatry.* 2016; 21(11):1624–32.
112. Lyall DM, Ward J, Ritchie SJ, Davies G, Cullen B, Celis C, et al. Alzheimer disease genetic risk factor APOE $\epsilon 4$ and cognitive abilities in 111,739 UK Biobank participants. *Age Ageing.* 2016; 45(4):511–17. <https://doi.org/10.1093/ageing/afw068> PMID: 27103599
113. Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaaers SP, Harris SE, et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N = 112 151). *Mol Psychiatry.* 2016; 21(6):758–67.

114. Tank R, Ward J, Flegal KE, Smith DJ, Bailey ME, Cavanagh J, et al. Association between polygenic risk for Alzheimer's disease, brain structure and cognitive abilities in UK Biobank. *Neuropsychopharmacology*. 2022; 47(2):564–69. <https://doi.org/10.1038/s41386-021-01190-4> PMID: 34621014
115. Bartzokis G. Alzheimer's disease as homeostatic responses to age-related myelin breakdown. *Neurobiol Aging*. 2011; 32(8):1341–71. <https://doi.org/10.1016/j.neurobiolaging.2009.08.007> PMID: 19775776
116. Tavares RM, Mendelsohn A, Grossman Y, Williams CH, Shapiro M, Trope Y, et al. A map for social navigation in the human brain. *Neuron*. 2015; 87(1):231–43. <https://doi.org/10.1016/j.neuron.2015.06.011> PMID: 26139376
117. Morrish N, Medina-Lara A. Does unemployment lead to greater levels of loneliness? A systematic review. *Soc Sci Med*. 2021;114339. <https://doi.org/10.1016/j.socscimed.2021.114339> PMID: 34455335
118. Penninkilampi R, Casey A-N, Singh MF, Brodaty H. The Association between Social Engagement, Loneliness, and Risk of Dementia: A Systematic Review and Meta-Analysis. *J Alzheimers Dis*. 2018; 66:1619–33. <https://doi.org/10.3233/JAD-180439> PMID: 30452410
119. Kuiper JS, Zuidersma M, Voshaar RCO, Zuidema SU, van den Heuvel ER, Stolk RP, et al. Social relationships and risk of dementia: A systematic review and meta-analysis of longitudinal cohort studies. *Ageing Res Rev*. 2015; 22:39–57. <https://doi.org/10.1016/j.arr.2015.04.006> PMID: 25956016
120. Bassuk SS, Glass TA, Berkman LF. Social disengagement and incident cognitive decline in community-dwelling elderly persons. *Ann Intern Med*. 1999; 131(3):165–73. <https://doi.org/10.7326/0003-4819-131-3-199908030-00002> PMID: 10428732
121. Scarmeas N, Levy G, Tang M-X, Manly J, Stern Y. Influence of leisure activity on the incidence of Alzheimer's disease. *Neurology*. 2001; 57(12):2236–42. <https://doi.org/10.1212/wnl.57.12.2236> PMID: 11756603
122. Hostinar CE, Gunnar MR. Social support can buffer against stress and shape brain activity. *AJOB Neurosci*. 2015; 6(3):34–42. <https://doi.org/10.1080/21507740.2015.1047054> PMID: 26478822
123. Cacioppo JT, Cacioppo S, Capitanio JP, Cole SW. The Neuroendocrinology of Social Isolation. *Annu Rev Psychol*. 2015; 66(1):733–67. <https://doi.org/10.1146/annurev-psych-010814-015240> PMID: 25148851
124. Bzdok D, Dunbar RI. The neurobiology of social distance. *Trends Cogn Sci*. 2020; 24(9):717–33. <https://doi.org/10.1016/j.tics.2020.05.016> PMID: 32561254
125. Kong X, Wei D, Li W, Cun L, Xue S, Zhang Q, et al. Neuroticism and extraversion mediate the association between loneliness and the dorsolateral prefrontal cortex. *Exp Brain Res*. 2015; 233(1):157–64. <https://doi.org/10.1007/s00221-014-4097-4> PMID: 25234401
126. Lelic-Tosevski D, Vukovic O, Stepanovic J. Stress and personality. *Psychiatriki*. 2011; 22(4):290–7. PMID: 22271841
127. Swanson LW, Cowan WM. Hippocampo-Hypothalamic Connections: Origin in Subicular Cortex, Not Ammon's Horn. *Science*. 1975; 189(4199):303–4. <https://doi.org/10.1126/science.49928> PMID: 49928
128. Wilson RS, Schneider JA, Boyle PA, Arnold SE, Tang Y, Bennett DA. Chronic distress and incidence of mild cognitive impairment. *Neurology*. 2007; 68(24):2085–92. <https://doi.org/10.1212/01.wnl.0000264930.97061.82> PMID: 17562829
129. Wilson R, Evans D, Bienias J, De Leon CM, Schneider J, Bennett D. Proneness to psychological distress is associated with risk of Alzheimer's disease. *Neurology*. 2003; 61(11):1479–85. <https://doi.org/10.1212/01.wnl.0000096167.56734.59> PMID: 14663028
130. Wilson R, Barnes L, Bennett D, Li Y, Bienias J, De Leon CM, et al. Proneness to psychological distress and risk of Alzheimer disease in a biracial community. *Neurology*. 2005; 64(2):380–2. <https://doi.org/10.1212/01.WNL.0000149525.53525.E7> PMID: 15668449
131. Wilson RS, Arnold SE, Schneider JA, Kelly JF, Tang Y, Bennett DA. Chronic psychological distress and risk of Alzheimer's disease in old age. *Neuroepidemiology*. 2006; 27(3):143–53. <https://doi.org/10.1159/000095761> PMID: 16974109
132. Fiocco A, Nair N, Schwartz G, Kin FNY, Joobor R, Poirier J, et al. Influence of genetic polymorphisms in the apolipoprotein (APOE) and the butyrylcholinesterase (BCHE) gene on stress markers in older adults: a 3-year study. *Neurobiol Aging*. 2009; 30(6):1001–5. <https://doi.org/10.1016/j.neurobiolaging.2007.09.008> PMID: 17996334
133. Levy BR, Slade MD, Pietrzak RH, Ferrucci L. When Culture Influences Genes: Positive Age Beliefs Amplify the Cognitive-Aging Benefit of APOE ε2. *J Gerontol B Psychol Sci Soc Sci*. 2020; 75(8):e198–203.

134. Alfaro-Almagro F, Jenkinson M, Bangerter NK, Andersson JL, Griffanti L, Douaud G, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage*. 2018; 166:400–24. <https://doi.org/10.1016/j.neuroimage.2017.10.034> PMID: 29079522
135. Zhang Q, Sidorenko J, Couvy-Duchesne B, Marioni RE, Wright MJ, Goate AM, et al. Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture. *Nat Commun*. 2020; 11(1):1–11.
136. Liu JZ, Erlich Y, Pickrell JK. Case-control association mapping by proxy using family history of disease. *Nat Genet*. 2017; 49(3):325–31. <https://doi.org/10.1038/ng.3766> PMID: 28092683
137. Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, et al. GWAS on family history of Alzheimer's disease. *Transl Psychiatry*. 2018; 8(1):1–7.
138. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat Genet*. 2019; 51(3):414–30.
139. Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013; 45(12):1452–58. <https://doi.org/10.1038/ng.2802> PMID: 24162737
140. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci*. 2016; 19(11):1523. <https://doi.org/10.1038/nn.4393> PMID: 27643430
141. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002; 17(3):143–55. <https://doi.org/10.1002/hbm.10062> PMID: 12391568
142. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal*. 2001; 5(2):143–56. [https://doi.org/10.1016/s1361-8415\(01\)00036-6](https://doi.org/10.1016/s1361-8415(01)00036-6) PMID: 11516708
143. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 2002; 17(2):825–41. [https://doi.org/10.1016/s1053-8119\(02\)91132-8](https://doi.org/10.1016/s1053-8119(02)91132-8) PMID: 12377157
144. Andersson JL, Jenkinson M, Smith S. Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2. FMRIB Analysis Group of the University of Oxford. 2007; 2(1):e21.
145. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging*. 2001; 20(1):45–57. <https://doi.org/10.1109/42.906424> PMID: 11293691
146. Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, et al. Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis. *Neuroimage*. 2002; 17(1):479–89. <https://doi.org/10.1006/nimg.2002.1040> PMID: 12482100
147. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb Cortex*. 2017;1–20.
148. Spreng RN, Dimas E, Mwilambwe-Tshilobo L, Dagher A, Koellinger P, Nave G, et al. The default network of the human brain is associated with perceived social isolation. *Nat Commun*. 2020; 11(1):1–11.
149. Schurz M, Uddin L, Kanske P, Lamm C, Sallet J, Bernhardt B, et al. Variability in brain structure and function reflects lack of peer support. *Cereb Cortex*. 2021. <https://doi.org/10.1093/cercor/bhab109> PMID: 33982758
150. Witten DM, Tibshirani R, Hastie T. A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics*. 2009:kxp008. <https://doi.org/10.1093/biostatistics/kxp008> PMID: 19377034
151. Bzdok D, Nichols TE, Smith SM. Towards algorithmic analytics for large-scale datasets. *Nat Mach Intell*. 2019; 1(7):296–306. <https://doi.org/10.1038/s42256-019-0069-5> PMID: 31701088
152. Zajner C, Spreng RN, Bzdok D. Lacking social support is associated with structural divergences in hippocampus-default network co-variation patterns. *Soc Cogn Affect Neurosci*. 2022; 17(9):802–18. <https://doi.org/10.1093/scan/nsac006> PMID: 35086149
153. Efron B, Tibshirani RJ. An introduction to the bootstrap. CRC Press; 1994.
154. Millard LAC, Davies NM, Gaunt TR, Davey Smith G, Tilling K. Software Application Profile: PHESANT: a tool for performing automated phenome scans in UK Biobank. *Int J Epidemiol*. 2018; 47(1):29–35. <https://doi.org/10.1093/ije/dyx204> PMID: 29040602
155. Wolpert DH. Stacked generalization. *Neural Netw*. 1992; 5(2):241–59.
156. Breiman L. Stacked regressions. *Mach Learn*. 1996; 24(1):49–64.
157. Dong Q, Zhang W, Wu J, Li B, Schron EH, McMahon T, et al. Applying surface-based hippocampal morphometry to study APOE-E4 allele dose effects in cognitively unimpaired subjects. *Neuroimage Clin*. 2019; 22:101744. <https://doi.org/10.1016/j.nicl.2019.101744> PMID: 30852398

158. Hostage CA, Roy Choudhury K, Doraiswamy PM, Petrella JR. Initiative AsDN. Dissecting the gene dose-effects of the APOE ϵ 4 and ϵ 2 alleles on hippocampal volumes in aging and Alzheimer's disease. *PLoS ONE*. 2013; 8(2):e54483.
159. West HL, Rebeck GW, Hyman BT. Frequency of the apolipoprotein E ϵ 2 allele is diminished in sporadic Alzheimer disease. *Neurosci Lett*. 1994; 175(1–2):46–8.
160. Benjamin R, Leake A, McArthur F, Ince P, Candy J, Edwardson J, et al. Protective effect of apoE epsilon 2 in Alzheimer's disease. *Lancet (London, England)*. 1994; 344(8920):473. [https://doi.org/10.1016/S0140-6736\(94\)91804-x](https://doi.org/10.1016/S0140-6736(94)91804-x) PMID: 7914580
161. Lecarpentier J, Silvestri V, Kuchenbaecker KB, Barrowdale D, Dennis J, McGuffog L, et al. Prediction of Breast and Prostate Cancer Risks in Male BRCA1 and BRCA2 Mutation Carriers Using Polygenic Risk Scores. *J Clin Oncol*. 2017; 35(20):2240–50. <https://doi.org/10.1200/JCO.2016.69.4935> PMID: 28448241
162. Fan BJ, Bailey JC, Igo RP Jr, Kang JH, Boumenna T, Brilliant MH, et al. Association of a Primary Open-Angle Glaucoma Genetic Risk Score With Earlier Age at Diagnosis. *JAMA Ophthalmology*. 2019; 137(10):1190–94. <https://doi.org/10.1001/jamaophthalmol.2019.3109> PMID: 31436842
163. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Methodol*. 1995; 57(1):289–300.
164. Sha Z, Schijven D, Carrion-Castillo A, Joliot M, Mazoyer B, Fisher SE, et al. The genetic architecture of structural left-right asymmetry of the human brain. *Nat Hum Behav*. 2021; 5(9):1226–39. <https://doi.org/10.1038/s41562-021-01069-w> PMID: 33723403
165. Raizada RD, Richards TL, Meltzoff A, Kuhl PK. Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *Neuroimage*. 2008; 40(3):1392–401. <https://doi.org/10.1016/j.neuroimage.2008.01.021> PMID: 18308588
166. Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*. 2002; 15(4):870–78. <https://doi.org/10.1006/nimg.2001.1037> PMID: 11906227
167. Ward JH Jr. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc*. 1963; 58(301):236–44.
168. Pearson KLIII. On lines and planes of closest fit to systems of points in space. *Lond Edin Dubl Phil Mag J Sci*. 1901; 2(11):559–572.
169. Gerbrands JJ. On the relationships between SVD, KLT and PCA. *Pattern Recognit*. 1981; 14(1–6):375–81.
170. Tremblay-Mercier J, Madjar C, Das S, Binette AP, Dyke SO, Étienne P, et al. Open science datasets from PREVENT-AD, a longitudinal cohort of pre-symptomatic Alzheimer's disease. *Neuroimage Clin*. 2021; 31:102733. <https://doi.org/10.1016/j.nicl.2021.102733> PMID: 34192666